

Praktikum primijenjene matematike II. Metoda konačnih
elemenata

Mladen Jurak

23. prosinca 2006.

Sadržaj

1	Uvod. Jednodimenzionalan primjer	2
1.1	Varijacijska formulacija	2
1.2	Egzistencija rješenja varijacijske jednadžbe	7
1.3	Prostori Soboljeva	9
1.4	Varijacijska aproksimacija	14
1.5	Ocjena greške aproksimacije	16
1.6	Konstrukcija konačnodimenzionalnog prostora S	17
1.7	Argument dualnosti	22
1.8	Uniformna ocjena	23
1.9	Algoritamski aspekti metode konačnih elemenata	24
2	Varijacijska formulacija eliptičkih rubnih zadaća	38
2.1	Uvodni primjer	39
2.2	Prostori $L^2(\Omega)$, $H^1(\Omega)$ i $H_0^1(\Omega)$	43
2.3	Homogena Dirichletova zadaća	46
2.4	Nehomogena Dirichletova zadaća	50
2.5	Neumannova i mješovita zadaća	52
2.6	Nesimetrične zadaće	56
2.7	Spektralna zadaća	59
2.8	Nelinearne zadaće	64
2.9	Varijacijska aproksimacija	67
2.10	Dodatak. Lax-Milgramova lema	72
2.11	Zadaci	84
3	Konstrukcija prostora konačnih elemenata u $H^1(\Omega)$	86
3.1	Osnovna svojstva prostora konačnih elemenata	86
3.2	Primjer: \mathbb{P}_1 elementi na trokutima	88
3.3	Simplicijalni elementi	90
3.4	Interpolacijski operatori	96
3.5	Afina familija elemenata	97
3.6	Zadaci	98

4 Greška interpolacije	99
4.1 Lokalne ocjene	99
4.2 Globalne ocjene	105
4.3 Aubin-Nitscheova lema	107
4.4 O regularnosti točnog rješenja	108
4.5 Adaptivne metode	112
4.6 Zadaci	113
4.7 Dodatak 1. Dokaz Deny-Lionsove leme	114
4.8 Dodatak 2. Inverzne ocjene	116
5 Numerička integracija	121
5.1 Primjena formula numeričke integracije	121
5.2 Dovoljan uvjet uniformne V_h -eliptičnosti	125
5.3 Konzistentnost. Bramble-Hilbertova lema	127
6 Princip maksimuma	132
6.1 Slabi i jaki princip maksimuma za eliptičke jednadžbe	132
6.2 Slabi princip maksimuma za eliptičke jednadžbe u divergentnoj formi	134
6.3 Diskretan princip maksimuma	138
6.3.1 Primjer: metoda konačnih diferencija	138
6.3.2 Ocjena greške metode konačnih diferencija	140
6.3.3 Ireducibilne matrice	141
6.3.4 Diskretan princip maksimuma	146
6.4 Primjena na metodu konačnih elemenata	153
7 Adaptivne metode i a posteriori ocjena greške	157
7.1 Adaptivni konačni elementi	157
7.2 A posteriori procjenitelj greške	158
7.3 Indikator za Laplaceovu jednadžbu	160
8 Paraboličke rubne zadaće	164
8.1 Primjeri inicijalno-rubnih zadaća	164
8.2 Varijacijska formulacija	165
8.3 Korektnost zadaće (8.8)	167
8.4 Diskretizacija	169
8.5 Korektnost poludiskretizirane zadaće	170
8.6 Ocjena greške za poludiskretiziranu zadaću	171
A Funkcijski prostori	175
A.1 Diferencijalni operatori	175
A.1.1 Multiindeksi	175
A.1.2 Taylorov polinom	176

A.1.3	Diferencijalni operatori	177
A.2	Prostori glatkih funkcija	180
A.3	L^p prostori	181
A.4	$W^{1,p}$ prostori	184
A.4.1	Singulariteti $W^{1,p}$ funkcija	186
A.4.2	Regularnost granice	187
A.5	Zadaci	189

1

Uvod. Jednodimenzionalan primjer

U uvodnom poglavlju želimo objasniti metodu konačnih elemenata na vrlo jednostavnoj jednodimenzionalnoj rubnoj zadaći. Jednostavnost zadaće omogućava da se bez tehničkih poteškoća izlože osnovni elementi matematičke teorije. Nedostatak je u tome što nije moguće prezentirati sve elemente važne za implementaciju metode u punoj općenitosti. Pored opisa metode cilj nam je motivirati uvođenje prostora Soboljeva.

Neka je zadana funkcija $f: (0, 1) \rightarrow \mathbb{R}$. Promatramo sljedeću zadaću: treba naći funkciju $u: [0, 1] \rightarrow \mathbb{R}$ koja zadovoljava

$$\begin{cases} -\frac{d^2u}{dx^2} = f & \text{na } (0, 1) \\ u(0) = 0, \quad u'(1) = 0. \end{cases} \quad (1.1)$$

Iako rješenje zadaće (1.1) možemo izraziti jednostavnom formulom korisno je postaviti pitanje precizne definicije rješenja.

1.1 Varijacijska formulacija

Da bismo precizirali smisao rješenja moramo uvesti neke funkcijske prostore. Svi prostori koje ćemo uvesti imaju strukturu linearnog prostora u odnosu na uobičajeno zbrajanje funkcija i množenja skalarom, tako da to nadalje nećemo naglašavati.

Neka $\Omega \subset \mathbb{R}$ otvoren skup, a $\bar{\Omega}$ njegov zatvarač. Tada definiramo:

$C(\Omega)$ = prostor svih neprekidnih funkcija $u: \Omega \rightarrow \mathbb{R}$

$C(\bar{\Omega})$ = prostor svih uniformno neprekidnih funkcija $u: \Omega \rightarrow \mathbb{R}$

$C^k(\Omega)$ = prostor svih k -puta neprekidno derivabilnih funkcija $u: \Omega \rightarrow \mathbb{R}$

$C^k(\bar{\Omega})$ = prostor funkcija $u: \Omega \rightarrow \mathbb{R}$ za koje je $\frac{d^j u}{dx^j} \in C(\bar{\Omega})$ za $j = 0, 1, \dots, k$.

Za zadano $f \in C(0, 1)$ funkciju $u \in C^2(0, 1) \cap C^1([0, 1])$ nazivamo klasično rješenje zadaje (1.1) ako zadovoljava rubne uvjete (1.1)₂, te ako diferencijalnu jednadžbu (1.1)₁ zadovoljava u svakoj točki $x \in (0, 1)$.

U fizikalnim promjenama funkcija f ne mora nužno biti neprekidna. Na primjer, moguće je imati na desnoj strani funkciju

$$f(x) = \begin{cases} 0 & \text{za } x < \frac{1}{2} \\ 1 & \text{za } x \geq \frac{1}{2}. \end{cases}$$

Što je rješenje zadaje (1.1) u tom slučaju? Ono evidentno nije funkcija iz $C^2(0, 1)$ već rješenje treba tražiti u prostoru funkcija koje imaju po dijelovima neprekidnu drugu derivaciju. U točki prekida desne strane, $x = 1/2$, druga derivacija ne mora biti definirana, i diferencijalna jednadžba u toj točki nije zadovoljena.

Ako idemo korak dalje, postavlja se pitanje koliko je moguće smanjiti glatkoću desne strane f , a da još uvijek možemo na smislen način definirati rješenje problema (1.1)? Odgovor je u varijacijskoj formulaciji rubne zadaća.

Neka je u klasično rješenje zadaje (1.1). Odaberimo funkciju $v \in C^1([0, 1])$ takvu da je $v(0) = 0$ i s njom pomnožimo jednadžbu (1.1)₁. Nakon integracije dobivamo

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

Parcijalnom integracijom u prvom integralu dobivamo:

$$-u'(1)v(1) + u'(0)v(0) + \int_0^1 u'(x)v'(x) dx = \int_0^1 f(x)v(x) dx.$$

Zbog $v(0) = 0$ i $u'(1) = 0$ izlazi

$$\int_0^1 u'(x)v'(x) dx = \int_0^1 f(x)v(x) dx, \quad \forall v \in C^1([0, 1]) \text{ takvo da je } v(0) = 0.$$

Dobivena se jednadžba naziva **varijacijska** stoga što vrijedi za svako $v \in C^1([0, 1])$ takvo da je $v(0) = 0$.

Zapisat ćemo dobivenu varijacijsku zadaću u malo apstraktnijoj formi. Prvo uvodimo linearni prostor funkcija

$$V = \{\phi \in C^1([0, 1]): v(0) = 0\},$$

te funkcionalne

$$a: V \times V \rightarrow \mathbb{R}, \quad F: V \rightarrow \mathbb{R}$$

definirane formulama:

$$a(u, v) = \int_0^1 u'(x)v'(x) dx, \tag{1.2}$$

$$F(v) = \int_0^1 f(x)v(x) dx. \tag{1.3}$$

Evidentno je F linearan funkcional, odnosno zadovoljava

$$\forall \alpha, \beta \in \mathbb{R}, \forall u, v \in V, \quad F(\alpha u + \beta v) = \alpha F(u) + \beta F(v).$$

Funkcional $a(\cdot, \cdot)$ je **bilinearna forma** (ili bilinearni funkcional), što znači da je linearan funkcional u svakoj varijabli zasebno. Preciznije,

$$\begin{aligned} \forall \alpha, \beta \in \mathbb{R}, \forall u, v, w \in V, \quad a(\alpha u + \beta v, w) &= \alpha a(u, w) + \beta a(v, w), \\ \forall \alpha, \beta \in \mathbb{R}, \forall u, v, w \in V, \quad a(u, \alpha v + \beta w) &= \alpha a(u, v) + \beta a(u, w). \end{aligned}$$

Varijacijska formulacija zadatka (1.1) sada prima oblik

$$\begin{cases} \text{naći } u \in V \\ a(u, v) = F(v), \quad \forall v \in V. \end{cases} \quad (1.4)$$

Funkciju v iz varijacijske jednadžbe nazivamo **test funkcija**.

Lema 1.1 Neka je $f \in C([0, 1])$ i $u \in C^2([0, 1])$ rješenje zadatka (1.4). Tada je u klasično rješenje zadatka (1.1).

Dokaz. Zbog pretpostavljene glatkoće u varijacijskoj jednadžbi možemo izvršiti parcijalnu integraciju.

$$\int_0^1 [(u'(x)v(x))' - u''(x)v(x)] dx = \int_0^1 f(x)v(x) dx, \quad \forall v \in V. \quad (1.5)$$

Odaberemo li test funkciju $v \in V$ koja zadovoljava $v(0) = v(1) = 0$, dobivamo

$$\int_0^1 (u''(x) + f(x))v(x) dx = 0 \quad \forall v \in C^1([0, 1]), v(0) = v(1) = 0.$$

Funkcija $w = u'' + f$ je iz $C([0, 1])$. Kada bi bila različita od nule, onda bi morao postojati interval $(x_0, x_1) \subset (0, 1)$ na kome je strogo pozitivna ili strogo negativna. Uzmimo, b.s.o., da postoji jedan takav interval na kome je w strogo pozitivna i definirajmo test funkciju

$$v(x) = \begin{cases} (x - x_0)^2(x - x_1)^2 & \text{za } x_0 < x < x_1 \\ 0 & \text{inače.} \end{cases}$$

Kako je $v(0) = v(1) = 0$ dobivamo

$$\int_{x_0}^{x_1} w(x)v(x) dx = 0,$$

no budući da je podintegralna funkcija strogo pozitivna, dobili smo kontradikciju i zaključujemo da je nužno $w \equiv 0$. Time smo dokazali da je diferencijalna jednadžba zadovoljena u svakoj točki $x \in (0, 1)$. Vratimo se sada u (1.5) gdje uzimamo proizvoljnu funkciju $v \in V$ i dobivamo

$$u'(1)v(1) - u'(0)v(0) = \int_0^1 (u''(x) + f(x))v(x) dx = 0,$$

gdje zadnja jednakost slijedi iz do sada dokazanog. Kako je $v(0) = 0$, dovoljno je uzeti test funkciju koja zadovoljava $v(1) = 1$ da bismo zaključili da je $u'(1) = 0$; zbog $u \in V$, automatski je zadovoljeno $u(0) = 0$ i time su rubni uvjeti zadovoljeni. \square

Napomena 1.1 *Pokazali smo da varijacijska formulacija sadrži sve informacije o rubnoj zadaći kao i klasična. Pri tome se u varijacijskoj formulaciji javljaju samo derivacije prvog reda, što ju čini općenitijom od klasične.*

Napomena 1.2 *Dirchletov rubni uvjet, $u(0) = 0$, u varijacijskoj formulaciji je ugrađen u prostor V , pa se stoga naziva i esencijalni rubni uvjet. Neumannov rubni uvjet, $u'(1) = 0$, zadovoljen je implicitno samom varijacijskom jednadžbom, pa ga nazivamo i prirodni rubni uvjet.*

Napomena 1.3 *Korisno je uvesti pojam nosača funkcije: to je zatvarač skupa na kojem je funkcija različita od nule. Oznaka je $\text{supp}(u)$, i ako je (a, b) domena funkcije $u(x)$, onda je*

$$\text{supp}(u) = \overline{\{x \in (a, b) : u(x) \neq 0\}}.$$

Potez iznad skupa označava zatvarač. Nosač funkcije je stoga uvijek zatvoren skup i sadrži pojedine točke u kojima se funkcija poništava.

Funkcija ima kompaktan nosač ako je njen nosač kompaktan skup. (Da bi nosač bio kompaktan dovoljno je da je ograničen.) Prostori neprekidnih (neprekidno derivabilnih) funkcija s kompaktnim nosačem označavaju se slovom c : $C_c(a, b)$, $C_c^k(a, b)$ itd. Uočimo da ako funkcija definirana na otvorenom intervalu (a, b) i ima kompaktan nosač, onda ona postaje identički jednaka nuli na pozitivnoj udaljenosti od rubova intervala a i b .

Napomena 1.4 *U dokazu Leme 1.1 dokazali smo i osnovnu lemu varijacijskog računa koja kaže da ako funkcija $w \in C(a, b)$ zadovoljava*

$$\int_a^b w(x)v(x) dx = 0, \quad \forall v \in C_c^\infty(a, b),$$

onda je $w \equiv 0$. Ovdje je s $C_c^\infty(a, b)$ označen prostor svih beskonačno puta derivabilnih funkcija

$$C_c^\infty(a, b) = \bigcap_{k=1}^{\infty} C_c^k(a, b).$$

Lema 1.1 nam govori da je varijacijska zadaća (1.4) generalizacija rubne zadaće (1.1). Stoga rješenje zadaće (1.4) nazivamo **poopćeno** (ili **slabo**) rješenje zadaće (1.1). Slabo rješenje ima smisla za svaku funkciju f za koju je linearni funkcional (1.3) dobro definiran i neprekidan, što ćemo precizirati kasnije.

Varijacijska formulacija složenijih rubnih zadaća radi s na isti način, parcijalnom integracijom. Homogeni Dirichletovi rubni uvjeti se ugrađuju u prostor, a Neumannovi uvjeti su zadovoljeni implicitno.

Zadatak 1.1 *Nađite varijacijsku formulaciju ovih rubnih zadaća postavljenih na intervalu (a, b) :*

$$-u'' + cu = f, \quad u(a) = u(b) = 0, \quad (1.6)$$

$$-(ku')' + cu = f, \quad u(a) = 0, u'(b) + \alpha u(b) = 0, \quad (1.7)$$

$$-\varepsilon u'' + cu' + du = f, \quad u'(a) = u'(b) = 0. \quad (1.8)$$

Ovdje su c , d i f funkcije iz $C([a, b])$, $k \in C^1([a, b])$ i $\varepsilon > 0$. Pokažite da je svako dovoljno glatko rješenje varijacijske jednadžbe ujedno i klasično rješenje.

Rješenje. *ad) (1.6).* $V = \{u \in C^1([a, b]): v(a) = v(b) = 0\}$; Za svako $v \in V$,

$$\int_a^b (u'(x)v'(x) + c(x)u(x)v(x)) dx = \int_a^b f(x)v(x) dx;$$

ad) (1.7). $V = \{u \in C^1([a, b]): v(a) = 0\}$; Za svako $v \in V$,

$$\int_a^b (k(x)u'(x)v'(x) + c(x)u(x)v(x)) dx + \alpha k(b)u(b)v(b) = \int_a^b f(x)v(x) dx;$$

ad) (1.8). $V = C^1([a, b])$; Za svako $v \in V$,

$$\int_a^b (\varepsilon u'(x)v'(x) + c(x)u'(x)v(x) + d(x)u(x)v(x)) dx = \int_a^b f(x)v(x) dx.$$

Dirichletovi rubni uvjeti mogu biti nehomogeni (različiti od nule) i tada ih ne možemo ugraditi u prostor u kome biramo test funkcije jer bi time izgubio linearnost (postao bi afina mnogostrukost). Stoga se nehomogeni rubni uvjeti prvo homogeniziraju i na taj se način rubna zadaća svodi na rubnu zadaću s homogenim rubnim uvjetom.

Na primjer, u zadaći

$$-u'' + du = f, \quad u(0) = A, u(1) = B, \quad (1.9)$$

postavljenoj na $(0, 1)$, uveli bismo funkciju $w(x) = (B - A)x + A$, koja zadovoljava rubne uvjete. Traženu funkciju u napišemo kao $u(x) = U(x) + w(x)$, gdje sada funkcija U zadovoljava

$$-U'' + dU = f - dw, \quad U(0) = U(1) = 0.$$

Rubni uvjeti su time homogenizirani, a u jednadžbi se promijenila samo desna strana.

Zadatak 1.2 *Homogenizirajte rubne uvjete u sljedećim zadaćama postavljenim na (a, b) :*

$$-u'' + 3u = f, \quad u(a) = A, u(b) = B,$$

$$-(ku')' + cu = f, \quad u(a) = A, u(b) = B, (k(x) = 1 + x^2, c(x) = \cos(x))$$

$$-u'' + u' + u = f, \quad u(a) = A, u'(b) = B.$$

Nehomogeni Neumannov rubni uvjet nije potrebno homogenizirati.

Zadatak 1.3 *Formirajte varijacijsku zadaću za*

$$-u'' + u' + u = f, \quad x \in (a, b) \quad u'(a) = A, u'(b) = B.$$

Pokažite da je svako dovoljno glatko slabo rješenje ujedno i klasično rješenje.

Rješenje. $V = C^1([a, b])$; Za svako $v \in V$,

$$\int_a^b (\varepsilon u'(x)v'(x) + u'(x)v(x) + u(x)v(x)) \, dx = \int_a^b f(x)v(x) \, dx + Bv(b) - Av(a).$$

1.2 Egzistencija rješenja varijacijske jednadžbe

Da bismo vidjeli pod kojim uvjetima varijacijska jednadžba (1.4) ima jedinstveno rješenje trebamo problem postaviti u okviru funkcionalne analize.

Prije svega, uočimo da je prostor V beskonačnodimenzionalan linearan prostor. Da bismo mogli koristiti standardne rezultate funkcionalne analize moramo postaviti zahtjev da je V potpun normiran prostor. To znači da je na V definirana norma, koju ćemo označavati s $\|\cdot\|_V$, te da je u njemu svaki Cauchyjev niz konvergentan (potpunost). Funkcijski prostori koji su vezani uz linearne diferencijalne jednadžbe obično imaju strukturu potpunog unitarnog prostora. To je prostor u kojem je definiran skalarni produkt, kojeg označavamo s (\cdot, \cdot) , a pripadna norma je dana formulom $\|u\|_V = \sqrt{(u, u)}$. U toj normi prostor mora biti potpun.

Standardna terminologija funkcionalne analize je sljedeća: potpun normiran prostor naziva se Banachov prostor, a potpun unitaran prostor je Hilbertov prostor. Osnovne definicije i rezultati funkcionalne analize koji su nam potrebni dani su u Dodatku A.

Dualni prostor prostora V je prostor svih linearnih i neprekidnih funkcionala $F: V \rightarrow \mathbb{R}$. To je linearan normiran prostor koji označavamo V' , a normu definiramo formulom

$$\|F\|_{V'} = \sup\{|F(v)| : v \in V, \|v\|_V \leq 1\}.$$

On je potpun čak ako V to nije.

Uobičajeno je primjenu funkcionala $F \in V'$ na nekom elementu označavati oštrim zagradama

$$F(v) = \langle F, v \rangle_{V', V}.$$

Oznake prostora se ispuštaju kad je jasno o kojim se prostorima radi. Uočimo još da je po definiciji dualne norme

$$|\langle F, v \rangle_{V', V}| \leq \|F\|_{V'} \|v\|_V. \quad (1.10)$$

Varijacijska formulacija linearne rubne zadaće prirodno se postavlja u Hilbertovom prostoru. Na osnovu Rieszovog teorema reprezentacije svaki linearan i neprekidan funkcional nad Hilbertovim prostorom može se prikazati kao skalarni produkt s nekim elementom prostora.

O bilinearnoj formi ćemo napraviti dvije pretpostavke: ograničenost i koercitivnost.

Definicija 1.1 Kažemo da je bilinearna forma $a: V \times V \rightarrow \mathbb{R}$ ograničena ako postoji konstanta M , takva da vrijedi

$$\forall u, v \in V, \quad |a(u, v)| \leq M \|u\|_V \|v\|_V. \quad (1.11)$$

Definicija 1.2 Kažemo da je bilinearna forma $a: V \times V \rightarrow \mathbb{R}$ koercitivna ako postoji konstanta $\alpha > 0$ takva da je

$$\forall v \in V, \quad \alpha \|v\|_V^2 \leq a(v, v). \quad (1.12)$$

Kasnije ćemo dokazati ovaj teorem.

Teorem 1.1 (Lax-Milgramova lema) Neka je V Hilbertov prostor i $a: V \times V \rightarrow \mathbb{R}$ bilinearna, ograničena i koercitivna forma. Tada za svako $F \in V'$ problem (1.4) ima jedinstveno rješenje $u \in V$.

Lax-Milgramova lema nam kaže sljedeće: ako želimo dokazati da varijacijska zadaća (1.4) ima jedinstveno rješenje dovoljno je provjeriti da su zadovoljeni sljedeći uvjeti:

1. Prostor V je Hilbertov;
2. $a: V \times V \rightarrow \mathbb{R}$ je bilinearna forma;
3. $F: V \rightarrow \mathbb{R}$ linearan funkcional;
4. $a: V \times V \rightarrow \mathbb{R}$ zadovoljava uvjet ograničenosti (1.11);
5. $a: V \times V \rightarrow \mathbb{R}$ zadovoljava uvjet koercitivnosti (1.12);
6. $F: V \rightarrow \mathbb{R}$ zadovoljava uvjet neprekidnosti (1.10).

Napomena 1.5 * Ovdje želimo pokazati da su uvjeti Lax-Milgramove leme prirodni zahtjevi (s matematičke strane gledano) koje treba postaviti na elemente varijacijske jednadžbe (1.4).

Bilinearna forma $a(\cdot, \cdot)$ definira jedinstven linearan operator $A: V \rightarrow V'$ pomoću formule

$$\forall u, v \in V, \quad \langle Au, v \rangle_{V', V} = a(u, v).$$

Stoga, zadaću (1.4) možemo zapisati u obliku operatorske jednadžbe

$$Au = F. \quad (1.13)$$

Problem (1.13) je dobro postavljen ili korektan ako ima jedinstveno rješenje u koje neprekidno ovisi o desnoj strani F . To znači da operator A mora biti bijekcija, a inverz A^{-1} neprekidan operator. Prema teoremu o otvorenom preslikavanju (vidi S. Kurepa [3], str. 389) neprekidnost bijektivnog operatora A povlači neprekidnost inverznog operatora, pa stoga dolazimo do sljedećeg zaključka: Problem (1.13) je dobro postavljen ako je operator A neprekidna bijekcija.

Prostor svih neprekidnih linearnih operatora $A: V \rightarrow V'$ označava se $\mathcal{L}(V, V')$. To je linearan prostor u koji se uvodi operatorska norma

$$\|A\| = \sup\{\|Av\|_{V'} : v \in V, \|v\|_V \leq 1\}.$$

Ako je V potpun, onda je i $\mathcal{L}(V, V')$ potpun prostor.

Ograničenost (=neprekidnost) linearnog operatora A bit će osigurana ako imamo ograničenost bilinearne forme. Iz (1.11) se lako pokazuje da za pripadni operator A vrijedi $\|A\| \leq M$. Nadalje, potrebno je osigurati injektivnost operatora A . Jedan prirodan i jednostavan uvjet na formu $a(\cdot, \cdot)$, koji osigurava injektivnost, je uvjet koercitivnosti. Iz (1.12) slijedi

$$\alpha \|v\|_V^2 \leq a(v, v) = \langle Av, v \rangle_{V', V} \leq \|Av\|_{V'} \|v\|_V,$$

što povlači da za svako $v \in V$ vrijedi $\alpha \|v\|_V \leq \|Av\|_{V'}$. Iz te nejednakosti lako slijedi injektivnost operatora A i na isti način injektivnost adjungiranog operatora, te zatvorenost slike operatora A . Tada prema Banachovom teoremu o zatvorenoj slici dobivamo surjektivnost operatora A (za detalje vidi H. Brezis [2], II.7). Time dolazimo do egzistencije i jedinstvenosti rješenja jednadžbe (1.13).

1.3 Prostori Soboljeva

Vratimo se sada našem polaznom primjeru i provjerimo možemo li zadovoljiti gornje zahtjeve. U prostor $V = \{\phi \in C^1([0, 1]): \phi(0) = 0\}$ prirodno je uvesti normu

$$\|\phi\| = \sup_{0 \leq x \leq 1} |\phi(x)| + \sup_{0 \leq x \leq 1} |\phi'(x)|,$$

uz koji je on potpun. Zatim je lako pokazati da je uz tu normu bilinearna forma $a(\cdot, \cdot)$ ograničen:

$$\forall u, v \in V, \quad |a(u, v)| \leq \|u\| \|v\|.$$

Nasuprot tome, ne postoji konstanta α takva da je

$$a(u, u) = \int_0^1 u'(x)^2 dx \geq \alpha \|u\|^2 \geq \alpha \sup_{0 \leq x \leq 1} |u'(x)|^2.$$

Zadatak 1.4 Konstruirajte niz funkcija (u_n) iz prostora V za koji vrijedi

$$\int_0^1 u_n'(x)^2 dx \rightarrow 0 \quad \text{kada } n \rightarrow \infty,$$

$$\sup_{0 \leq x \leq 1} |u_n'(x)| = 1, \quad \forall n \in \mathbb{N}.$$

Problem koji se ovdje javlja je u tome što je prirodna norma prostora V suviše jaka. Varijacijska formulacija nam diktira normu integralnog oblika.

Uvedimo prostor

$$L^2(a, b) = \{\phi: (a, b) \rightarrow \mathbb{R}: \int_a^b \phi^2(x) dx < +\infty\}.$$

To je prostor svih kvadratno integrabilnih funkcija i u njemu dobro definiran skalarni produkt

$$(\phi, \psi)_{L^2(a,b)} = \int_a^b \phi(x)\psi(x) dx, \quad (1.14)$$

te pripadna norma

$$\|\phi\|_{L^2(a,b)} = \left(\int_a^b \phi^2(x) dx \right)^{1/2}. \quad (1.15)$$

Prostor $L^2(a, b)$ s ovom normom je Hilbertov.

Napomena 1.6 * Teorija $L^2(a, b)$ -prostora se oslanja na Lebesgueovu teoriju integracije. Formulom (1.15) definirana je norma na prostoru svih klasa ekvivalencije skoro svuda jednakih funkcija (u odnosu na Lebesgueovu mjeru). Prema tome, elementi $L^2(a, b)$ -prostora nisu funkcije nego klase ekvivalencija skoro svuda jednakih funkcija. Ta se tehnička komplikacija može izbjeći tako da se $L^2(a, b)$ promatra kao prostor funkcija, a da se uzima da su dvije funkcije jednake ako su jednake svuda osim na skupu (Lebesgueove) mjere nula. Uz takav dogovor ne možemo govoriti o vrijednosti funkcije u točki ili na bilo kojem skupu mjere nula.

Iz činjenice da je formulom (1.14) dobro definiran skalarni produkt slijedi:

Lema 1.2 (Schwarz–Cauchyjeva nejednakost) Za svake dvije funkcije $\phi, \psi \in L^2(a, b)$ vrijedi

$$\left| \int_a^b \phi(x)\psi(x) dx \right| \leq \left(\int_a^b \phi^2(x) dx \right)^{1/2} \left(\int_a^b \psi^2(x) dx \right)^{1/2}.$$

Za dokaz vidi [3].

Prostor $L^2(a, b)$ će biti zamjena za $C([a, b])$. Prostor $C^1([a, b])$ zamjenjujemo prostorom

$$H^1(a, b) = \left\{ \phi \in L^2(a, b) : \int_a^b \phi'(x)^2 dx < +\infty \right\}.$$

To je prostor kvadratno integrabilnih funkcija koje imaju kvadratno integrabilnu derivaciju. U pitanje kako precizno definirati derivaciju funkcije iz $L^2(a, b)$ ovdje nećemo ulaziti (vidi Dodatak A).

U prostoru $H^1(a, b)$ definiramo skalarni produkt

$$(\phi, \psi)_{H^1(a,b)} = \int_a^b \phi(x)\psi(x) dx + \int_a^b \phi'(x)\psi'(x) dx,$$

i pripadnu normu

$$\|\phi\|_{H^1(a,b)} = \left(\int_a^b \phi(x)^2 dx + \int_a^b \phi'(x)^2 dx \right)^{1/2}. \quad (1.16)$$

Pokazuje se da je $H^1(a, b)$ s ovim skalarnim produktom Hilbertov prostor. To jedan od prostora iz familije prostora Soboljeva.

Reformulirajmo sada varijacijsku jednadžbu (1.4) tako da za prostor V uzmemo

$$V = \{\phi \in H^1(0, 1) : \phi(0) = 0\}. \quad (1.17)$$

To je Hilbertov prostor s normom prostora $H^1(0, 1)$. Štoviše, u njemu možemo uzeti ekvivalentnu normu

$$\|\phi\| = \left(\int_0^1 \phi'(x)^2 dx \right)^{1/2}. \quad (1.18)$$

To je posljedica Poincaréove nejednakosti koju dokazujemo sada.

Lema 1.3 Za svaku funkciju $\phi \in H^1(a, b)$, za koju je $\phi(a) = 0$, vrijedi Poincaréova nejednakost

$$\|\phi\|_{L^2(a,b)} \leq C \|\phi'\|_{L^2(a,b)}. \quad (1.19)$$

Dokaz. Za svako $\phi \in V$ imamo

$$\phi(x) = \phi(x) - \phi(a) = \int_a^x \phi'(t) dt \leq \left(\int_a^x 1^2 dx \right)^{1/2} \left(\int_a^x \phi'(t)^2 dt \right)^{1/2},$$

gdje smo iskoristili Cauchyjevu nejednakost. Time dobivamo

$$|\phi(x)| \leq \left((b-a) \int_a^x \phi'(t)^2 dt \right)^{1/2}$$

pa kvadriranjem i integriranjem izlazi

$$\int_a^b \phi(x)^2 dx \leq (b-a) \int_a^b \int_a^x \phi'(t)^2 dt dx \leq (b-a)^2 \int_a^b \phi'(x)^2 dx. \quad \square$$

Iz (1.19) lako slijedi da je na potprostoru svih funkcija iz $H^1(a, b)$, koje se poništavaju u $x = a$, norma (1.16) ekvivalentna s normom

$$|\phi|_{H^1(a,b)} = \left(\int_a^b \phi'(x)^2 dx \right)^{1/2}.$$

Napomena 1.7 *Ekvivalentne norme generiraju istu topologiju na vektorskom prostoru. Drugim riječima, ako jedan niz konvergira u jednoj normi, onda on konvergira i u svakoj drugoj, njoj ekvivalentnoj, normi. Prostor koji je potpun u jednoj normi, potpun je i u njoj ekvivalentnoj normi.*

Zadatak 1.5 Funkcije iz prostora $H^1(a, b)$ su uniformno neprekidne. Naime, za svaku funkciju $v \in H^1(a, b)$ i za svake dvije točke $x, y \in (a, b)$ vrijedi

$$|v(x) - v(y)| \leq \sqrt{|x - y|} \|v'\|_{L^2(a, b)}.$$

Dokažite.

Zadatak 1.6 Uvedimo oznaku $\|v\|_\infty = \sup_{x \in (a, b)} |v(x)|$. Prema zadatku 1.5 $\|v\|_\infty$ je konačno za svaku funkciju $v \in H^1(a, b)$. Pokažite da za svaku funkciju $v \in H^1(a, b)$, za koju je $v(a) = 0$ ili $v(b) = 0$, vrijedi

$$\|v\|_\infty \leq \sqrt{b - a} \|v'\|_{L^2(a, b)}.$$

Vratimo se sada našem primjeru. Sada znamo da na prostoru V , definiranom u (1.17), možemo uzeti normu $\|\cdot\|$, definiranu s (1.18), budući da je ona prema dokazanom ekvivalentna s $H^1(0, 1)$ normom na $V \subset H^1(0, 1)$. Imamo,

$$\begin{aligned} |a(u, v)| &= \left| \int_0^1 u'(x)v'(x) dx \right| \leq \|u'\|_{L^2(0, 1)} \|v'\|_{L^2(0, 1)} = \|u\| \|v\|, \\ a(v, v) &= \int_0^1 v'(x)^2 dx = \|v\|^2. \end{aligned}$$

Uz pretpostavku da je $f \in L^2(0, 1)$ dobivamo neprekidnost linearnog funkcionala F :

$$\begin{aligned} |F(v)| &= \left| \int_0^1 f(x)v(x) dx \right| \leq \left(\int_0^1 f^2(x) dx \right)^{1/2} \left(\int_0^1 v^2(x) dx \right)^{1/2} \\ &= \|f\|_{L^2(0, 1)} \|v\|_{L^2(0, 1)} \leq \|f\|_{L^2(0, 1)} \|v\|. \end{aligned}$$

U zadnjoj smo nejednakosti ponovo koristili ekvivalenciju normi (1.16) i (1.18) na prostoru V . Dakle, promijenjenom osnovnog prostora s lakoćom se pokazuje da varijacijska zadaća (1.4) ima jedinstveno rješenje za svako $f \in L^2(0, 1)$.

Zaključak. Prelaskom na varijacijsku formulaciju rubne zadaće proširili smo pojam rješenja rubne zadaće. Rješenje varijacijske zadaće nazivamo slabim rješenjem zadaće postavljene u diferencijalnom obliku. Na varijacijsku zadaću možemo primijeniti teoreme funkcionalne analize koji na relativno jednostavan način daju korektnost zadaće, pod uvjetom da je funkcijski prostor u kome tražimo rješenje ispravno odabran. To nas vodi do zamjene prostora neprekidno derivabilnih funkcija prostorima Soboljeva. Istodobno dobivamo i nove, slabije, uvjete na funkciju $f(x)$ uz koju je slabo rješenje zadaće (1.1) dobro definirano.

Napomena 1.8 Analogno kao $H^1(a, b)$ definiraju se prostori Soboljeva višeg reda:

$$H^k(a, b) = \{\phi \in L^2(a, b) : \phi', \phi'', \dots, \phi^{(k)} \in L^2(a, b)\}.$$

Pripadna norma je

$$\|\phi\|_{H^k(a, b)} = \left(\int_a^b \phi(x)^2 dx + \sum_{i=1}^k \int_a^b \phi^{(i)}(x)^2 dx \right)^{1/2}.$$

Vrlo je korisna sljedeća elementarna nejednakost.

Lema 1.4 Za svaka dva broja $a, b \geq 0$ i svako $\eta > 0$ vrijedi

$$ab \leq \eta a^2 + \frac{1}{4\eta} b^2. \quad (1.20)$$

Dokaz. Slijedi iz

$$ab \leq \frac{1}{2}(a^2 + b^2),$$

transformacijom $a \mapsto \sqrt{2\eta}a$ i $b \mapsto b/\sqrt{2\eta}$. \square

Zadatak 1.7 Dokažite Poincaréovu nejednakost za funkcije iz $H^1(a, b)$ koje se poništavaju u $x = b$.

Zadatak 1.8 Pokažite primjerom da Poincaréova nejednakost ne vrijedi na čitavom prostoru $H^1(a, b)$.

Zadatak 1.9 Provjerite uz koje uvjete na koeficijente diferencijanih jednadžbi su ispunjeni uvjeti Lax-Milgramove leme za primjere iz Zadatka 1.1.

Rješenje. Uočimo odmah da u varijacijskim zadaćama prostor $C^1([a, b])$ treba zamijeniti s $H^1(a, b)$, te da u svim primjerima funkcija $f(x)$ mora zadovoljavati $f \in L^2(a, b)$.

ad) (1.6). Dovoljno je uzeti $c(x) \geq 0$ za sve $x \in (a, b)$ i $\|c\|_\infty < +\infty$;

ad) (1.7). Dovoljno je uzeti $\alpha \geq 0$, $k(x) \geq k_0 > 0$, za sve $x \in (a, b)$, gdje je k_0 zadana pozitivna konstanta, $\|k\|_\infty < +\infty$ te $c(x) \geq 0$ za sve $x \in (a, b)$ i $\|c\|_\infty < +\infty$. U dokazu ograničenosti bilinearne forme treba se koristiti rezultat zadatka 1.6.

ad) (1.8). Ovaj primjer je složeniji zbog pojave člana $c(x)u'(x)v(x)$ pod integralom. Za ograničenost bilinearne forme moramo imati $\|c\|_\infty < +\infty$ i $\|d\|_\infty < +\infty$. U ocjeni konstante koercitivnosti možemo iskoristiti (1.20) na sljedeći način:

$$\begin{aligned} \int_a^b c(x)u'(x)v(x) dx &\leq \|c\|_\infty \int_a^b \left(\frac{\varepsilon}{2\|c\|_\infty} u'(x)^2 + \frac{\|c\|_\infty}{2\varepsilon} u(x)^2 \right) dx \\ &= \frac{\varepsilon}{2} \int_a^b u'(x)^2 dx + \frac{\|c\|_\infty^2}{2\varepsilon} \int_a^b u(x)^2 dx. \end{aligned}$$

Sada se koercitivnost dobija uz uvjet

$$d(x) \geq d_0 > \frac{\|c\|_\infty^2}{2\varepsilon} > 0, \quad \forall x \in (a, b).$$

Zadatak 1.10 Dokažite neprekidnost funkcionala na desnoj strani u zadatku 1.3.

Definicija 1.3 Za bilinearnu formu $a: V \times V \rightarrow \mathbb{R}$ kažemo da je simetrična ako vrijedi

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

Zadatak 1.11 Ispitajte simetriju bilinearnih formi u primjerima iz Zadatka 1.1.

Napomena 1.9 Varijacijske jednadžbe sa simetričnom bilinearnom formom mogu se interpretirati kao Eulerove jednadžbe za odgovarajući funkcional energije. Preciznije, tada varijacijskoj jednadžbi (1.4) možemo pridružiti funkcional energije

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

Varijacija funkcionala je

$$\left. \frac{d}{dt} J(u + tv) \right|_{t=0} = \left. \frac{d}{dt} \left(\frac{1}{2}a(u + tv, u + tv) - F(u + tv) \right) \right|_{t=0} = a(u, v) - F'(v) = 0,$$

gdje smo iskoristili da je $u \in V$ rješenje varijacijske zadaće (1.4). Stoga rješenje varijacijske zadaće možemo interpretirati kao minimim funkcionala energije (za detalje vidi Poglavlje 2).

1.4 Varijacijska aproksimacija

Varijacijska formulacija rubne zadaće za diferencijalnu jednadžbu vrlo je pogodna za numeričku aproksimaciju. Rekapitulirajmo prvo ono što je do sada učinjeno.

Problem (1.1) smo reformulirali u obliku

$$\begin{cases} \text{naći } u \in V \\ a(u, v) = F(v), \quad \forall v \in V. \end{cases}$$

gdje je

- V beskonačnodimenzionalan Hilbertov prostor zadan s (1.17), i s normom (1.18);
- $a(\cdot, \cdot)$ je bilinearna, ograničena i koercitivna forma zadana s (1.2);
- $F \in V'$ funkcional zadan s (1.3), pri čemu je pretpostavljeno $f \in L^2(0, 1)$.

Da bismo aproksimirali rješenje ove zadaće uvodimo konačnodimenzionalni prostor $S \subset V$ koji predstavlja “aproksimaciju” prostora V i formiramo novu varijacijsku zadaću:

$$\begin{cases} \text{naći } u_S \in S \\ a(u_S, v) = F(v), \quad \forall v \in S. \end{cases} \quad (1.21)$$

Dobiveni problem je konačnodimenzionalan pa se njegovo rješenje može efektivno izračunati. Opisani se postupak naziva varijacijska aproksimacija. Uočimo da diskretna zadaća (1.21), prema

Lax-Milgramovoj lemi, ima jedinstveno rješenje ako su uvjeti Lax-Milgramove leme ispunjeni za kontinuiranu varijacijsku zadaću (1.4).

Neka je $\{\phi_1, \phi_2, \dots, \phi_n\}$ baza u S ($n = \dim(S)$). Tada rješenje u_S možemo raspisati po vektorima baze,

$$u_S(x) = \sum_{j=1}^n U_j \phi_j(x).$$

Uzimajući test funkciju $v = \phi_i$ u zadaći (1.21), dobivamo

$$\sum_{j=1}^n U_j a(\phi_j, \phi_i) = F(\phi_i), \quad i = 1, \dots, n. \quad (1.22)$$

Sustav (1.22) je jedna matrična jednadžba, stoga uvedimo oznake

$$\mathbf{U} = (U_j) \in \mathbb{R}^n, \quad \mathbf{F} = (F_j) \in \mathbb{R}^n, \quad F_j = F(\phi_j), \quad \mathbf{K} = (K_{i,j}), \quad K_{i,j} = a(\phi_j, \phi_i).$$

Jednadžbe (1.22) se zapisuju sada u obliku

$$\mathbf{K}\mathbf{U} = \mathbf{F}.$$

Pokažimo da iz koercitivnosti bilinearne forme slijedi da je matrica \mathbf{K} pozitivno definitna, odnosno da vrijedi

$$\forall \mathbf{V} \neq 0, \quad \mathbf{K}\mathbf{V} \cdot \mathbf{V} > 0. \quad (1.23)$$

Formirajmo funkciju

$$v(x) = \sum_{j=1}^n V_j \phi_j(x).$$

Imamo,

$$\begin{aligned} \mathbf{K}\mathbf{V} \cdot \mathbf{V} &= \sum_{i,j=1}^n a(\phi_j, \phi_i) V_i V_j \\ &= a\left(\sum_{i=1}^n V_i \phi_j, \sum_{j=1}^n V_j \phi_j\right) = a(v, v) \\ &\geq \alpha \|v\|^2. \end{aligned}$$

Kako je $\|\cdot\|$ norma na prostoru V , znamo da

$$\|v\| = 0 \Rightarrow v = 0.$$

Budući da funkcija $v(x)$ ne može biti identički jednaka nuli za netrivialan vektor \mathbf{V} , zaključujemo da vrijedi (1.23).

Matrica \mathbf{K} je, dakle, pozitivno definitna pa je i regularna. Time smo na drugi način pokazali egzistenciju i jedinstvenost diskretnog problema.

Napomena 1.10 Matrica \mathbf{K} se tradicionalno naziva matrica krutosti. \square

1.5 Ocjena greške aproksimacije

Željeli bismo ocijeniti s kojom točnošću rješenje zadatke (1.21) aproksimira rješenje zadatke (1.4), odnosno htjeli bismo ocijeniti grešku aproksimacije $u - u_S$. U toj ocjeni osnovnu ulogu ima svojstvo ortogonalnosti koje izlazi iz toga što je $S \subset V$ (konformnost aproksimacije). Do njega dolazimo oduzimanjem jednadžbi

$$\begin{aligned} a(u, v) &= F(v), \quad \forall v \in S \subset V \\ a(u_S, v) &= F(v), \quad \forall v \in S, \end{aligned}$$

te dobivamo

$$a(u - u_S, v) = 0, \quad \forall v \in S. \quad (1.24)$$

Pretpostavimo da je naša bilinearna forma $a(\cdot, \cdot)$ simetrična, tj da vrijedi

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

pa zbog koercitivnosti ona ima sva svojstva skalarnog produkta. Stoga je

$$\|v\|_E = \sqrt{a(v, v)}$$

dobro definirana norma koju nazivamo **energetska norma**. Kako za svaki skalarni produkt vrijedi Cauchyjeva nejednakost imamo

$$\forall u, v \in V, \quad |a(u, v)| \leq \|u\|_E \|v\|_E. \quad (1.25)$$

Koristeći ortogonalnost (1.24) i Cauchyjevu nejednakost, za proizvoljno $v \in S$ dobivamo:

$$\begin{aligned} \|u - u_S\|_E^2 &= a(u - u_S, u - u_S) \\ &= a(u - u_S, u - v) + a(u - u_S, v - u_S) \\ &= a(u - u_S, u - v) \quad (\text{ortogonalnost}) \\ &\leq \|u - u_S\|_E \|u - v\|_E. \end{aligned}$$

Time smo dobili ocjenu

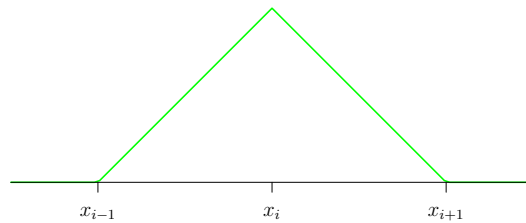
$$\forall v \in S, \quad \|u - u_S\|_E \leq \|u - v\|_E.$$

Uzimanjem infimuma po svim $v \in S$ dobivamo

Teorem 1.2 Neka je $u \in V$ rješenje varijacijske zadatke (1.4), a $u_S \in S$ rješenje zadatke (1.21). Tada je

$$\|u - u_S\|_E \leq \inf\{\|u - v\|_E : v \in S\}.$$

Teorem 1.2 kaže da je varijacijska aproksimacija ujedno najbolja aproksimacija u energetske norme. Infimum na desnoj strani se naravno postiže u $v = u_S$. Štoviše, za potpunu ocjenu greške aproksimacije treba samo vidjeti kako dobro prostor S aproksimira funkcije iz prostora V .



Slika 1.1: Primjer bazne funkcije tipa krovića.

1.6 Konstrukcija konačnodimenzionalnog prostora S

Metoda konačnih elemenata bazira se varijacijskoj aproksimaciji i specifičnom izboru konačnodimenzionalnog prostora S . Općenito metoda konstrukcije prostora S je sljedeća: domena u kojoj je postavljena rubna zadaća razbije se na konačno mnogo podskupova jednostavne geometrije: segmenata u jednoj dimenziji, trokuta ili pravokutnika u dvije, tetraedara i heksaedara u tri itd. Takav rastav domene naziva se *triangulacija*. Prostor S se tada sastoji od funkcija koje su na svakom elementu triangulacije polinomi određenog stupnja, a globalno su neprekidni. Takav izbor dozvoljava konstrukciju baze koja ima povoljna numerička svojstva.

Napomena 1.11 *Pri aproksimaciji diferencijalnih jednadžbi četvrtog reda obično se traži da funkcije iz prostora S imaju neprekidne prve parcijalne derivacije. S druge strane, postoji klasa metoda konačnih elemenata koje koriste prostor S čije funkcije nisu globalno neprekidne.*

Uvedimo najjednostavniji primjer prostora S koji sadrži po dijelovima affine funkcije. Uzmimo da je zadana jedna particija segmenta $[0, 1]$ (triangulacija)

$$0 = x_0 < x_1 < x_2 < \dots < x_n = 1.$$

Prostor S definiramo kao prostor svih funkcija $v(x)$ koje zadovoljavaju sljedeća tri svojstva:

- $v \in C([0, 1])$;
- $v|_{[x_{i-1}, x_i]}$ je afina funkcija za sve $i = 1, 2, \dots, n$;
- $v(0) = 0$.

Lako se pokazuje da je $S \subset H^1(0, 1)$, tj. da je naša aproksimacija konformna.

Prvi zadatak je konstruirati bazu prostora $\{\phi_1, \phi_2, \dots, \phi_n\}$. Standardna baza je sastavljena od *krovića*:

$$\phi_i(x) = \begin{cases} 0 & x \notin (x_{i-1}, x_{i+1}) \\ \frac{x - x_{i-1}}{x_i - x_{i-1}} & x \in [x_{i-1}, x_i] \cap [0, 1] \\ \frac{x_{i+1} - x}{x_{i+1} - x_i} & x \in [x_i, x_{i+1}] \cap [0, 1] \end{cases}$$

ili jednostavnije, $\phi_i(x)$ je funkcija iz S sa svojstvom

$$\phi_i(x_j) = \delta_{i,j},$$

za sve $i, j = 1, 2, \dots, n$. Linearna nezavisnost takvih funkcija je evidentna. Lako je vidjeti da se svaka funkcija $v \in S$ može na jedinstven način napisati u obliku

$$v(x) = \sum_{i=1}^n V_i \phi_i(x),$$

gdje je $V_i = v(x_i)$.

Deriviranjem dobivamo

$$v'(x) = \sum_{i=1}^n V_i \phi_i'(x) = \text{po dijelovima konstantna funkcija}$$

pa je stoga $v' \in L^2(0, 1)$ i time smo provjerili da je $S \subset H^1(0, 1)$.

Uz prostor S definira se i **interpolacijski operator**. Za svaku funkciju $v \in C([0, 1]) \cap V$ definiramo funkciju $v_I \in S$ formulom

$$v_I(x) = \sum_{i=1}^n v(x_i) \phi_i(x).$$

Tu smo evidentno interpolirali funkciju v u čvorovima mreže pomoću afinih funkcija. Interpolacijski operator $\mathcal{I}: C([0, 1]) \cap V \rightarrow S$ definira se formulom $\mathcal{I}v = v_I$. On ima važnu ulogu u ocjeni veličine $\inf\{\|u - v\|_E: v \in S\}$, budući da koristeći ocjenu

$$\inf\{\|u - v\|_E: v \in S\} \leq \|u - u_I\|_E,$$

ocjenu greške svodimo na ocjenu greške interpolacije.

Napomena 1.12 *Točke u kojima uzimamo vrijednosti funkcije u interpolacijskom operatoru nazivaju se **nodalne točke**. U ovom primjeru, nodalne točke su točke x_i , za $i = 1, \dots, n$.*

Teorem 1.3 Neka je $u \in H^2(0, 1) \cap V$ i $h = \max_{1 \leq i \leq n} (x_i - x_{i-1})$. Tada postoji konstanta C , neovisna o h i u , takva da je

$$\|u - u_I\|_E \leq Ch \|u''\|_{L^2(0,1)}.$$

Dokaz. Ako pogledamo definiciju normi, dovoljno je za svako $j = 1, \dots, n$ dokazati

$$\int_{x_{j-1}}^{x_j} |(u - u_I)'(x)|^2 dx \leq c(x_j - x_{j-1})^2 \int_{x_{j-1}}^{x_j} |u''(x)|^2 dx.$$

Tvrđnja se dobije sumiranjem po j . Uvedimo oznaku $e(x) = u(x) - u_I(x)$. Kako je u_I afina funkcija, prethodna nejednakost je ekvivalentna s

$$\int_{x_{j-1}}^{x_j} |e'(x)|^2 dx \leq c(x_j - x_{j-1})^2 \int_{x_{j-1}}^{x_j} |e''(x)|^2 dx.$$

Napravimo li zamjenu varijabli

$$x = x_{j-1} + y(x_j - x_{j-1}), \quad \tilde{e}(y) = e(x_{j-1} + y(x_j - x_{j-1})),$$

izlazi

$$\int_0^1 |\tilde{e}'(y)|^2 dy \leq c \int_0^1 |e''(y)|^2 dy. \quad (1.26)$$

Potrebno je dokazati samo posljednju nejednakost. Prethodna onda slijedi zamjenom varijabli. Takav se postupak često koristi pa je dobio ime **argument skaliranja**.

Da bismo dokazali (1.26) uočimo da je $\tilde{e}(0) = \tilde{e}(1) = 0$ (tu koristimo činjenicu da je u_I interpoland funkcije u). Po Rolleovom teoremu postoji točka $\xi \in (0, 1)$ u kojoj je $\tilde{e}'(\xi) = 0$. Stoga je

$$\tilde{e}'(y) = \int_{\xi}^y \tilde{e}''(t) dt.$$

Pomoću Cauchyjeve nejednakosti izlazi

$$\begin{aligned} |\tilde{e}'(y)| &= \left| \int_{\xi}^y \tilde{e}''(t) dt \right| \leq \left(\int_{\xi}^y 1^2 dt \right)^{1/2} \left(\int_{\xi}^y \tilde{e}''(t)^2 dt \right)^{1/2} \\ &\leq |y - \xi|^{1/2} \left(\int_0^1 \tilde{e}''(t)^2 dt \right)^{1/2}. \end{aligned}$$

Uzmemo li u obzir da je

$$\max_{0 \leq \xi \leq 1} \int_0^1 |y - \xi| dy = \frac{1}{2}, \quad (1.27)$$

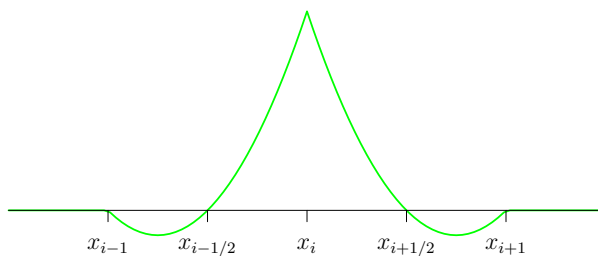
integriranjem dobivamo

$$\int_0^1 |\tilde{e}'(y)|^2 dy \leq \int_0^1 |y - \xi| dy \left(\int_0^1 \tilde{e}''(t)^2 dt \right) \leq \frac{1}{2} \int_0^1 \tilde{e}''(t)^2 dt.$$

Time je (1.26) dokazano s $c = 1/2$. \square

Zadatak 1.12 Dokažite (1.27).

Sada možemo formulirati osnovni rezultat o grešci aproksimacije.

Slika 1.2: Primjer bazne funkcije drugog reda. Funkcija ϕ_i .

Teorem 1.4 Neka je $u \in H^2(0,1) \cap V$ rješenje zadaće (1.4), a u_S rješenje zadaće (1.21). Tada postoji konstanta C , neovisna o $h = \max_{1 \leq i \leq n} (x_i - x_{i-1})$ i u , takva da je

$$\|u - u_S\|_E \leq Ch \|u''\|_{L^2(0,1)}.$$

Dokaz. Pomoću Teorema 1.2 i Teorema 1.3 dobiva se

$$\|u - u_S\|_E \leq \inf\{\|u - v\|_E : v \in S\} \leq \|u - u_I\|_E \leq Ch \|u''\|_{L^2(0,1)}. \quad \square$$

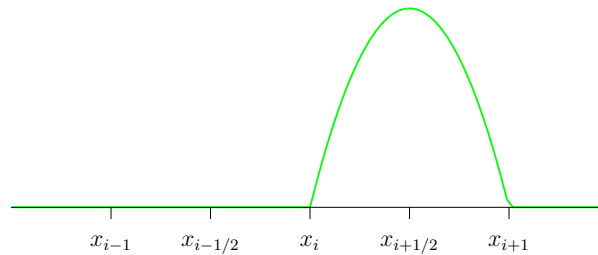
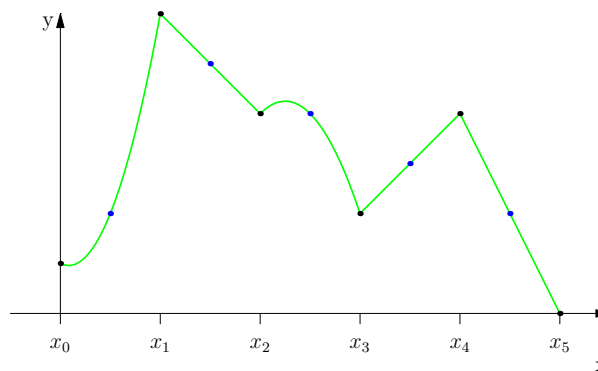
Prostor S koji smo konstruirali je tzv. prostor \mathbb{P}_1 elemenata, nazvanih tako jer su funkcije iz S na svakom elementu $[x_i, x_{i+1}]$ polinomi prvog stupnja. Slično se konstruiraju i elementi višeg reda: elementi tipa \mathbb{P}_k su neprekidne funkcije koje su na svakom elementu polinomi stupnja manjeg ili jednakog k .

Detalji konstrukcije prostora S tipa \mathbb{P}_2 : Prostor možemo konstruirati tako da direktno konstruiramo njegovu bazu. Budući da na svakom elementu $[x_i, x_{i+1}]$ funkcija $v \in S$ mora biti polinom drugog stupnja uvest ćemo dodatnu *nodalnu* točku u element $[x_i, x_{i+1}]$, $x_{i+1/2} = (x_i + x_{i+1})/2$. Funkcija je na $[x_i, x_{i+1}]$ posve određena svojim vrijednostima u točkama x_i , $x_{i+1/2}$ i x_{i+1} . Bazne funkcije se ponovo definiraju kao funkcije iz prostora S koje su jednake jedan u jednoj nodalnoj točki, a nula svim ostalim. Kako sada imamo dvije različite vrste nodalnih točaka, one koje su na rubovima elemenata i one koje su u sredini elemenata, imat ćemo dva različita tipa baznih funkcija: za $i = 0, 1, \dots, n$:

$$\begin{aligned} \phi_i \in S, \quad \phi_i(x_j) &= \delta_{i,j}, \quad j = 0, 1, \dots, n \\ \phi_i(x_{j-1/2}) &= 0, \quad j = 1, \dots, n; \end{aligned}$$

za $i = 1, \dots, n$:

$$\begin{aligned} \phi_{i-1/2} \in S, \quad \phi_{i-1/2}(x_j) &= 0, \quad j = 0, 1, \dots, n \\ \phi_{i-1/2}(x_{j-1/2}) &= \delta_{i,j}, \quad j = 1, \dots, n. \end{aligned}$$

Slika 1.3: Primjer bazne funkcije drugog reda. Funkcija $\phi_{i+1/2}$.

Slika 1.4: Primjer funkcije iz prostora konačnih elemenata drugog reda.

Primjeri baznih funkcija dani su na slikama 1.2 i 1.3. Skup S sada možemo definirati na ovaj način:

$$S = \left\{ \sum_{i=0}^n V_i \phi_i(x) + \sum_{i=1}^n V_{i-1/2} \phi_{i-1/2}(x) : V_0, \dots, V_n, V_{1/2}, \dots, V_{n-1/2} \in \mathbb{R} \right\} \quad (1.28)$$

Zadatak 1.13 Pokažite da je (1.28) linearan prostor svih neprekidnih funkcija koje su na svakom elementu $[x_i, x_{i+1}]$ kvadratni polinomi.

Za funkciju

$$v(x) = \sum_{i=0}^n V_i \phi_i(x) + \sum_{i=1}^n V_{i-1/2} \phi_{i-1/2}(x)$$

vrijedi

$$V_i = v(x_i), \quad V_{i-1/2} = v(x_{i-1/2}).$$

Te se vrijednosti nazivaju stupnjevi slobode. Neprekidnost cijele funkcije postignuta je upravo ovakvim izborom stupnjeva slobode. Jedan primjer funkcije iz ovog prostora dan je na slici 1.4.

Zadatak 1.14 Nađite formule za bazne funkcije tipa \mathbb{P}_2 .

1.7 Argument dualnosti

Teorem 1.4 rješava problem konvergencije metode konačnih elemenata u energetskej normi. Podsjetimo se u slučaju zadaće (1.1) energetska norma ima oblik

$$\|u\|_E = \|u'\|_{L^2(0,1)}.$$

Dakle radi se o konvergenciji derivacije. Ali što možemo reći o ocjeni greške u $L^2(0,1)$ normi? Ta bi ocjena trebala biti bolja od ocjene za derivaciju. Iz Poincaréove nejednakosti (1.19) i Teorema 1.4 odmah dobivamo ocjenu

$$\|u - u_S\|_{L^2(0,1)} \leq Ch \|u''\|_{L^2(0,1)},$$

no pokazat ćemo da vrijedi jača ocjena.

Tehnika koja se naziva argument dualnosti dozvoljava izvod precizne ocjene greške u $L^2(0,1)$ normi. U njoj se prvo definira tzv. dualni problem:

$$\begin{cases} -w'' = u - u_S & \text{na } (0, 1) \\ w(0) = 0, & w'(1) = 0. \end{cases} \quad (1.29)$$

Nadalje, potrebno je uvesti određenu pretpostavku regularnosti. U ovom slučaju pretpostavljamo da zadaća (1.29) ima rješenje u $H^2(0,1)$. U jednodimenzionalnom slučaju to je evidentno jer je $u - u_S \in C([0,1])$ pa (1.29) ima klasično rješenje. Sada imamo:

$$\begin{aligned} \|u - u_S\|_{L^2(0,1)}^2 &= \int_0^1 (u - u_S)(u - u_S) dx = - \int_0^1 (u - u_S)w'' dx \\ &= - \int_0^1 ((u - u_S)w')' dx + \int_0^1 (u - u_S)'w' dx \\ &= -(u - u_S)(1)w'(1) + (u - u_S)(0)w'(0) + \int_0^1 (u - u_S)'w' dx \\ &= \int_0^1 (u - u_S)'w' dx = a(u - u_S, w) \end{aligned}$$

gdje smo iskoristili $w'(1) = 0$ i $u(0) = u_S(0) = 0$. Koristeći relaciju ortogonalnosti (1.24) dobivamo za svako $v \in S$

$$\|u - u_S\|_{L^2(0,1)}^2 = a(u - u_S, w - v) \leq \|u - u_S\|_E \|w - v\|_E.$$

To ćemo pisati u obliku

$$\|u - u_S\|_{L^2(0,1)} \leq \|u - u_S\|_E \frac{\|w - v\|_E}{\|u - u_S\|_{L^2}}.$$

Kako je $-w'' = u - u_S$, a $v \in S$ proizvoljno, imamo

$$\|u - u_S\|_{L^2(0,1)} \leq \|u - u_S\|_E \inf_{v \in S} \frac{\|w - v\|_E}{\|w''\|_{L^2}}.$$

Primijenjujući Teorem 1.3 na funkciju w dobivamo

$$\inf_{v \in S} \frac{\|w - v\|_E}{\|w''\|_{L^2}} \leq \frac{\|w - w_I\|_E}{\|w''\|_{L^2}} \leq Ch.$$

Time smo dobili

$$\|u - u_S\|_{L^2(0,1)} \leq Ch \|u - u_S\|_E,$$

pa još jedna primjena Teorema 1.4 daje

$$\|u - u_S\|_{L^2(0,1)} \leq Ch^2 \|u''\|_{L^2(0,1)}. \quad (1.30)$$

Konvergencija u $L^2(0, 1)$ normi je dakle kvadratična.

1.8 Uniformna ocjena

Željeli bismo imati ocjene u normama koje nisu integralne. Na primjer, uzmimo normu

$$\|v\|_\infty = \sup_{0 \leq x \leq 1} |v(x)|.$$

U tu svrhu se koristi Greenova funkcija za diferencijalni operator. Tehnika je dosta složena u slučaju parcijalne diferencijalne jednadžbe, no ovdje je vrlo jednostavna jer je diferencijalni operator naprosto druga derivacija. U našem primjeru Greenova funkcija je familija funkcija (x je parametar)

$$g_x(t) = \begin{cases} t & \text{za } t < x, \\ x & \text{za } t \geq x \end{cases}$$

To je funkcija iz prostora V pa imamo za proizvoljno $v \in V$

$$a(v, g_x) = \int_0^1 v'(t) g'_x(t) dt = \int_0^x v'(t) dt = v(x).$$

Primijenjujući tu relaciju na $v = u - u_S$, slijedi

$$(u - u_S)(x) = a(u - u_S, g_x).$$

Uočimo sada da za $x = x_i$ funkcija g_x pripada prostoru S , pa koristeći relaciju ortogonalnosti (1.24) dobivamo

$$(u - u_S)(x_i) = 0, \quad i = 1, 2, \dots, n.$$

Dakle $u_S = u_I$, odnosno, rješenje problema (1.21) je interpoland točnog rješenja. Drugim riječima, naša je aproksimacija egzaktna u nodalnim točkama. Sada nam treba odgovarajuća ocjena za grešku interpolacije u max-normi, no takva se greška može izvesti po analogiji s dokazom Teorema 1.3.

Zadatak 1.15 Pokažite da vrijedi sljedeća tvrdnja:

Neka je $u \in H^2(0, 1) \cap V$, funkcija s ograničenom drugom derivacijom i $h = \max_{1 \leq i \leq n} (x_i - x_{i-1})$. Tada postoji konstanta C , neovisna o h i u , takva da je

$$\|u - u_I\|_\infty \leq Ch^2 \|u''\|_\infty.$$

Budući da je $u_S = u_I$ dobivamo ocjenu greške

$$\|u - u_S\|_\infty \leq Ch^2 \|u''\|_\infty.$$

1.9 Algoritamski aspekti metode konačnih elemenata

U ovoj sekciji diskutiramo implementaciju metode konačnih elemenata na modelnoj zadaći

$$-(k(x)u')' + c(x)u = f(x), \quad x \in (a, b) \quad (1.31)$$

$$k(a)u'(a) = \alpha_A, \quad u(b) = u_B, \quad (1.32)$$

u kojoj pretpostavljamo da vrijedi

$$k(x) \geq k_0 > 0, \quad c(x) \geq 0, \quad \forall x \in [a, b]. \quad (1.33)$$

Varijacijska formulacija: Naći $u \in V = H^1(a, b)$, $u(b) = u_B$, za koje vrijedi

$$\int_a^b (k(x)u'\phi' + c(x)u\phi) dx = \int_a^b f(x)\phi dx - \alpha_A\phi(a), \quad \forall \phi \in V^0,$$

gdje je $V^0 = \{\phi \in H^1(a, b) : \phi(b) = 0\}$.

Diskretna zadaća se definira uvođenjem konačnodimenzionalnog prostor $V_h \subset V$ i zadavanjem varijacijske jednadžbe na tom prostoru: treba naći $u_h \in V_h$, $u_h(b) = u_B$, koje zadovoljava

$$\int_a^b (k(x)u_h'\phi' + c(x)u_h\phi) dx = \int_a^b f(x)\phi dx - \alpha_A\phi(a), \quad \forall \phi \in V_h^0, \quad (1.34)$$

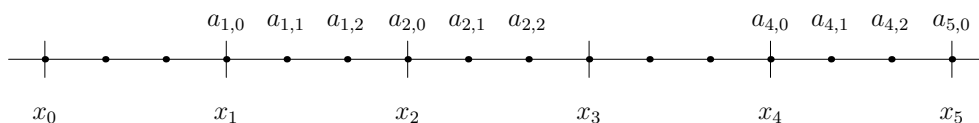
gdje je $V_h^0 = \{\phi \in V_h : \phi(b) = 0\}$.

Za konstrukciju prostora V_h uvodimo triangulaciju domene (a, b) . Ona je zadana nizom točaka

$$a = x_0 < x_1 < x_2 < \dots < x_N = b.$$

Koristit ćemo oznaku $K_{e+1/2} = [x_e, x_{e+1}]$ za *elemente* triangulacije, i

$$\mathcal{T}_h = \{K_{e-1/2} : e = 1, \dots, N\}$$

Slika 1.5: Nodalne točke u slučaju $k = 3$ i $N = 5$.

za samu *triangulaciju*. Triangulacija je dakle skup segmenata koji prekrivaju domenu bez preklapanja, osim u dodirnim točkama susjednih elemenata. Parametar h kojim indeksiramo *triangulaciju* je definiran kao

$$h = \max_{0 \leq e < N} (x_{e+1} - x_e),$$

i naziva se *finoća triangulacije*; triangulacija je to finija što je parametar h manji.

Triangulacija je obično ekvidistantna, ali ako imamo određene informacije o rješenju, ona može biti finija na mjestima gdje se rješenje brže mijenja.

Za fiksno $k = 1, 2, \dots$ definiramo prostor

$$V_h = V_h^{(k)} = \{\phi \in C([a, b]): \forall e = 1, \dots, N, \quad \phi|_{K_{e-1/2}} \in \mathbb{P}_k\}.$$

Dakle u $V_h^{(k)}$ su sve neprekidne funkcije koje su na svakom elementu polinomi stupnja manjeg ili jednakog k . U $V_h^{(1)}$ su po dijelovima affine funkcije, u $V_h^{(2)}$ po dijelovima kvadratne itd. Treba uočiti da se u slučaju $k > 1$ ne radi o splineovima, jer ne zahtjevamo maksimalnu moguću glatkoću funkcija – one trebaju biti jedino neprekidne.

Dimenzija prostora $V_h^{(k)}$: Na svakom elementu imamo $k+1$ stupanj slobode. Kako triangulacija ima N elemenata, to čini ukupno $N(k+1)$ stupnjeva slobode. Neprekidnost u čvorovima x_1, \dots, x_{N-1} je osigurana s $N-1$ uvjeta što daje $N(k+1) - (N-1) = Nk+1$ stupnjeva slobode. To je dimenzija prostora.

Nodalna baza prostora $V_h^{(k)}$: Prvo uvedemo nodalne točke

$$a_{e,l} = x_e + \frac{l}{k} dx_{e+1/2}, \quad e = 0, 1, \dots, N-1; \quad l = 0, 1, \dots, k-1,$$

$$a_{N,0} = b,$$

gdje je $dx_{e+1/2} = x_{e+1} - x_e$. U slučaju $k = 1$ nodalne točke su samo krajevi elemenata. Za $k = 2$ pojavljuju se i centri elemenata, itd. Slučaj $k = 3$ je skiciran na Slici 1.5.

Uočimo da je broj nodalnih točaka jednak $Nk+1$, što je dimenzija prostora $V_h^{(k)}$. Nodalna baza je skup funkcija

$$\varphi_{e,l} \in V_h^{(k)} \quad e = 0, 1, \dots, N-1; \quad l = 0, 1, \dots, k-1, \quad \text{te } \varphi_{N,0} \in V_h^{(k)} \quad (1.35)$$

definiranih svojstvom

$$\varphi_{e,l}(a_{e',l'}) = \begin{cases} 1 & \text{ako je } a_{e,l} = a_{e',l'} \\ 0 & \text{inače} \end{cases}.$$

Funkcija $\varphi_{e,l}$ ima svojstvo da je jednaka jedan u nodalnoj točki $a_{e,l}$, a nula u svim ostalim nodalnim točkama i tim je pravilom definirana u svim nodalnim točkama. Na svakom elementu se nalazi $k + 1$ nodalnih točaka koje u potpunosti određuju polinom stupnja k , pa je stoga funkcija $\varphi_{e,l}$ dobro definirana na svakom pojedinom elementu i evidentno je neprekidna na granici susjednih elemenata (jer su granične točke elemenata ujedno i nodalne točke).

Linearna nezavisnost funkcija (1.35) je evidentna, pa one stoga čine bazu u $V_h^{(k)}$ (broj im je jednak dimenziji prostora). Svaka se funkcija iz $V_h^{(k)}$ može raspisati po toj bazi, a koeficijenti u tom raspisu (*stupnjevi slobode*) su vrijednosti funkcije u nodalnim točkama:

$$v \in V_h^{(k)}, \quad v(x) = \sum_{e=0}^{N-1} \sum_{l=0}^{k-1} v(a_{e,l}) \varphi_{e,l}(x) + v(b) \varphi_{N,0}(x),$$

($b = a_{N,0}$).

Generiranje baznih funkcija pomoću referentnog elementa. Prilikom asembliranja diskretnog sustava pokazuje se praktično sve račune provoditi na jednom fiksiranom elementu, koji se naziva *referentni element*. Svi drugi elementi dobivaju se afnim preslikavanjem referentnog.

Označimo s $\hat{K} = [-1, 1]$ referentni element, a s $T_{e+1/2}$ afino preslikavanje s referentnog elementa na element $K_{e+1/2} = [x_e, x_{e+1}]$:

$$T_{e+1/2}(\xi) = \frac{x_{e+1} - x_e}{2} \xi + \frac{x_{e+1} + x_e}{2}. \quad (1.36)$$

Inverzno preslikavanje je dano formulom

$$T_{e+1/2}^{-1}(x) = \frac{2x - (x_{e+1} + x_e)}{x_{e+1} - x_e}.$$

U referentnom elementu definiramo *nodalne točke*,

$$\xi_l = -1 + 2\frac{l}{k}, \quad l = 0, 1, \dots, k.$$

Nodalnim točkama pridružimo lokalne bazne funkcije

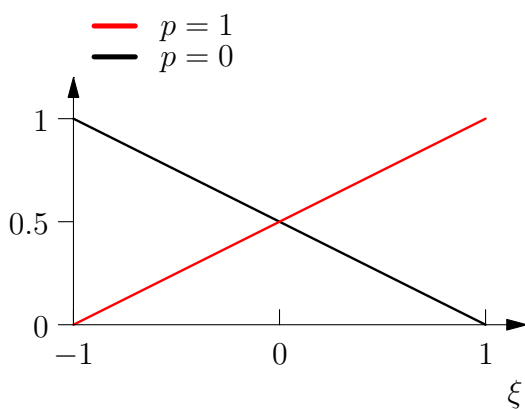
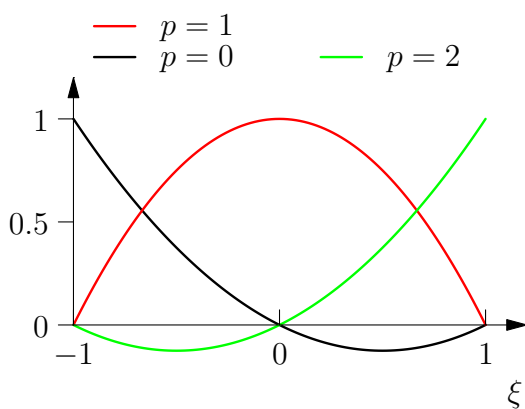
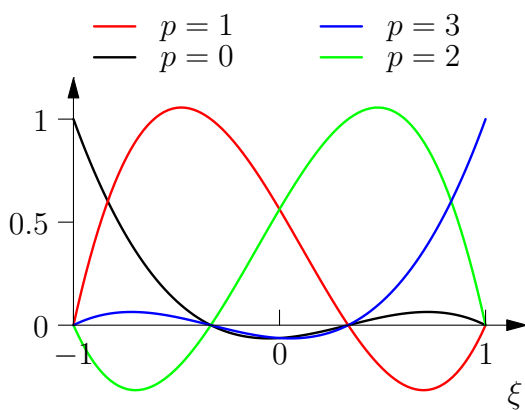
$$\hat{\phi}_0(\xi), \hat{\phi}_1(\xi), \dots, \hat{\phi}_k(\xi) \in \mathbb{P}_k$$

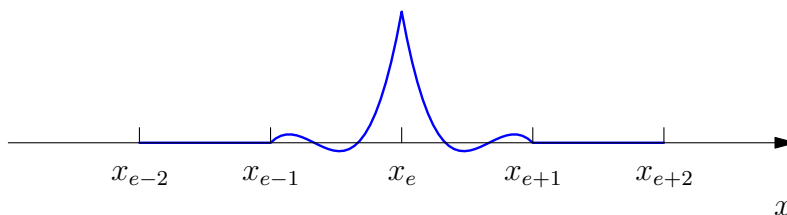
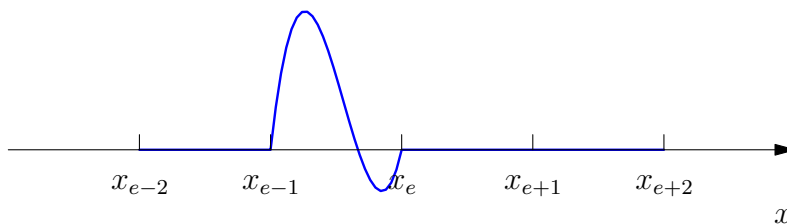
koje su definirane pravilom

$$\hat{\phi}_l(\xi_r) = \delta_{l,r}, \quad l, r = 0, 1, \dots, k,$$

($\delta_{l,r}$ je Kronekerov delta-simbol). Te se funkcije lako konstruiraju pomoću formule

$$\hat{\phi}_l(\xi) = \frac{\prod_{\substack{j=0 \\ j \neq l}}^k (\xi - \xi_j)}{\prod_{\substack{j=0 \\ j \neq l}}^k (\xi_l - \xi_j)}.$$

Slika 1.6: Lokalne bazne funkcije, $k = 1$.Slika 1.7: Lokalne bazne funkcije, $k = 2$.Slika 1.8: Lokalne bazne funkcije, $k = 3$.

Slika 1.9: Primjer globalne bazne funkcije $\varphi_{e,0}$, $k = 3$.Slika 1.10: Primjer globalne bazne funkcije $\varphi_{e,1}$, $k = 3$.

Na slikama 1.6, 1.7 i 1.8 prikazane su funkcije za $k = 1, 2$ i 3 .

Pomoću lokalnih baznih funkcija i afinog preslikavanja referentnog elementa na ciljni element lako konstruiramo globalne bazne funkcije:

$$\varphi_{e,0}(x) = \begin{cases} \hat{\phi}_k(T_{e-1/2}^{-1}(x)) & x \in [x_{e-1}, x_e] \\ \hat{\phi}_0(T_{e+1/2}^{-1}(x)) & x \in [x_e, x_{e+1}] \\ 0 & \text{inače} \end{cases}, \quad (1.37)$$

te za $0 < l < k$,

$$\varphi_{e,l}(x) = \begin{cases} \hat{\phi}_l(T_{e+1/2}^{-1}(x)) & x \in [x_e, x_{e+1}] \\ 0 & \text{inače} \end{cases}. \quad (1.38)$$

Uočimo da funkcije $\varphi_{e,0}$ imaju za nosač dva susjedna elementa (osim u slučaju $e = 0$ i $e = N$), dok sve ostale imaju za nosač jedan element. Primjer globalnih baznih funkcija u slučaju $k = 3$ dani su na Slikama 1.9 i 1.10.

Napomena 1.13 Do sada smo u konstrukciji prostora konačnih elemenata zanemarivali Dirichletov rubni uvjet. Funkcije iz V_h^0 moraju se poništavati u graničnim točkama u kojima je Dirichletov rubni uvjet zadan. Kako su granične točke uvijek i nodalne točke taj dodatni uvjet je lako ispuniti. Treba samo odgovarajući stupanj slobode postaviti na nulu. Prostor time gubi jednu (ili dvije) bazne funkcije. U našem slučaju je

$$V_h^{0(k)} = \{\phi \in V_h^{(k)} : \phi(b) = 0\}$$

prostor razapet baznim funkcijama $\varphi_{e,l} \in V_h^{(k)}$, $e = 0, 1, \dots, N-1$; $l = 0, 1, \dots, k-1$ (bez $\phi_{N,0}$).

Konstrukcija diskretnog sustava. Pri konstrukciji sustava koristit ćemo jednostavnije indeksiranje baznih funkcija, odnosno indeksirat ćemo ih jednim indeksom od 1 do $Nk + 1$. To pretpostavlja da smo indeksirali na isti način nodalne točke. Skup tako indeksiranih nodalnih točaka označimo s $\{b_1, b_2, \dots, b_M, b_{M+1}\}$ (uveli smo pokratu $M = Nk$), a skup baznih funkcija s $\{\phi_1, \phi_2, \dots, \phi_M, \phi_{M+1}\}$. Funkcija $\phi_j \in V_h^{(k)}$ je dakle definirana s

$$\phi_j(b_i) = \delta_{i,j}, \quad i, j = 1, 2, \dots, M + 1.$$

Napomena 1.14 *U jednoj dimenziji postoji prirodan način indeksiranja nodalnih točaka i baznih funkcija. Treba uzeti*

$$b_i = a_{e,l} \quad \text{za } i = k(e - 1) + l + 1, \quad e = 1, \dots, N, \quad l = 0, 1, \dots, k.$$

Tada je $\phi_i = \varphi_{e,l}$.

Uvrstimo aproksimaciju oblika

$$u_h(x) = \sum_{j=1}^{M+1} U_j \phi_j(x),$$

u varijacijsku jednadžbu (1.34) i uzmimo za test funkciju $v = \phi_i$. Pretpostavimo još da smo odabrali prirodan način indeksiranja nodalnih točaka tako da je $b_{M+1} = b$, pa stoga imamo $U_{M+1} = u_B$ i funkcija ϕ_{M+1} ne ulazi u prostor V_h (jer se funkcije u V_h poništavaju u $x = b$). Sada je

$$u_h(x) = \sum_{j=1}^M U_j \phi_j(x) + u_B \phi_{M+1}(x). \quad (1.39)$$

Za vektor

$$\mathbf{U} = (U_j)_{j=1}^M \in \mathbb{R}^M$$

dobivamo sljedeći sustav algebarskih jednadžbi:

$$\sum_{j=1}^{M+1} U_j a(\phi_j, \phi_i) = F(\phi_i), \quad i = 1, \dots, M,$$

gdje je

$$a(\phi_j, \phi_i) = \int_a^b (k(x) \phi_j' \phi_i' + c(x) \phi_j \phi_i) dx, \quad (1.40)$$

$$F(\phi_i) = \int_a^b f(x) \phi_i dx - \alpha_A \phi_i(a). \quad (1.41)$$

Budući da je vrijednost U_{M+1} poznata, treba ju prebaciti na desnu stranu. Time dolazimo do jednadžbi,

$$\sum_{j=1}^M U_j a(\phi_j, \phi_i) = F(\phi_i) - u_B a(\phi_{M+1}, \phi_i), \quad i = 1, \dots, M.$$

(Uočimo da smo time napravili homogenizaciju rubnog uvjeta u diskretnoj zadaći.) Matrica sustava i vektor desne strane su

$$\mathbf{K} = (K_{i,j}) \in \mathbb{R}^{M \times M}, \quad K_{i,j} = a(\phi_j, \phi_i), \quad (1.42)$$

$$\mathbf{F} = (F_i) \in \mathbb{R}^M \quad F_i = F(\phi_i) - u_B a(\phi_{M+1}, \phi_i), \quad (1.43)$$

te treba riješiti

$$\mathbf{K}\mathbf{U} = \mathbf{F}.$$

Aproksimacija u_h dobiva se tada po formuli (1.39).

Napomena 1.15 *U praksi je zgodnije formirati proširenu matricu sustava \mathbf{K}^p i vektor desne strane \mathbf{F}^p u kojima je Dirichletov rubni uvjet zanemaren:*

$$\mathbf{K}^p = (K_{i,j}) \in \mathbb{R}^{M+1 \times M+1}, \quad K_{i,j} = a(\phi_j, \phi_i),$$

$$\mathbf{F}^p = (F_i) \in \mathbb{R}^{M+1} \quad F_i = F(\phi_i),$$

i zatim naknadnom modifikacijom sustava implementirati Dirichletov rubni uvjet.

Osnovno svojstvo matrice sustava \mathbf{K} , odnosno \mathbf{K}^p , je sljedeće: ako indeksi i i j ne odgovaraju nodalnim točkama koje se nalaze u istom elementu, onda je $K_{i,j} = 0$. To je posljedica činjenice što je svaka bazna funkcija različita od nule jedino na onim elementima kojima pripada njena nodalna točka. To može biti jedan ili dva susjedna elementa. Matrica sustava stoga ima veliki broj nula u odnosu na elemente različite od nule. Svaki redak matrice koji odgovara nodalnoj točki koja se nalazi na rubu dva elementa ima $2k + 1$ elemenata različitih od nule. Redak koji odgovara nodalnoj točki iz unutrašnjosti elementa ima $k + 1$ elemenata različitih od nule. Ako primijenimo prirodan način indeksiranja nodalnih točaka, onda matrica sustava dobiva trakastu strukturu što je vrlo povoljno za rješavanje sustava direktnim metodama.

Formiranje matrice sustava moglo bismo vršiti u dvostrukoj petlji oblika

```

for  $i, j = 1, 2, \dots, M$  do
   $K_{i,j} = 0$ 
  for  $e = 0, 1, \dots, N - 1$  do
     $K_{i,j} = K_{i,j} + \int_{K_{e+1/2}} (k(x)\phi_j'\phi_i' + c(x)\phi_j\phi_i) dx$ 
  end for
end for

```

no takav je postupak vrlo neefikasan, budući da se radi o trostrukoj petlji. Postupak se može učiniti efikasnim ako se petlja po svim elementima koristi kao vanjska petlja:

```

for  $e = 0, 1, \dots, N - 1$  do
  for  $i, j = 1, 2, \dots, M$  do
     $K_{i,j} = K_{i,j} + \int_{K_{e+1/2}} (k(x)\phi'_j\phi'_i + c(x)\phi_j\phi_i) dx$ 
  end for
end for

```

pri čemu matricu trebamo inicijalizirati nulama prije ulaska u petlju po elementima. Prednost ovog pristupa je u tome što za dani element $K_{e+1/2}$ ne moramo imati dvostruku petlju po svim elementima matrice. Dovoljno je računati doprinose elementima matrice $K_{i,j}$ koji odgovaraju nodalnim točkama $b_i, b_j \in K_{e+1/2}$; svi ostali doprinosi jednaki su nuli.

Da bismo reducirali dvostruku unutarnju petlju potrebna nam je tabela

$$\text{nodal}(1:N, 0:k)$$

koja će izraziti globalne indekse nodalnih točaka preko lokalnog indeksa nodalne točke u elementu i indeksa elementa. Prisjetimo se da smo pri konstrukciji baznih funkcija i nodalnih točaka iste indeksirali pomoću dva indeksa – indeksom elementa i lokalnim indeksom koji pokazuje položaj točke unutar elementa, dok smo pri konstrukciji sustava algebarskih jednadžbi koristili jedan, globalni indeks. Tabela `nodal` treba povezati ta dva načina indeksiranja.

Za zadani konačni element $K_{e+1/2}$ s indeksom e , $e = 0, 1, \dots, N - 1$, i lokalni indeks nodalne točke $l = 0, 1, \dots, k$, $j = \text{nodal}(e, l)$ je globalni indeks nodalne točke. To znači

$$j = \text{nodal}(e, l) \quad \Rightarrow \quad b_j = a_{e,l}, \quad l = 0, 1, \dots, k - 1;$$

Nadalje, za $l=k$ i $j = \text{nodal}(e, k)$ imamo $b_j = a_{e+1,0}$. (To znači da tabela mora zadovoljavati $\text{nodal}(e+1,0) = \text{nodal}(e,k)$.)

Napomena 1.16 *Uzmemo li prirodan način indeksiranja nodalnih točaka i baznih funkcija, onda je*

$$\text{nodal}(e, l) = k(e-1) + l + 1$$

Proces asbliranja možemo sada napisati u sljedećem obliku:

```

for  $e = 0, 1, \dots, N - 1$  do
  for  $l, r = 0, 1, \dots, k$  do
     $\hat{k}_{l,r} = \int_{K_{e+1/2}} (k(x)\varphi'_{e,r}\varphi'_{e,l} + c(x)\varphi_{e,r}\varphi_{e,l}) dx$ 
     $i = \text{nodal}(e, l), \quad j = \text{nodal}(e, r)$ 
     $K_{i,j} = K_{i,j} + \hat{k}_{l,r}$ 
  end for
end for

```

Sada na svakom konačnom elementu računamo samo lokalnu matricu krutosti koja je dimenzije $k + 1 \times k + 1$, što je optimalna količina računanja za assembliranje globalne matrice sustava.

Računanje integrala. Integriranje ćemo vršiti po referentnom elementu. Za svaki element $K_{e+1/2}$ imamo bijektivno afino preslikavanje $T_{e+1/2}: \hat{K} \rightarrow K_{e+1/2}$, dano formulom (1.36) ($K_{e+1/2} = T_{e+1/2}(\hat{K})$). Prema (1.37) i (1.38), za bazne funkcije imamo za sve $l = 0, 1, \dots, k$,

$$\varphi_{e,l}(x)|_{K_{e+1/2}} = \hat{\phi}_l(T_{e+1/2}^{-1}(x)) \quad (1.44)$$

(uz konvenciju da je $\varphi_{e,k} = \varphi_{e+1,0}$.) Koristeći zamjenu varijabli

$$\int_{K_{e+1/2}} g(x) dx = \int_{\hat{K}} g(T_{e+1/2}(\xi)) |T'_{e+1/2}(\xi)| d\xi = \frac{x_{e+1} - x_e}{2} \int_{\hat{K}} g(T_{e+1/2}(\xi)) d\xi$$

dobivamo ($dx_{e+1/2} = x_{e+1} - x_e$)

$$\int_{K_{e+1/2}} c(x) \varphi_{e,r}(x) \varphi_{e,l}(x) dx = \frac{dx_{e+1/2}}{2} \int_{\hat{K}} c(T_{e+1/2}(\xi)) \hat{\phi}_r(\xi) \hat{\phi}_l(\xi) d\xi. \quad (1.45)$$

Deriviranjem jednakosti (1.44) dobivamo

$$\frac{d}{dx} \varphi_{e,l}(x)|_{K_{e+1/2}} = \frac{2}{x_{e+1} - x_e} \frac{d\hat{\phi}_l}{d\xi}(T_{e+1/2}^{-1}(x)),$$

i stoga

$$\int_{K_{e+1/2}} k(x) \varphi'_{e,r}(x) \varphi'_{e,l}(x) dx = \frac{2}{dx_{e+1/2}} \int_{\hat{K}} k(T_{e+1/2}(\xi)) \frac{d\hat{\phi}_r}{d\xi}(\xi) \frac{d\hat{\phi}_l}{d\xi}(\xi) d\xi. \quad (1.46)$$

Time smo u lokalnim doprinosima matrici sustava posve eliminirali globalne bazne funkcije i cijeli se račun oslanja samo na lokalne bazne funkcije.

Numerička integracija. Ako su koeficijenti diferencijalne jednadžbe konstantni, onda se s integrali (1.45) i (1.46) mogu izračunati egzaktno. U općenitoj situaciji moramo iskoristiti neku formulu numeričke integracije. Pri tome je dovoljno zadati takvu formulu na referentnom elementu.

Odaberimo na referentnom elementu Gaussovu integracijsku formulu oblika

$$\int_{-1}^1 g(\xi) d\xi \approx \sum_{s=1}^L \hat{\omega}_s g(z_s).$$

Ovdje su $z_1, \dots, z_L \in [-1, 1]$ integracijske točke, a $\hat{\omega}_1, \dots, \hat{\omega}_L$ integracijske težine. Time dobivamo

$$\int_{K_{e+1/2}} c(x) \varphi_{e,r}(x) \varphi_{e,l}(x) dx \approx \frac{dx_{e+1/2}}{2} \sum_{s=1}^L \hat{\omega}_s c(T_{e+1/2}(z_s)) \hat{\phi}_r(z_s) \hat{\phi}_l(z_s)$$

$$\int_{K_{e+1/2}} k(x) \varphi'_{e,r}(x) \varphi'_{e,l}(x) dx \approx \frac{2}{dx_{e+1/2}} \sum_{s=1}^L \hat{\omega}_s k(T_{e+1/2}(z_s)) \frac{d\hat{\phi}_r}{d\xi}(z_s) \frac{d\hat{\phi}_l}{d\xi}(z_s).$$

U efikasnoj implementaciji vrijednosti

$$\hat{\phi}_r(z_s), \quad \frac{d\hat{\phi}_r}{d\xi}(z_s)$$

se izračunaju jednom, prije početka petlje po svim elementima. Konačan algoritam asembliranja će imati sljedeću formu, u kojoj smo radi jednostavnosti zapisa stavili $z_s^e = T_{e+1/2}(z_s)$:

```

for  $e = 0, 1, \dots, N - 1$  do
  for  $l, r = 0, 1, \dots, k$  do
     $\hat{k}_{l,r} = 0$ 
    for  $s = 1, \dots, L$  do
       $\hat{k}_{l,r} = \hat{k}_{l,r} + 2\hat{\omega}_s k(z_s^e) \hat{\phi}'_r(z_s) \hat{\phi}'_l(z_s) / dx_{e+1/2} + dx_{e+1/2} \hat{\omega}_s c(z_s^e) \hat{\phi}_r(z_s) \hat{\phi}_l(z_s) / 2$ 
    end for
     $i = \text{nodal}(e, l), j = \text{nodal}(e, r)$ 
     $K_{i,j} = K_{i,j} + \hat{k}_{l,r}$ 
  end for
end for

```

Analogan postupak asembliranja matrice krutosti primjenjuje se i u višedimenzionalnim zadacima. Detalji se mogu pogledati, na primjer, u [1].

Zadatak 1.16 *Modificirajte algoritam asembliranja matrice krutosti tako da umjesto pune matrice koristite matricu u trakastoj formi.*

Zadatak 1.17 *Pokažite da će asembliranje desne strane, bez doprinosa od Neumannovog rubnog uvjeta, ići prema sljedećem algoritmu:*

```

for  $e = 0, 1, \dots, N - 1$  do
  for  $l = 0, 1, \dots, k$  do
     $\hat{f}_{l,r} = 0$ 
    for  $s = 1, \dots, L$  do
       $\hat{f}_l = \hat{f}_l + dx_{e+1/2} \hat{\omega}_s f(z_s^e) \hat{\phi}_l(z_s) / 2$ 
    end for
     $i = \text{nodal}(e, l)$ 
     $F_i = F_i + \hat{f}_l$ 
  end for
end for

```

Modificirajte algoritam tako da uvažite doprinosa on Neumannovog rubnog uvjeta u $x = a$.

Napomena 1.17 Ako koristimo elemnete reda k , onda je dovoljno uzeti formulu numeričke integracije koja je točna na polinomima reda $2k - 2$ kako ne bismo izgubili asimptotski red točnosti metode. Najefikasnije je uzeti Gaussove integracijske formule.

Prethodni algoritmi asembliraju proširenu matricu krutosti \mathbf{K}^p i prošireni vektor desne strane \mathbf{F}^p .

Zadatak 1.18 Pokažite da prošireni sustav $\mathbf{K}^p \mathbf{U}^p = \mathbf{F}^p$, gdje je $\mathbf{U}^p = (U_1, \dots, U_{M+1})^\tau$, odgovara rubnoj zadaći s dva Neumannova rubna uvjeta (homogena ako nema modifikacije vektora desne strane).

Dirichletov rubni uvjet. Sustav s Dirichletovim rubnim uvjetom je oblika

$$\begin{bmatrix} K_{1,1} & \cdots & K_{1,M} \\ \vdots & \cdots & \vdots \\ K_{M,1} & \cdots & K_{M,M} \end{bmatrix} \begin{bmatrix} U_1 \\ \vdots \\ U_M \end{bmatrix} = \begin{bmatrix} F_1 \\ \vdots \\ F_M \end{bmatrix} - \begin{bmatrix} K_{1,M+1}u_B \\ \vdots \\ K_{M,M+1}u_B \end{bmatrix} \quad (1.47)$$

pa bismo morali smanjivati dimenziju proširene matrice krutosti koju smo asemblirali. Kako je to nepraktično, rješava se sustav

$$\begin{bmatrix} K_{1,1} & \cdots & K_{1,M} & K_{1,M+1} \\ \vdots & \cdots & \vdots & \vdots \\ K_{M,1} & \cdots & K_{M,M} & K_{M,M+1} \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} U_1 \\ \vdots \\ U_M \\ U_{M+1} \end{bmatrix} = \begin{bmatrix} F_1 \\ \vdots \\ F_M \\ u_B \end{bmatrix}$$

koji je ekvivalentan s (1.47). Modifikacija proširenog sustava sastoji se dakle u tome da se u retku matrice koji odgovara Dirichletovom čvoru svi vandijagonalni elementi stave na nulu, dijagonalni element se postavi na jedan, a odgovarajuća komponenta desne strane se postavi na vrijednost Dirichelovog rubnog uvjeta. Loša strana ovog postupka je gubitak simetrije matrice. Da bismo to izbjegli (simetrija je važna za neke iterativne algoritme, kao što je algoritam konjugiranih gradijenata) možemo modificirati sustav na sljedeći način:

$$\begin{bmatrix} K_{1,1} & \cdots & K_{1,M} & 0 \\ \vdots & \cdots & \vdots & \vdots \\ K_{M,1} & \cdots & K_{M,M} & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} U_1 \\ \vdots \\ U_M \\ U_{M+1} \end{bmatrix} = \begin{bmatrix} F_1 \\ \vdots \\ F_M \\ u_B \end{bmatrix} - \begin{bmatrix} K_{1,M+1}u_B \\ \vdots \\ K_{M,M+1}u_B \\ 0 \end{bmatrix}$$

Vidimo da je ovdje jednačba koja odgovara Dirichletovom čvoru $i = M + 1$ zamijenjena s jednačbom $u_{M+1} = u_B$.

Treći način implementacije Dirichletovog rubnog uvjeta je putem *penalizacije*. To je "lijena" metoda koja prošireni sustav modificira minimalno i ne kviri simetriju susatva. Radi se o tome da se jednačbi u Dirichletovom čvoru $i = M + 1$ doda jednačba

$$\frac{1}{\varepsilon} U_{M+1} = \frac{1}{\varepsilon} u_B,$$

gdje je ε vrlo mali broj (npr. $\varepsilon = 10^{-20}$). Time će rubni uvjet biti približno zadovoljen, naravno pod uvjetom da je $1/\varepsilon$ vrlo velik u odnosu na elemente $M+1$ -og retka matrice krutosti i u odnosu na F_{M+1} . Jedina modifikacija koju treba učiniti na sustavu je

$$K_{M+1,M+1} = K_{M+1,M+1} + \frac{1}{\varepsilon}, \quad F_{M+1} = F_{M+1} + \frac{1}{\varepsilon}u_B.$$

Zadatak 1.19 *Matrica krutosti diskretnog sustava je simetrična jer polazna diferencijalna jednadžba ima odgovarajuće svojstvo simetrije. Formirajte algoritam za diskretizaciju rubne zadaće (k , b , c i f su zadane funkcije)*

$$\begin{aligned} -(k(x)u')' + b(x)u'(x) + c(x)u &= f(x), \quad x \in (a, b) \\ u(a) &= u_A, \quad u(b) = u_B, \end{aligned}$$

i uvjerite se da matrica krutosti nije simetrična ako je $b \neq 0$.

Neumannova zadaća. Regularnost matrice krutosti slijedi iz njene pozitivne definitnosti koju osiguravaju svojstva koeficijenata diferencijalne jednadžbe. Tako na primjer, kod zadaće (1.31), (1.32) pretpostavili smo uvjete (1.33) koji za proširenu matricu krutosti daju

$$\sum_{i,j=1}^{M+1} K_{i,j}\xi_i\xi_j = \int_a^b (k(x)U'(x)^2 + c(x)U(x)^2) dx$$

gdje je

$$U(x) = \sum_{i=1}^{M+1} \xi_i\phi(x).$$

Sada iz $k(x) \geq k_0 > 0$ i $c(x) \geq 0$, za sve $x \in [a, b]$, imamo

$$\sum_{i,j=1}^{M+1} K_{i,j}\xi_i\xi_j \geq k_0 \int_a^b U'(x)^2 dx,$$

i stoga

$$\sum_{i,j=1}^{M+1} K_{i,j}\xi_i\xi_j = 0 \quad \Rightarrow \quad U' \equiv 0.$$

To je pak moguće ako i samo ako je $\xi_1 = \xi_2 = \dots = \xi_{M+1}$. Za matricu krutosti Dirichletove zadaće imamo analogni zaključak:

$$\sum_{i,j=1}^M K_{i,j}\xi_i\xi_j = \int_a^b (k(x)V'(x)^2 + c(x)V(x)^2) dx$$

gdje je

$$V(x) = \sum_{i=1}^M \xi_i \phi_i(x).$$

Funkcija V sada zadovoljava $V(b) = 0$ i uvjet $V' \equiv 0$ nužno povlači $V \equiv 0$, što je moguće ako i samo ako je $\xi_1 = \xi_2 = \dots = \xi_M = 0$. Dirichletova matrica je stoga pozitivno definitna.

Uočimo da je uz pretpostavku $c(x) \geq c_0 > 0$, za sve $x \in [a, b]$, i proširena matrica krutosti pozitivno definitna.

Neumannova zadaća može stoga vodiri na diskretnu zadaću s neregularnom matricom. Na primjer, zadaća

$$\begin{aligned} -(k(x)u')' &= f(x), & x \in (a, b) \\ u'(a) &= u'(b) = 0, \end{aligned}$$

ima matricu krutosti $\mathbf{K} \in \mathbb{R}^{M+1, M+1}$ koja je singularna. Štoviše,

$$\mathcal{N}(\mathbf{K}) = \{\alpha \mathbf{1} : \alpha \in \mathbb{R}\}. \quad (1.48)$$

Naime,

$$(\mathbf{K}\mathbf{1})_i = \sum_{j=1}^{M+1} K_{i,j} = \sum_{j=1}^{M+1} \int_a^b (k(x)\phi_j'\phi_i') dx = \int_a^b (k(x)1'\phi_i') dx = 0,$$

jer je

$$\sum_{j=1}^{M+1} \phi_j \equiv 1.$$

To pokazuje da je $\mathbf{1}$ element jezgre matrice krutosti. S druge strane, pokazali smo više da svaki vektor iz jezgre nužno zadovoljava $\xi_1 = \xi_2 = \dots = \xi_{M+1}$, što dokazuje (1.48).

U ovom slučaju desna strana mora zadovoljavati uvjet kompatibilnosti

$$\int_a^b f(x) dx = 0,$$

koji u diskretnom slušaju daje

$$\sum_{j=1}^{M+1} F_j = \mathbf{F} \cdot \mathbf{1} = 0.$$

Time dobivamo da diskretna desna strana mora biti ortogonalna na jezgru matrice sustava, što je nužan i dovoljan uvjet egzistencije rješenja diskretnog sustava. Rješenje diskretnog sustava, kao i kontinuiranog, određeno je do na aditivnu konstantu.

U praksi je nezgodno raditi sa singularnom matricom. Stoga se postupa tako da se rješenje zada u jednoj točki. Svako drugo rješenje dobiva se dodavanjem aditivne konstante. Posebno je zgodno staviti rješenje na nulu jedno točki, na primjer u $x = b$. Sustav se tada transformira u

$$\begin{bmatrix} K_{1,1} & \cdots & K_{1,M} & 0 \\ \vdots & \cdots & \vdots & \vdots \\ K_{M,1} & \cdots & K_{M,M} & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} U_1 \\ \vdots \\ U_M \\ U_{M+1} \end{bmatrix} = \begin{bmatrix} F_1 \\ \vdots \\ F_M \\ 0 \end{bmatrix}. \quad (1.49)$$

Zadatak 1.20 Pokažite da je matrica sustava (1.49) regularna.

Bibliografija

- [1] Jean-Luc Guermond Alexandre Ern. *Éléments finis: théorie, applications, mise en oeuvre*, volume 36 of *Mathématiques & Applications*. Springer-Verlag, Berlin, 2002.
- [2] H. Brezis. *Analyse Fonctionnelle*. Masson, Paris, 1983.
- [3] Svetozar Kurepa. *Funkcionalna analiza. Elementi teorije operatora*. Školska knjiga, Zagreb, 1990.

2

Varijacijska formulacija eliptičkih rubnih zadaća

Rubni problemi za diferencijalne jednačbe eliptičkog tipa predstavljaju široku klasu matematičkih zadaća koje opisuju ravnotežna stanja različitih fizikalnih sustava. *Klasična rješenja* tih zadaća moraju zadovoljiti diferencijalnu jednačbu i rubni uvjet u svakoj točki domene i njezine granice te stoga koeficijenti diferencijalne jednačbe moraju biti neprekidne funkcije, odnosno neprekidno derivabilne ako se nalaze pod znakom derivacije. S fizikalnog (nematematičkog) stajališta prirodno je zahtijevati da koeficijenti diferencijalne jednačbe, čak i kad se nalaze pod znakom deriviranja, nisu neprekidne funkcije, odnosno da imaju *skokove*. Time dolazimo do potrebe da se pojam rješenja rubne zadaće generalizira kako bi se mogao primijeniti i na zadaće s diskontinuiranim koeficijentima.

Postoji više načina da se generalizira pojam rješenja rubne zadaće. Dva su osnovna, zamjena diferencijalne jednačbe *integralnom jednačbom* te uvođenje *varijacijske jednačbe*. Integralna formulacija rubne zadaće je na određen način prirodnija od varijacijske jer u izvodu diferencijalne jednačbe iz fizikalnog zakona najčešće polazimo od integralne jednačbe koja formulira određen zakon sačuvanja te koristimo princip lokalizacije. Integralna jednačba ima stoga strogo fizikalno značenje, a zahtjevi glatkoće koeficijenata potrebni su da bi se mogla izvršiti *lokalizacija* koja vodi do diferencijalne jednačbe (vidi npr. [1]). Integralna formulacija stoga predstavlja prirodan način generalizacije rubne zadaće, ali s teorijskog stajališta nije jako korisna pa se ne koristi za matematičku analizu rubnih zadaća. S druge strane, integralna formulacija je osnova klase numeričkih metoda koje se nazivaju *metodama konačnih volumena* i koje pored metoda konačnih elemenata predstavljaju najrašireniji način numeričkog rješavanja rubnih zadaća.

Formulacija rubne zadaće u obliku varijacijske jednačbe ima dosta zajedničkih točaka s integralnom formulacijom i puno je pogodnija za teorijsku analizu rubnih zadaća. Pomoću varijacijske jednačbe formulira se tzv. *slabo rješenje* rubne zadaće koje postoji uz puno slabije uvjete na koeficijente diferencijalne jednačbe od onih nužnih za egzistenciju klasičnog rješenja. K tome, egzistenciju slabog rješenja je lakše dokazati te je svako slabo rješenje ujedno i klasično ako ima dovoljnu glatkoću. Varijacijska se formulacija koristi kao polazna točka za klasu numeričkih pos-

tupaka koje nazivamo *metoda konačnih elemenata*. Stoga ćemo u ovom poglavlju pokazati kako se postavlja varijacijska formulacija eliptičke rubne zadaće i kako se na toj osnovi formulira numerička metoda. U dodatku ćemo dati rezultate iz funkcionalne analize koji omogućuju dokazivanje egzistencije i jedinstvenosti slabog rješenja linearne eliptičke zadaće.

2.1 Uvodni primjer

Razmotrit ćemo sljedeću Dirichletovu rubnu zadaću:

$$\begin{cases} -\Delta u + u = f & \text{u } \Omega \\ u = 0 & \text{na } \partial\Omega \end{cases} \quad (2.1)$$

gdje je $\Omega \subset \mathbb{R}^d$ otvoren i povezan skup, a $\partial\Omega$ njegova granica. Klasično rješenje zadaće (2.1) je funkcija $u \in C^2(\Omega) \cap C_0(\Omega)$ koja diferencijalnu jednadžbu¹ (2.1)₁ zadovoljava u svakoj točki domene Ω i poništava se u svakoj točki granice $\partial\Omega$. Da bi takvo rješenje postojalo evidentno desna strana f mora biti neprekidna funkcija.

Metoda konačnih elemenata se temelji na varijacijskoj formulaciji rubne zadaće, koja se u slučaju zadaće (2.1) dobiva množenjem jednadžbe (2.1)₁ s test funkcijom $v \in C^1(\overline{\Omega}) \cap C_0(\Omega)$ i integracijom po Ω :

$$\int_{\Omega} (-\Delta u + u)v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x}.$$

Parcijalnom integracijom u integralu na lijevoj strani dobivamo

$$-\int_{\partial\Omega} (\nabla u \cdot \mathbf{n})v \, d\sigma + \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x},$$

gdje je \mathbf{n} jedinična vanjska normala na $\partial\Omega$. Budući da se test funkcija v poništava na rubu domene Ω , integral po $\partial\Omega$ propada i dobivamo

$$\int_{\Omega} (\nabla u \cdot \nabla v + uv) \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x}, \quad \forall v \in C^1(\overline{\Omega}) \cap C_0(\Omega). \quad (2.2)$$

To je varijacijska formulacija rubne zadaće (2.1). Uočimo da se u njoj ne pojavljuju druge derivacije nepoznate funkcije $u(\mathbf{x})$ te da ju može zadovoljavati funkcija iz prostora $C^1(\overline{\Omega}) \cap C_0(\Omega)$. Na toj činjenici se bazira generalizacija pojma rješenja diferencijalne zadaće. Pri tome, da bismo imali korektnu generalizaciju, treba pokazati da je svako rješenje zadaće (2.2), koje je dva puta neprekidno derivabilno, ujedno i rješenje zadaće (2.1). Pretpostavimo stoga da imamo funkciju

¹Za definiciju funkcijskih prostora vidi Poglavlje A. Prostor $C_0(\Omega)$ čine funkcije iz $C(\overline{\Omega})$ koje se poništavaju na rubu.

$u(\mathbf{x})$ koja zadovoljava (2.2) i koja pripada prostoru $C^2(\overline{\Omega}) \cap C_0(\Omega)$. Parcijalnom integracijom u (2.2), prebacujući derivacije na funkciju u , dobivamo,

$$-\int_{\partial\Omega} (\nabla u \cdot \mathbf{n})v \, d\sigma = \int_{\Omega} (\Delta u - u + f)v \, d\mathbf{x}, \quad \forall v \in C_0^1(\Omega).$$

Kako je lijeva strana jednaka nuli izlazi

$$\int_{\Omega} (\Delta u - u + f)v \, d\mathbf{x} = 0, \quad \forall v \in C_0^1(\Omega).$$

Sada trebamo iskoristiti osnovnu lemu varijacijskog računa:

Propozicija 2.1 Neka je $g \in C(\Omega)$ funkcija koja zadovoljava

$$\int_{\Omega} gv \, d\mathbf{x} = 0, \quad \forall v \in C_0^\infty(\Omega).$$

Tada je $g(\mathbf{x}) = 0$ za svako $\mathbf{x} \in \Omega$.

[Napomena o dokazu i općenitijem rezultatu ...]

Primjenjujući Propoziciju 2.1 na $g = \Delta u - u + f$ dobivamo da je diferencijalna jednadžba (2.1)₁ zadovoljena u svakoj točki domene Ω . Kako je rubni uvjet zadovoljen već pretpostavkom da je $u \in C_0(\Omega)$ možemo zaključiti da je funkcija u klasično rješenje zadaće (2.1).

Time smo posve opravdali *varijacijsku jednadžbu* (2.2) kao generalizaciju zadaće (2.1). Pogledajmo sada strukturu jednadžbe (2.2). Prvo uvedimo oznaku $V = C^1(\overline{\Omega}) \cap C_0(\Omega)$ i uočimo da je V linearan prostor uz algebarske uobičajene operacije nad funkcijama. Nadalje, definirajmo preslikavanja

$$a: V \times V \rightarrow \mathbb{R}, \quad F: V \rightarrow \mathbb{R},$$

formulama

$$a(u, v) = \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, d\mathbf{x}, \quad F(v) = \int_{\Omega} fv \, d\mathbf{x}. \quad (2.3)$$

Zadaću (2.2) možemo sada apstraktno zapisati u obliku: Treba naći $u \in V$ koje zadovoljava

$$\forall v \in V, \quad a(u, v) = F(v). \quad (2.4)$$

Iz formule (2.3) lako se uočava da je F linearan funkcional, a da je $a(\cdot, \cdot)$ *bilineran*, tj. linearan u svakoj varijabli zasebno (vidi Definiciju 2.4). Sada možemo iskoristiti vrlo općenit egzistencijalni rezultat iz funkcionalne analize koji nam daje uvjete uz koje zadaća (2.4) ima jedinstveno rješenje:

Teorem 2.1 (Lax-Milgramova lema) Neka je V Hilbertov prostor s normom $\|\cdot\|$ i $a: V \times V \rightarrow \mathbb{R}$ neka je bilinearna forma sa sljedećim svojstvima:

$$\begin{aligned} \exists M > 0 \text{ takvo da je } \forall u, v \in V, \quad |a(u, v)| &\leq M\|u\|\|v\| \text{ (ograničenost);} \\ \exists \alpha > 0 \text{ takvo da je } \forall v \in V, \quad a(v, v) &\geq \alpha\|v\|^2 \text{ (koercitivnost).} \end{aligned}$$

Tada za svaki linearan i neprekidan funkcional $F \in V'$ zadaća

$$\begin{cases} \text{Naći } u \in V \\ \forall v \in V, \quad a(u, v) = \langle F, v \rangle \end{cases} \quad (2.5)$$

ima jedinstveno rješenje $u \in V$ i vrijedi

$$\|u\| \leq \frac{1}{\alpha} \|F\|_{V'}.$$

Dakle da bismo primijenili Lax-Milgramovu lemu trebamo pokazati da je odabrani prostor V Hilbertov, da su bilinearna forma i linearan funkcional ograničeni te uvjet koercitivnosti bilinearne forme. Uvjet koercitivnosti je pri tome najvažniji i usko je vezan uz izbor prostora V .

Budući da je bilinearna forma dana u integralnom obliku nije moguće postići njenu koercitivnost ako za prostor V uzmemo prostor neprekidno derivabilnih funkcija. Da bismo se u to uvjerali, uzmimo $V = C^1(\bar{\Omega}) \cap C_0(\Omega)$, što je prirodan odabir prostora u našem slučaju. Norma, u kojoj je taj prostor funkcija potpun dana je formulom

$$\|u\|_{C^1(\bar{\Omega})} = \sup_{\mathbf{x} \in \Omega} |u(\mathbf{x})| + \sum_{i=1}^d \sup_{\mathbf{x} \in \Omega} \left| \frac{\partial u(\mathbf{x})}{\partial x_i} \right|.$$

Uvjet koercitivnosti sada glasi: postoji konstanta $\alpha > 0$ takva da je

$$\forall v \in C_0^1(\Omega), \quad \int_{\Omega} (|\nabla v|^2 + v^2) d\mathbf{x} \geq \alpha \|v\|_{C^1(\bar{\Omega})}^2 \quad (2.6)$$

Uvjet (2.6) nije moguće zadovoljiti, što se može vidjeti tako da se konstruira kontraprimjer u jednoj prostornoj dimenziji. Razlog se nalazi u tome što na lijevoj strani imamo integralni izraz, dok je na desnoj strani max-norma. Naprosto moramo oslabiti normu prostora $C^1(\bar{\Omega})$ da bismo zadovoljili uvjet koercitivnosti, a to onda znači i proširiti prostor u kome tražimo rješenje, jer on mora biti potpun. [Uputa na zadatak -> uvodno poglavlje.]

Izlaz je u tome da prijedemo na itegralne norme. Najjednostavnije bi bilo u prostor $C_0^1(\bar{\Omega})$ uvesti novu normu:

$$\|v\| = \left(\int_{\Omega} |v(\mathbf{x})|^2 dx + \sum_{i=1}^d \int_{\Omega} \left| \frac{\partial v(\mathbf{x})}{\partial x_i} \right|^2 dx \right)^{1/2}. \quad (2.7)$$

Time je evidentno uvjet koercitivnosti sada zadovoljen, ali smo izgubili potpunost prostora: prostor $C_0^1(\Omega)$ nije potpun u novoj normi. Taj se problem rješava prijelazom na širi prostor koji je *upotpunjenje* prostora $C_0^1(\Omega)$ u normi (2.7) i koji se označava $H_0^1(\Omega)$. To je jedan iz klase prostora Soboljeva.

Uzimajući za prostor $V = H_0^1(\Omega)$ lako se pokazuje da su svi uvjeti Lax-Milgramove leme zadovoljeni i stoga poopćena zadaća

$$\begin{cases} \text{Naći } u \in H_0^1(\Omega) \\ \forall v \in H_0^1(\Omega), \quad \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} \end{cases} \quad (2.8)$$

ima jedinstveno rješenje. Time smo došli do generalizacije pojma rješenja zadaće (2.1) u kojoj se ne pojavljuju druge derivacije rješenja, a niti prve derivacije ne moraju nužno biti neprekidne. Egzistenciju i jedinstvenost osigurava jednostavan rezultat funkcionalne analize, a jedini nezgodan dio ovog pristupa je u tome što rješenje tražimo u prostoru definiranom na vrlo apstraktan način – kao upotpunjenje prostora $C^1(\bar{\Omega}) \cap C_0(\Omega)$. Detaljnija prezentacija prostora $H_0^1(\Omega)$ i drugih prostora Soboljeva dana je u Dodatku A, dok osnovne definicije i rezultate iznosimo u sljedećoj sekciji.

U našem je primjeru bilinearna forma simetrična (Definicija 2.7), pa je prema Lemi 2.5 varijacijska jednadžba (2.2) Eulerova jednadžba za funkcional energije

$$J(v) = \frac{1}{2} \int_{\Omega} |\nabla v|^2 \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x},$$

i slabo rješenje zadaće (2.1) predstavlja minimum funkcionala energije na prostoru $H_0^1(\Omega)$. Ta se činjenica može iskoristiti za konstrukciju numeričkog algoritma.

Varijacijska jednadžba (2.5) predstavlja temelj za sustavnu konstrukciju numeričkih postupaka za aproksimaciju rubnih zadaća. Dovoljno je zamijeniti beskonačnodimenzionalan prostor V s nekim njegovim konačnodimenzionalnim potprostorom $V_h \subset V$. Time dobivamo konačnodimenzionalnu varijacijsku zadaću:

$$\begin{cases} \text{Naći } u_h \in V_h \\ \forall v \in V_h, \quad a(u_h, v) = \langle F, v \rangle \end{cases} \quad (2.9)$$

Ako imamo niz prostora $V_h \subset V$ koji sve bolje aproksimiraju prostor V , kada parametar $h \rightarrow 0$, onda očekujemo da će $u_h \rightarrow u$ kada $h \rightarrow 0$. Mi ćemo dokazati i više od toga odnosno, odredit ćemo red konvergencije postupka u ovisnosti o načinu konstrukcije prostora V_h . Pokažimo sada da je konačnodimenzionalni problem (2.9) dobro postavljen.

Lema 2.1 Uz pretpostavke Lax-Milgramove leme problem (2.9) ima jedinstveno rješenje.

Dokaz. Neka je $\{\phi_1, \phi_2, \dots, \phi_n\}$ baza u V_h ($n = \dim(V_h)$). Tada je

$$u_h(\mathbf{x}) = \sum_{j=1}^n U_j \phi_j(\mathbf{x}).$$

Uzimajući za test funkciju $v = \phi_i$ u (2.9), dobivamo sustav

$$\sum_{j=1}^n U_j a(\phi_j, \phi_i) = \langle F, \phi_i \rangle, \quad i = 1, \dots, n.$$

U oznakama

$$\mathbf{U} = (U_j) \in \mathbb{R}^n, \quad \mathbf{F} = (F_j) \in \mathbb{R}^n, \quad F_j = \langle F, \phi_j \rangle, \quad \mathbf{K} = (K_{i,j}), \quad K_{i,j} = a(\phi_j, \phi_i),$$

imamo sustav linearnih algebarskih jednadžbi $\mathbf{K}\mathbf{U} = \mathbf{F}$. Matrica \mathbf{K} je pozitivno definitna. Zaista, za $\boldsymbol{\xi} = (\xi_j) \in \mathbb{R}^n$, zbog koercitivnosti, imamo

$$\mathbf{K}\boldsymbol{\xi} \cdot \boldsymbol{\xi} = \sum_{i,j=1}^n a(\phi_j, \phi_i) \xi_i \xi_j = a\left(\sum_{j=1}^n \xi_j \phi_j, \sum_{i=1}^n \xi_i \phi_i\right) \geq \alpha \left\| \sum_{i=1}^n \xi_i \phi_i \right\|^2 \geq 0.$$

Ako je $\mathbf{K}\boldsymbol{\xi} \cdot \boldsymbol{\xi} = 0$, onda zbog linearne nezavisnosti baznih funkcija slijedi $\boldsymbol{\xi} = 0$ i pozitivna definitnost je dokazana. Iz pozitivne definitnosti slijedi regularnost matrice \mathbf{K} , pa problem (2.9) ima jedinstveno rješenje. \square

Matricu \mathbf{K} tradicionalno nazivamo **matrica krutosti** i vidimo da koercitivnost bilinearne forme implicira njenu pozitivnu definitnost. K tome, ako je bilinearna forma simetrična, onda je i matrica krutosti simetrična.

Opisanu shemu definicije slabog rješenja rubne zadaće primijenit ćemo sada na općenitije eliptičke jednadžbe drugog reda no prije toga ćemo s više detalja opisati prostore Soboljeva.

2.2 Prostori $L^2(\Omega)$, $H^1(\Omega)$ i $H_0^1(\Omega)$

U prošloj smo sekciji uveli prostor $H_0^1(\Omega)$ kao upotpunjenje prostora $C^1(\overline{\Omega}) \cap C_0(\Omega)$ u odgovarajućoj normi. U ovoj sekciji dajemo konkretniji opis tog prostora.

Kao i do sada neka je $\Omega \subset \mathbb{R}^d$ otvoren i ograničen skup. Funkcija $u: \Omega \rightarrow \mathbb{R}$ je apsolutno integrabilna ako je

$$\int_{\Omega} |u(\mathbf{x})| d\mathbf{x} < +\infty,$$

gdje je integral uzet u odnosu na Lebesgueovu mjeru na \mathbb{R}^d . Skup svih apsolutno integrabilnih funkcija ima strukturu vektorskog prostora i označava se s $L^1(\Omega)$. Štoviše, on postaje potpun normiran prostor kada u njega uvedemo normu

$$\|u\|_{L^1(\Omega)} = \int_{\Omega} |u(\mathbf{x})| d\mathbf{x}.$$

S $L_{\text{loc}}^1(\Omega)$ označavamo linearan prostor svih realnih funkcija na Ω koje su apsolutno integrabilne na svakom kompaktnom skupu $K \subset \Omega$. Takve funkcije ne moraju nužno biti apsolutno integrabilne na cijelom skupu Ω , pa imamo strogu inkluziju $L^1(\Omega) \subset L_{\text{loc}}^1(\Omega)$.

Slično se definira $L^2(\Omega)$, kao linearan prostor svih funkcija $u(\mathbf{x})$ na Ω sa svojstvom da je $u^2 \in L^1(\Omega)$, odnosno

$$L^2(\Omega) = \{u: \Omega \rightarrow \mathbb{R}: \int_{\Omega} |u(\mathbf{x})|^2 d\mathbf{x} < +\infty\}.$$

Norma u $L^2(\Omega)$ je dana formulom

$$\|u\|_{L^2(\Omega)} = \left[\int_{\Omega} |u(\mathbf{x})|^2 d\mathbf{x} \right]^{1/2}.$$

Uvest ćemo još oznaku $L^\infty(\Omega)$ za prostor svih ograničenih (izmjerivih) funkcija. Norma u tom prostoru definira se formulom

$$\|u\|_{L^\infty(\Omega)} = \inf\{C: |u(\mathbf{x})| \leq C \text{ s.s.}\}.$$

(s.s. znači skoro svuda u odnosu na Lebesgueovu mjeru, tj. svugdje osim eventualno na skupu mjere nula.) Uvijek imamo

$$|u(\mathbf{x})| \leq \|u\|_{L^\infty} \text{ s.s.}$$

Ako je f neprekidna funkcija, onda je

$$\|u\|_{L^\infty(\Omega)} = \sup_{\mathbf{x} \in \Omega} |u(\mathbf{x})|.$$

Istaknimo sljedeću činjenicu: Prostori $L^1(\Omega)$ i $L^\infty(\Omega)$ su Banchovi uz uvedene norme, a prostor $L^2(\Omega)$ je Hilbertov u skalarnom produktu

$$(u, v)_{L^2(\Omega)} = \int_{\Omega} u(\mathbf{x})v(\mathbf{x}) d\mathbf{x}.$$

Često ćemo koristiti sljedeći rezultat:

Lema 2.2 Hölderova nejednakost. Neka su $u, v \in L^2(\Omega)$. Tada je $uv \in L^1(\Omega)$ i vrijedi

$$\int_{\Omega} |u(\mathbf{x})v(\mathbf{x})| d\mathbf{x} \leq \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}.$$

Prostore Soboljeva definiramo generalizacijom pojma parcijalne derivacije.

Definicija 2.1 Za funkciju $u \in L^2(\Omega)$ kažemo da ima slabu parcijalnu derivaciju po varijabli x_i , ako postoji funkcija $g \in L_{\text{loc}}^1(\Omega)$, takva da za svako $\phi \in C_c^1(\Omega)$ vrijedi²

$$\int_{\Omega} u(\mathbf{x}) \frac{\partial \phi}{\partial x_i}(\mathbf{x}) d\mathbf{x} = - \int_{\Omega} g(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}.$$

Funkcija g je tada slaba parcijalna derivacija funkcije u po x_i pa koristimo standardnu oznaku

$$g = \frac{\partial u}{\partial x_i}.$$

²Funkcije iz $C_c^1(\Omega)$ imaju kompaktni nosač, pa je stoga integral na desnoj strani dobro definiran već za lokalno integrabilnu funkciju (vidi Poglavlje A).

Iz ove definicije je jasno da je klasična parcijalna derivacija, ako postoji, ujedno i slaba derivacija. Nadalje, lako se pokazuje da je slaba derivacija jedinstvena (vidi Poglavlje A). Stoga je ovaj pojam dobro proširenje klasičnog pojma derivacije. U tekstu koji slijedi sve parcijalne derivacije se uzimaju u slabom smislu.

Sada možemo definirati

$$H^1(\Omega) = \{u \in L^2(\Omega) : \frac{\partial u}{\partial x_i} \in L^2(\Omega), 1 \leq i \leq d\}.$$

Prostor $H^1(\Omega)$ je očito linearan i normu u njemu definiramo formulom

$$\|u\|_{H^1(\Omega)} = \left(\int_{\Omega} |u(\mathbf{x})|^2 d\mathbf{x} + \sum_{i=1}^d \int_{\Omega} \left| \frac{\partial u}{\partial x_i}(\mathbf{x}) \right|^2 d\mathbf{x} \right)^{1/2}.$$

Ta je norma evidentno generirana skalarnim produktom

$$(u, v)_{H^1(\Omega)} = \int_{\Omega} u(\mathbf{x})v(\mathbf{x}) d\mathbf{x} + \sum_{i=1}^d \int_{\Omega} \frac{\partial u}{\partial x_i}(\mathbf{x}) \frac{\partial v}{\partial x_i}(\mathbf{x}) d\mathbf{x}.$$

Prostor $H^1(\Omega)$ se naziva prostor Soboljeva i može se dokazati:

Teorem 2.2 Prostor $H^1(\Omega)$ je Hilbertov.

Konačno, uvedimo potprostor $H_0^1(\Omega) \subset H^1(\Omega)$ kao zatvarač (upotpunjenje) prostora $C_0^1(\Omega)$ u normi prostora $H^1(\Omega)$. To je dakle skup svih funkcija iz $H^1(\Omega)$ koje se mogu dobiti kao limes funkcija iz $C_0^1(\Omega)$. To je evidentno linearan prostor i kao zatvoren potprostor potpunog prostora i sam je potpun. Iz konstrukcije je jasno da se $H_0^1(\Omega)$ sastoji od onih funkcija iz $H^1(\Omega)$ koje se poništavaju na rubu domene $\partial\Omega$ ([link na teorem o tragu](#)).

Definicija poopćenih prvih parcijalnih derivacija lako se generalizira na derivacije proizvoljnog reda. Tako na primjer, kažemo da funkcija $u \in L^2(\Omega)$ ima slabu parcijalnu derivaciju

$$\frac{\partial^2 u}{\partial x_i \partial x_j}$$

ako postoji funkcija $g \in L_{\text{loc}}^1(\Omega)$, takva da za svako $\phi \in C_c^2(\Omega)$ vrijedi

$$\int_{\Omega} u(\mathbf{x}) \frac{\partial^2 \phi}{\partial x_i \partial x_j}(\mathbf{x}) d\mathbf{x} = \int_{\Omega} g(\mathbf{x}) \phi(\mathbf{x}) d\mathbf{x}.$$

Negativan predznak je nestao zbog “dvostruke primjene formule parcijalne integracije”. Prostor Soboljevleva drugog reda definiramo kao

$$H^2(\Omega) = \{u \in H^1(\Omega) : \frac{\partial^2 u}{\partial x_i \partial x_j} \in L^2(\Omega), 1 \leq i, j \leq d\}.$$

To je ponovo Hilbertov prostor u skalarnom produktu

$$(u, v)_{H^2(\Omega)} = \int_{\Omega} u(\mathbf{x})v(\mathbf{x}) d\mathbf{x} + \sum_{i=1}^d \int_{\Omega} \frac{\partial u}{\partial x_i}(\mathbf{x}) \frac{\partial v}{\partial x_i}(\mathbf{x}) d\mathbf{x} + \sum_{i,j=1}^d \int_{\Omega} \frac{\partial^2 u(\mathbf{x})}{\partial x_i \partial x_j} \frac{\partial^2 v(\mathbf{x})}{\partial x_i \partial x_j} d\mathbf{x}.$$

Za definiciju prostora $H^k(\Omega)$ vidi Dodatak.

2.3 Homogena Dirichletova zadaća

Neka je L eliptički diferencijalni operator drugog reda u divergentnom obliku:

$$Lu = - \sum_{i,j=1}^d \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) + a_0 u, \quad (2.10)$$

pri čemu je matrica koeficijenata $\mathbf{A}(\mathbf{x}) = (a_{ij}(\mathbf{x}))_{i,j=1}^d$ simetrična i pozitivno definitna u svakoj točki $\mathbf{x} \in \Omega$, te vrijedi

$$a_0(\mathbf{x}) \geq 0, \quad \forall \mathbf{x} \in \Omega.$$

Promatramo homogenu Dirichletovu rubnu zadaću

$$\begin{aligned} Lu &= f & \text{u } \Omega \\ u &= 0 & \text{na } \partial\Omega. \end{aligned}$$

Za klasično rješenje u i test funkciju $v \in H_0^1(\Omega)$ vrijedi

$$\int_{\Omega} \left[- \sum_{i,j=1}^d \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) + a_0 u \right] v d\mathbf{x} = \int_{\Omega} f v d\mathbf{x}.$$

Parcijalnom integracijom dobivamo

$$- \sum_{i,j=1}^d \int_{\Omega} \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j} v) d\mathbf{x} + \int_{\Omega} \left[\sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + a_0 uv \right] d\mathbf{x} = \int_{\Omega} f v d\mathbf{x}.$$

Zbog toga što se test funkcije v poništava na rubu domene ($\mathbf{n} = (n_i)$ je vanjska normala na $\partial\Omega$) imamo

$$\sum_{i,j=1}^d \int_{\Omega} \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j} v) d\mathbf{x} = \sum_{i,j=1}^d \int_{\partial\Omega} a_{ij} \frac{\partial u}{\partial x_j} n_i v d\sigma = 0.$$

Stoga je klasično rješenje u istovremeno i rješenje varijacijske zadaće

$$a(u, v) = \langle F, v \rangle, \quad \forall v \in H_0^1(\Omega),$$

gdje je

$$a(u, v) = \int_{\Omega} \left[\sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + a_0 uv \right] d\mathbf{x} \quad (2.11)$$

$$\langle F, v \rangle = \int_{\Omega} f v d\mathbf{x}. \quad (2.12)$$

Definicija 2.2 Neka je L diferencijalni operator definiran u (2.10). Funkcija $u \in H_0^1(\Omega)$ je slabo rješenje homogene Dirichletove zadaće

$$Lu = f \quad \text{u } \Omega \quad (2.13)$$

$$u = 0 \quad \text{na } \partial\Omega, \quad (2.14)$$

ako vrijedi

$$\forall v \in H_0^1(\Omega), \quad a(u, v) = \langle F, v \rangle,$$

gdje je $a(\cdot, \cdot)$ pridružena bilinearna forma definirana s (2.11), a desna strana je dana s (2.12).

Kao i u uvodnom primjeru pokazuje se da je svako slabo rješenje koje leži u prostoru $C^2(\bar{\Omega}) \cap C_0(\Omega)$ ujedno i klasično rješenje.

Precizirajmo sada pretpostavke o operatoru L koje omogućavaju Lax-Milgramove leme:

$$\forall i, j, \quad a_{ij} \in L^\infty(\Omega), \quad a_0 \in L^\infty(\Omega), \quad f \in L^2(\Omega), \quad (2.15)$$

$$\forall \boldsymbol{\xi} \in \mathbb{R}^d, \quad \forall \mathbf{x} \in \Omega, \quad \sum_{i,j=1}^d a_{ij}(\mathbf{x}) \xi_i \xi_j \geq \alpha \sum_{i=1}^d \xi_i^2. \quad (2.16)$$

$$\forall \mathbf{x} \in \Omega, \quad \text{matrica } \mathbf{A}(\mathbf{x}) = (a_{ij}(\mathbf{x})) \text{ je simetrična,} \quad (2.17)$$

$$\forall \mathbf{x} \in \Omega, \quad a_0(\mathbf{x}) \geq \alpha_1 > 0. \quad (2.18)$$

Prije iskaza teorema uvedimo u $H^1(\Omega)$ sljedeću *polunormu*

$$|u|_{H^1(\Omega)} = \left(\sum_{i=1}^d \int_{\Omega} \left| \frac{\partial u(\mathbf{x})}{\partial x_i} \right|^2 d\mathbf{x} \right)^{1/2}. \quad (2.19)$$

Funkcija $|\cdot|$ je polunorma na nekom vektorskom prostoru ako ima sva svojstva norme osim svojstva da $|u| = 0$ implicira $u = 0$; iz $|u|_{H^1(\Omega)} = 0$ slijedi da je u konstantna funkcija na Ω .

Teorem 2.3 Neka su zadovoljene pretpostavke (2.15)–(2.18). Tada Dirichletova zadaća (2.13)–(2.14) ima jedinstveno slabo rješenje. Ono minimizira funkcional

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

na prostoru $H_0^1(\Omega)$.

Dokaz. Potrebno je provjeriti uvjete Lax-Milgramove leme. Iz simetrije matrice koeficijenata lako slijedi simetrija bilinearne forme. Da bismo provjerili ograničenost uvedimo konstantu $M = \max\{\|a_{ij}\|_{L^\infty}, \|a_0\|_{L^\infty}\}$. Imamo

$$\begin{aligned} \left| \sum_{i,j=1}^d \int_{\Omega} a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} d\mathbf{x} \right| &\leq M \sum_{i,j=1}^d \int_{\Omega} \left| \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} \right| d\mathbf{x} \\ &\leq M \sum_{i,j=1}^d \left(\int_{\Omega} \left| \frac{\partial u}{\partial x_i} \right|^2 d\mathbf{x} \right)^{1/2} \left(\int_{\Omega} \left| \frac{\partial v}{\partial x_j} \right|^2 d\mathbf{x} \right)^{1/2} \\ &\leq Md \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \end{aligned}$$

gdje smo iskoristili nejednakost (2.108) (vidi zadatke). Nadalje,

$$\left| \int_{\Omega} a_0 uv d\mathbf{x} \right| \leq M \int_{\Omega} |uv| d\mathbf{x} \leq M \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}.$$

Kombinirajući ove dvije ocjene, dobivamo

$$a(u, v) \leq Md \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} + M \|u\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq C \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)},$$

gdje je konstanta C jednaka Md . Iz pozitivne definitnosti matrice koeficijenata slijedi

$$\sum_{i,j=1}^d a_{ij}(\mathbf{x}) \frac{\partial v}{\partial x_i} \frac{\partial v}{\partial x_j} \geq \alpha \sum_{i=1}^d \left| \frac{\partial v}{\partial x_i} \right|^2. \quad (2.20)$$

Integriranjem dobivamo

$$a(v, v) \geq \min(\alpha, \alpha_1) \int_{\Omega} \left(\sum_{i=1}^d \left| \frac{\partial v}{\partial x_i} \right|^2 + v^2 \right) d\mathbf{x} = \min(\alpha, \alpha_1) \|v\|_{H^1(\Omega)}^2. \quad (2.21)$$

Linearan funkcional F je neprekidan budući da vrijedi

$$\left| \int_{\Omega} fv d\mathbf{x} \right| \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)}.$$

Tvrđnja teorema sada slijedi primjenom Lax-Milgramove leme i leme o karakterizaciji varijacijske zadaće (Lema 2.5). \square

Pretpostavka (2.18) može se zamijeniti slabijom pretpostavkom

$$\forall \mathbf{x} \in \Omega, \quad a_0(\mathbf{x}) \geq 0, \quad (2.22)$$

a da Teorem 2.3 ostane vrijediti.

Teorem 2.4 (Poincaréova nejednakost) Ako je domena $\Omega \subset \mathbb{R}^d$ ograničena barem u jednom smjeru, onda postoji konstanta $c = c(\Omega)$ takva da za svako $u \in H_0^1(\Omega)$ vrijedi

$$\|u\|_{L^2(\Omega)} \leq c \|u\|_{H^1(\Omega)}. \quad (2.23)$$

Dokaz. Bez smanjenja općenitosti možemo uzeti da se domena nalazi unutar pruge $a < x_d < b$. Uzmimo proizvoljnu funkciju $v \in H_0^1(\Omega)$ i proširimo ju nulom izvan skupa Ω . Imamo

$$v(\mathbf{x}', x_d) = \int_a^{x_d} \frac{\partial v}{\partial x_d}(\mathbf{x}', t) dt.$$

Cauchyjeva nejednakost daje

$$|v(\mathbf{x}', x_d)|^2 \leq (x_d - a) \int_a^{x_d} \left| \frac{\partial v}{\partial x_d}(\mathbf{x}', t) \right|^2 dt,$$

pa integriranjem po prvih $d - 1$ varijabli slijedi

$$\begin{aligned} \int_{\mathbb{R}^{d-1}} |v(\mathbf{x}', x_d)|^2 d\mathbf{x}' &\leq (x_d - a) \int_a^{x_d} \int_{\mathbb{R}^{d-1}} \left| \frac{\partial v}{\partial x_d}(\mathbf{x}', t) \right|^2 d\mathbf{x}' dt. \\ \int_a^b \int_{\mathbb{R}^{d-1}} |v(\mathbf{x}', x_d)|^2 d\mathbf{x}' dx_d &\leq \frac{1}{2}(b - a)^2 \int_a^b \int_{\mathbb{R}^{d-1}} \left| \frac{\partial v}{\partial x_d}(\mathbf{x}', t) \right|^2 d\mathbf{x}' dt. \end{aligned}$$

Budući da je funkcija jednaka nuli izvan Ω slijedi

$$\int_{\Omega} |v(\mathbf{x})|^2 d\mathbf{x} \leq \frac{1}{2}(b - a)^2 \int_{\Omega} \left| \frac{\partial v}{\partial x_d}(\mathbf{x}) \right|^2 d\mathbf{x}. \quad \square$$

Promjenom Poincaréove nejednakosti dobivamo sljedeću ekvivalenciju normi na $H_0^1(\Omega)$:

$$\forall u \in H_0^1(\Omega), \quad \frac{1}{\sqrt{1 + c^2}} \|u\|_{H^1(\Omega)} \leq |u|_{H^1(\Omega)} \leq \|u\|_{H^1(\Omega)}.$$

Sada lako vidimo da (2.21) možemo uz pretpostavku (2.39) izvesti na ovaj način:

$$a(v, v) \geq \alpha \int_{\Omega} \left(\sum_{i=1}^d \left| \frac{\partial v}{\partial x_i} \right|^2 \right) d\mathbf{x} = \alpha |v|_{H^1(\Omega)}^2 \geq \frac{\alpha}{1 + c^2} \|u\|_{H^1(\Omega)}^2.$$

Time je koercitivnost bilinearne forme dokazana i Lax-Milgramova lema se može ponovo primijeniti.

2.4 Nehomogena Dirichletova zadaća

Promatramo nehomogenu Dirichletovu rubnu zadaću

$$\begin{aligned} Lu &= f & \text{u } \Omega \\ u &= g & \text{na } \partial\Omega. \end{aligned}$$

Da bi postojalo klasično rješenje te zadaće funkcija g , koja je zadana samo na rubu $\partial\Omega$, mora imati barem jedno proširenje u_0 na $\bar{\Omega}$ koje je klase C^2 . U tom se slučaju zadaća može svesti na zadaću s homogenim rubnim uvjetom postupkom koji zovemo *homogenizacija rubnog uvjeta*, a sastoji se u tome da se prijeđe na ekvivalentnu zadaću za funkciju $w = u - u_0$:

$$\begin{aligned} Lw &= f_1 & \text{u } \Omega \\ w &= 0 & \text{na } \partial\Omega. \end{aligned}$$

gdje je $f_1 = f - Lu_0$; traženo rješenje je $u = w + u_0$. Prema tome, barem teorijski, uvijek možemo homogenizirati rubni uvjet te je stoga dovoljno promatrati samo homogenu rubnu zadaću.

Postupak homogenizacije rubnog uvjeta možemo provesti i na slabom rješenju zadaće. Prvo uočimo da postupak izvoda slabe formulacije ne ovisi o homogenosti ili nehomogenosti Dirichletovog rubnog uvjeta jer test funkciju uzimamo jednaku nuli na rubu u oba slučaja. Varijacijska formulacija stoga prima sljedeći oblik:

$$\begin{cases} \text{Naći } u \in H^1(\Omega), & u = g \text{ na } \partial\Omega \\ \forall v \in H_0^1(\Omega), & a(u, v) = \langle F, v \rangle. \end{cases} \quad (2.24)$$

Jedina je razlika u tome što rješenje više ne tražimo u prostoru test funkcija, $H_0^1(\Omega)$, već u *linearnoj mnogostrukosti*

$$g + H_0^1(\Omega) = \{v \in H^1(\Omega) : v = g \text{ na } \partial\Omega\}.$$

To ima za posljedicu da ne možemo neposredno primijeniti Lax-Milgramovu lemu već se prvo mora homogenizirati rubni uvjet.

Da bi varijacijska zadaća (2.24) imala rješenje nužno moramo pretpostaviti da postoji barem jedna funkcija $u_g \in H^1(\Omega)$ takva da je $u_g = g$ na $\partial\Omega$. Funkcija $w = u - u_g \in H_0^1(\Omega)$ tada zadovoljava

$$\begin{cases} \text{Naći } w \in H_0^1(\Omega), \\ \forall v \in H_0^1(\Omega), & a(w, v) = \langle F, v \rangle - a(u_g, v), \end{cases} \quad (2.25)$$

To je ponovo varijacijska zadaća na koju možemo primijeniti Lax-Milgramovu lemu. Jedini novi element je provjeriti da je desna strana linearan i neprekidan funkcional na $H_0^1(\Omega)$, što slijedi iz ocjene:

$$\begin{aligned} |\langle F, v \rangle - a(u_g, v)| &\leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + M \|u_g\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \\ &\leq (\|f\|_{L^2(\Omega)} + M \|u_g\|_{H^1(\Omega)}) \|v\|_{H^1(\Omega)}. \end{aligned}$$

Napomena 2.1 Uočimo da konstanta $K = (\|f\|_{L^2} + M\|u_g\|_{H^1(\Omega)})$, koja se javlja u ovoj ocjeni, ne ovisi o načinu proširenja funkcije g u unutrašnjost domene stoga što u desnoj strani ocjene možemo uzeti infimum po svim takvim proširenjima.

Kako zadaća (2.25) ima jedinstveno rješenje $w \in H_0^1(\Omega)$ zaključujemo da i polazna nehomogena zadaća (2.24) ima rješenje $u = u_g + w \in H^1(\Omega)$. Jedinstvnost rješenja nehomogene zadaće slijedi iz toga što se razlika bilo koja dva rješenja $w = u_1 - u_2$ zadaće (2.24) nalazi u $H_0^1(\Omega)$ i zadovoljava homogenu varijacijsku jednadžbu

$$\forall v \in H_0^1(\Omega), \quad a(w, v) = 0.$$

Time smo dokazali

Teorem 2.5 Neka su ispunjeni uvjeti Teorema 2.3 i neka je Ω Lipschitzova domena. Tada za svaku funkciju $g \in L^2(\partial\Omega)$ koja ima barem jedno proširenje $G \in H^1(\Omega)$, $G|_{\partial\Omega} = g$, varijacijska zadaća (2.24) ima jedinstveno rješenje u $H^1(\Omega)$.

Nehomogena Dirichletova zadaća nas vodi do prostora funkcija definiranih na $\partial\Omega$ koji definiramo na ovaj način:

$$H^{1/2}(\partial\Omega) = \{\phi: \partial\Omega \rightarrow \mathbb{R}, \exists \Phi \in H^1(\Omega), \Phi|_{\partial\Omega} = \phi\}.$$

Linearnost ovakvog skupa funkcija lako se provjerava. S druge strane, restrikcija funkcije iz $H^1(\Omega)$ na rub domene $\partial\Omega$ ne može se neposredno definirati budući da su u $H^1(\Omega)$ funkcije koje su definirane samo na Ω i tamo su kvadratno integrabilne, zajedno sa svojim prvim slabim derivacijama. Ipak se može pokazati da se vrijednost funkcije na rubu može jednoznačno definirati za svaku funkciju iz $H^1(\Omega)$ ukoliko domena ima određenu *glatkoću*, na primjer ako je Lipschitzova (o tome više u Dodatku A), pa je stoga to uvjet Teorema 2.5. Restrikcija funkcije iz prostora Soboljeva na rub domene dobila je naziv *trag funkcije*, a operator $\gamma = \gamma_{\partial\Omega}$ koji pridružuje funkciji njen trag naziva se *operator traga*. Prostor $H^{1/2}(\partial\Omega)$ je u toj terminologiji prostor tragova funkcija iz $H^1(\Omega)$ i u njega se uvodi norma

$$\|\phi\|_{H^{1/2}(\partial\Omega)} = \inf_{\gamma\Phi = \phi} \|\Phi\|_{H^1(\Omega)}.$$

Ovdje se infimum uzima po svim funkcijama $\Phi \in H^1(\Omega)$ čiji trag na $\partial\Omega$ jednak ϕ . Operator traga

$$\gamma: H^1(\Omega) \rightarrow H^{1/2}(\partial\Omega)$$

je evidentno linearna surjekcija i očito je neprekidan u ovoj normi.

Nadalje, u teoriji prostora Soboljeva se pokazuje da je $H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega)$ i da je to pravi podskup. Kanonsko ulaganje je pri tome neprekidno, što znači da postoji konstanta $C > 0$ takva da je

$$\forall v \in H^{1/2}(\partial\Omega), \quad \|v\|_{L^2(\partial\Omega)} \leq C\|v\|_{H^{1/2}(\partial\Omega)}.$$

Drugim riječima Teorem 2.5 govori da je nehomogena Dirichletova zadaća rješiva za svaki rubni uvjet $g \in H^{1/2}(\partial\Omega)$, ali ne i za svako $g \in L^2(\partial\Omega)$.

Sumirajmo do sada rečeno.

Teorem 2.6 (Teorem o tragu) Neka je Ω Lipschitzova domena. Tada preslikavanje $\gamma_{\partial\Omega}$ koje svakoj funkciji $u \in C^1(\bar{\Omega})$ pridružuje njenu restrikciju na $\partial\Omega$ ima jedinstveno proširenje do linearnog i neprekidnog operatora

$$\gamma_{\partial\Omega}: H^1(\Omega) \rightarrow L^2(\partial\Omega).$$

Operator $\gamma_{\partial\Omega}$ se naziva operator traga i vrijedi $H_0^1(\Omega) = \{v \in H^1(\Omega): \gamma_{\partial\Omega}v = 0\}$. Slika operatora traga je prostor $H^{1/2}(\partial\Omega)$.

Uočimo da neprekidnost operatora traga znači da postoji konstanta $C > 0$ takva da za svako $v \in H^1(\Omega)$ vrijedi

$$\|\gamma_{\partial\Omega}v\|_{L^2(\partial\Omega)} \leq C\|v\|_{H^1(\Omega)}.$$

Nadalje, kako je operator traga tek generalizacija operatora restrikcije, nećemo ga eksplicitno pisati već ćemo ga *podrazumijevati*. Tako ćemo, na primjer, gornju nejednakost jednostavno pisati

$$\forall v \in H^1(\Omega), \quad \|v\|_{L^2(\partial\Omega)} \leq C\|v\|_{H^1(\Omega)}.$$

Napomena 2.2 *Homogenizacija rubnog uvjeta je nužna da bi se pokazalo da nehomogena zadaća ima rješenje. S druge strane, u numeričkom rješavanju homogenizacija rubnog uvjeta najčešće nije izvediva u varijacijskoj formulaciji već se sličan postupak radi na razini diskretne zadaće. [Vidi prvo poglavlje ...]*

2.5 Neumannova i mješovita zadaća

U ovoj sekciji formiramo slabo rješenje zadaće

$$\begin{cases} -\sum_{i,j=1}^d \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) + a_0 u = f & \text{u } \Omega \\ \sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} n_i = g & \text{na } \partial\Omega, \end{cases} \quad (2.26)$$

u kojoj smo postavili *Neumannov* odnosno *prirodni* rubni uvjet. Pretpostavljat ćemo da koeficijenti diferencijalne jednadžbe zadovoljavaju uvjete (2.15)–(2.18).

Postupak izvođenja varijacijske jednadžbe je isti kao i za Dirichletovu zadaću, s tom razlikom da se test funkcija ne mora poništavati na granici domene. Stoga množimo jednadžbu (2.26)₁ s test funkcijom $v \in H^1(\Omega)$ i integriramo po domeni:

$$\int_{\Omega} \left[-\sum_{i,j=1}^d \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) + a_0 u \right] v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x}.$$

Parcijalnom integracijom u prvom članu dobivamo kao i ranije

$$-\sum_{i,j=1}^d \int_{\partial\Omega} a_{ij} \frac{\partial u}{\partial x_j} n_i v \, d\sigma + \int_{\Omega} \left[\sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + a_0 uv \right] d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x}.$$

Funkcija koja se pojavljuje u integralu po rubu domene upravo nam je zadana pa iz toga proizlazi naziv *prirodan rubni uvjet*.

Definicija 2.3 Funkcija $u \in H^1(\Omega)$ je slabo rješenje zadaće (2.26) ako zadovoljava

$$\forall v \in H^1(\Omega), \quad \int_{\Omega} \left[\sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + a_0 uv \right] d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} + \int_{\partial\Omega} g v \, d\sigma. \quad (2.27)$$

Uočimo da Neumannov rubni uvjet nije *ugrađen u prostor*, već je implicitno zadovoljen varijacijskom jednadžbom. U to ćemo se uvjeriti tako da pretpostavimo da je domena Ω dovoljno glatka, da su svi koeficijenti jednadžbe najmanje neprekidni na $\bar{\Omega}$, te da varijacijska jednadžba ima rješenje $u \in C^2(\bar{\Omega})$. Tada u jednadžbi (2.27) možemo izvršiti parcijalnu integraciju u prvom članu (ponavljajući već izvršeni račun u suprotnom smjeru),

$$\begin{aligned} \int_{\Omega} \left[\sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + a_0 uv \right] d\mathbf{x} &= \sum_{i,j=1}^d \int_{\partial\Omega} \left(a_{ij} \frac{\partial u}{\partial x_j} n_i \right) v \, d\sigma \\ &\quad + \int_{\Omega} \left[- \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) + a_0 u \right] v \, d\mathbf{x}, \end{aligned}$$

iz (2.27) dobivamo

$$\int_{\partial\Omega} \left(\sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} n_i - g \right) v \, d\sigma + \int_{\Omega} \left[- \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) + a_0 u - f \right] v \, d\mathbf{x} = 0,$$

gdje je $\mathbf{n} = (n_i)$ jedinična vanjska normala na $\partial\Omega$. Uzimajući test funkcije koje se poništavaju na $\partial\Omega$ dobivamo da za svaku takvu vrijedi

$$\int_{\Omega} \left[- \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) + a_0 u - f \right] v \, d\mathbf{x} = 0.$$

Odavde se lako pokazuje, po osnovnoj lemi varijacijskog računa (Propozicija 2.1), da je diferencijalna jednadžba zadovoljena u svakoj točki domene Ω , pa onda integral propada za svaki izbor test funkcije (a ne samo one koja se poništava na rubu). Iz prethodne jednakosti stoga dobivamo

$$\int_{\partial\Omega} \left(\sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} n_i - g \right) v \, d\sigma = 0.$$

Odavde slijedi, ponovo prema verziji osnovne leme varijacijskog računa za plošni integral,

$$\sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} n_i = g \quad \text{na } \partial\Omega.$$

Time smo pokazali da je Neumannov rubni uvjet implicitno sadržan u varijacijskoj jednadžbi i ona stoga predstavlja dobru generalizaciju diferencijalne rubne zadaće.

Preostaje nam pokazati da na zadaću (2.27) možemo primijeniti Lax-Milgramovu lemu. Budući da je bilinearna forma ista kao u slučaju Dirichletove zadaće prethodni računi pokazuju ograničenost i koercitivnost forme na $H^1(\Omega)$ (vidi (2.21)).

Ostaje pokazati neprekidnost linearnog funkcionala. U tu svrhu treba pretpostaviti da je $f \in L^2(\Omega)$ i

$$g \in L^2(\partial\Omega). \quad (2.28)$$

Tada za funkcional

$$\langle F, v \rangle = \int_{\Omega} f v \, d\mathbf{x} + \int_{\partial\Omega} g v \, d\sigma, \quad (2.29)$$

koristeći teorem o tragu (Teorem 2.6), dobivamo ocjenu

$$\begin{aligned} |\langle F, v \rangle| &\leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + \|g\|_{L^2(\partial\Omega)} \|v\|_{L^2(\partial\Omega)} \\ &\leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + C \|g\|_{L^2(\partial\Omega)} \|v\|_{H^1(\Omega)} \\ &\leq (\|f\|_{L^2(\Omega)} + C \|g\|_{L^2(\partial\Omega)}) \|v\|_{H^1(\Omega)}. \end{aligned}$$

Sada Lax-Milgramova lema i Lema 2.5 daju sljedeći zaključak:

Teorem 2.7 Neka je Ω Lipschitzova domena i neka su zadovoljene pretpostavke (2.15)—(2.18) i (2.28). Tada Neumannova zadaća (2.26) ima jedinstveno slabo rješenje koje minimizira funkcional

$$J(v) = \frac{1}{2} a(v, v) - \int_{\Omega} f v \, d\mathbf{x} - \int_{\partial\Omega} g v \, d\sigma$$

na prostoru $H^1(\Omega)$.

Teorem 2.7 ne možemo primijeniti na problem

$$\begin{cases} -\Delta u = f & \text{u } \Omega \\ \frac{\partial u}{\partial \mathbf{n}} = g & \text{na } \partial\Omega. \end{cases} \quad (2.30)$$

budući da pripadna bilinearna forma

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x}$$

nije koecitivna na $H^1(\Omega)$. U (2.30)₂ koristimo oznaku

$$\frac{\partial u}{\partial \mathbf{n}} = \nabla u \cdot \mathbf{n} = \sum_{i=1}^d \frac{\partial u}{\partial x_i} n_i.$$

Uočimo da je klasično rješenje zadatke (2.30) određeno do na aditivnu konstantu, pa dakle nije jedinstveno i stoga je prirodno da Lax-Milgramova lema ne može biti primijenjena. Nadalje, lako je uočiti da postoji nužan uvjet egzistencije rješenja koji glasi

$$\int_{\Omega} f \, d\mathbf{x} + \int_{\partial\Omega} g \, d\sigma = 0, \quad (2.31)$$

i koji se dobiva integracijom diferencijalne jednadžbe (2.30)₁ po Ω .

Rješenje zadatke (2.30) možemo učiniti jedinstvenim tako da ga tražimo u užem funkcijskom prostoru $V \subset H^1(\Omega)$, iz kojeg su eliminirane sve konstante osim nule. Takav prostor možemo definirati na sljedeći način:

$$V = \{\phi \in H^1(\Omega) : \int_{\Omega} \phi(\mathbf{x}) \, d\mathbf{x} = 0\}. \quad (2.32)$$

Na njemu je bilinearna forma

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x}$$

ponovo koercitivna. To slijedi iz ove varijante Poincaréove nejednakosti:

Lema 2.3 Neka je Ω ograničena Lipschitzova domena. Tada postoji konstanta C , takva da za svako $v \in H^1(\Omega)$ vrijedi

$$\|v\|_{L^2(\Omega)} \leq C(|v|_{H^1(\Omega)} + \left| \int_{\Omega} v(\mathbf{x}) \, d\mathbf{x} \right|).$$

Za dokaz vidi Evans [2].

Na osnovu ove leme lako se pokazuje da je $v \mapsto |v|_{H^1(\Omega)}$ norma na V ekvivalentna s $H^1(\Omega)$ normom, pa se stoga može primijeniti Lax-Milgramova lema na problem: naći $u \in V$, takvo da

$$\forall v \in V, \quad a(u, v) = \int_{\Omega} f v \, d\mathbf{x} + \int_{\partial\Omega} g v \, d\sigma, \quad (2.33)$$

koji stoga ima jedinstveno rješenje $u \in V$. Da bi to rješenje bilo klasično, mora biti zadovoljen, pored glatkoće rješenja, i nužan uvjet egzistencije (2.31).

Pretpostavimo, konačno, da imamo mješovitu zadaću. Granicu $\partial\Omega$ ćemo podijeliti u dva disjunktna podskupa

$$\partial\Omega = \overline{\Gamma_D} \cup \overline{\Gamma_N}, \quad \Gamma_D \cap \Gamma_N = \emptyset.$$

Na Γ_D zadajemo Dirichletov, a na Γ_N Neumannov rubni uvjet. Na primjer,

$$\begin{cases} -\Delta u = f & \text{u } \Omega \\ u = 0 & \text{na } \Gamma_D \\ \frac{\partial u}{\partial \mathbf{n}} = g & \text{na } \Gamma_N. \end{cases}$$

Da bi se definirala slaba formulacija ove zadaće prvo se definira operator traga na *dio* granice Γ_D , koji ćemo označavati s γ_{Γ_D} . Taj operator ima svojstva analogna svojstvima operatora traga na čitavu granicu $\partial\Omega$. Zatim se pomoću njega definira funkcijski prostor V , u kome tražimo slabo rješenje:

$$H_{\Gamma_D}^1(\Omega) = \{v \in H^1(\Omega) : \gamma_{\Gamma_D} v = 0\}.$$

To su dakle funkcije iz $H^1(\Omega)$ koje se poništavaju na Γ_D . Uz uvjet da je površinska mjera skupa Γ_D strogo pozitivna može pokazatuje se da za funkcije iz $V = H_{\Gamma_D}^1(\Omega)$ vrijedi Poincaréova nejednakost. Stoga je $|\cdot|_{H^1(\Omega)}$ je norma na V ekvivalentna s $H^1(\Omega)$ normom. Formulacija slabog rješenja i primjena Lax-Milgramove leme sada slijedi analogno kao u prethodnim primjerima: treba naći $u \in H_{\Gamma_D}^1(\Omega)$ takvo da je

$$\forall v \in H_{\Gamma_D}^1(\Omega), \quad \int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} + \int_{\Gamma_N} g v \, d\sigma.$$

Ako je *Dirichletov rubni uvjet nehomogen*, na primjer $u = u_0$ na Γ_D , onda je nužno da postoji funkcija $U \in H^1(\Omega)$ koja se podudara s u_0 na Γ_D (u smislu da je $\gamma_{\Gamma_D} U = u_0$) i koja *ne kviri* Neumannov rubni uvjet na Γ_D , tj. takva za koju je

$$\frac{\partial U}{\partial \mathbf{n}} = 0 \quad \text{na } \Gamma_D.$$

Kad god takvu funkciju možemo naći, onda mješovita zadaća ima jedinstveno slabo rješenje u skupu $U + H_{\Gamma_D}^1(\Omega)$.

2.6 Nesimetrične zadaće

Diferencijalni operatori koje smo do sada promatrali vode na simetričnu bilinearnu formu. Varijacijska zadaća za takvu formu uvijek odgovara problemu minimizacije pripadnog funkcionala. Ako forma nije simetrična, onda to nije slučaj, što je lako provjeriti. Ipak tvrdnja Lax-Milgramove leme ostaje vrijediti.

Pogledajmo jedan primjer operatora koji vodi na nesimetričnu varijacijsku zadaću.

$$-\sum_{i,j=1}^d \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) + \sum_{i=1}^d b_i \frac{\partial u}{\partial x_i} + a_0 u = f \quad \text{u } \Omega \quad (2.34)$$

$$u = 0 \quad \text{na } \partial\Omega \quad (2.35)$$

Odabrali smo homogeni Dirichletov rubni uvjet, a novost u diferencijalnom operatoru je član s derivacijama prvog reda. Pretpostavke o koeficijentima jednadžbe su sljedeće:

$$\forall i, j, \quad a_{ij}, b_i, a_0 \in L^\infty(\Omega), \quad f \in L^2(\Omega), \quad (2.36)$$

$$\forall \boldsymbol{\xi} \in \mathbb{R}^d, \quad \forall \mathbf{x} \in \Omega, \quad \sum_{i,j=1}^d a_{ij}(\mathbf{x}) \xi_i \xi_j \geq \alpha \sum_{i=1}^d \xi_i^2. \quad (2.37)$$

$$\forall \mathbf{x} \in \Omega, \quad \text{matrica } \mathbf{A}(\mathbf{x}) = (a_{ij}(\mathbf{x})) \text{ je simetrična,} \quad (2.38)$$

$$\mathbf{b} = (b_i), \quad \text{div } \mathbf{b} \in L^\infty(\Omega). \quad (2.39)$$

Varijacijsku formulaciju izvodimo množenjem s test funkcijom $v \in H_0^1(\Omega)$ i integracijom po Ω ,

$$- \sum_{i,j=1}^d \int_{\Omega} \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) v \, d\mathbf{x} + \sum_{i=1}^d \int_{\Omega} b_i \frac{\partial u}{\partial x_i} v \, d\mathbf{x} + \int_{\Omega} a_0 u v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x},$$

a onda parcijalnom integracijom u prvom članu dovivamo

$$\sum_{i,j=1}^d \int_{\Omega} a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \, d\mathbf{x} + \sum_{i=1}^d \int_{\Omega} b_i \frac{\partial u}{\partial x_i} v \, d\mathbf{x} + \int_{\Omega} a_0 u v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x}.$$

Time dolazimo do varijacijskog problema: naći $u \in H_0^1(\Omega)$ takvo da je

$$\forall v \in H_0^1(\Omega), \quad a(u, v) = \langle F, v \rangle,$$

gdje je

$$a(u, v) = \sum_{i,j=1}^d \int_{\Omega} a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \, d\mathbf{x} + \sum_{i=1}^d \int_{\Omega} b_i \frac{\partial u}{\partial x_i} v \, d\mathbf{x} + \int_{\Omega} a_0 u v \, d\mathbf{x} \quad (2.40)$$

$$\langle F, v \rangle = \int_{\Omega} f v \, d\mathbf{x}. \quad (2.41)$$

Nova je forma evidentno bilinearna i nesimetrična, tj. općenito je $a(u, v) \neq a(v, u)$. Da bismo pokazali njenu neprekidnost, potrebno je samo ocijeniti dodatni član

$$\begin{aligned} \left| \sum_{i=1}^d \int_{\Omega} b_i \frac{\partial u}{\partial x_i} v \, d\mathbf{x} \right| &\leq \max_i \|b_i\|_{L^\infty} \sum_{i=1}^d \int_{\Omega} \left| \frac{\partial u}{\partial x_i} v \right| \, d\mathbf{x} \\ &\leq \max_i \|b_i\|_{L^\infty} \|u\|_{H^1(\Omega)} \|v\|_{L^2}. \end{aligned}$$

Uvedemo li oznaku $M = \max\{\|a_{ij}\|_{L^\infty}, \|a_0\|_{L^\infty}, \|b_i\|_{L^\infty}\}$ dobivamo

$$\begin{aligned} |a(u, v)| &\leq Md(\|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} + \|u\|_{H^1(\Omega)} \|v\|_{L^2} + \|u\|_{L^2} \|v\|_{L^2}) \\ &\leq 3Md \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}. \end{aligned}$$

Kod dokaza koercitivnosti treba ocijeniti samo drugi integral u formi $a(\cdot, \cdot)$ budući da imamo

$$\sum_{i,j=1}^d \int_{\Omega} a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial u}{\partial x_i} d\mathbf{x} \geq \alpha |u|_{H^1(\Omega)}^2.$$

To se može učiniti na razne načine. Uz pretpostavku $\operatorname{div} \mathbf{b} \in L^\infty(\Omega)$ imamo

$$\begin{aligned} \sum_{i=1}^d \int_{\Omega} b_i \frac{\partial u}{\partial x_i} u d\mathbf{x} &= \sum_{i=1}^d \frac{1}{2} \int_{\Omega} b_i \frac{\partial u^2}{\partial x_i} d\mathbf{x} \\ &= \sum_{i=1}^d \frac{1}{2} \int_{\Omega} \frac{\partial}{\partial x_i} (b_i u^2) d\mathbf{x} - \frac{1}{2} \int_{\Omega} \operatorname{div} \mathbf{b} u^2 d\mathbf{x} \\ &= -\frac{1}{2} \int_{\Omega} \operatorname{div} \mathbf{b} u^2 d\mathbf{x} \end{aligned}$$

gdje smo iskoristili rubni uvjet. Stoga dobivamo

$$a(u, u) \geq \alpha |u|_{H^1(\Omega)}^2 + \int_{\Omega} (a_0 - \frac{1}{2} \operatorname{div} \mathbf{b}) u^2 d\mathbf{x}.$$

Prema Poincaréovoj nejednakosti postoji konstanta C_Ω , koja ovisi samo o domeni Ω , takva da je

$$\forall v \in H_0^1(\Omega), \quad \|v\|_{L^2} \leq C_\Omega |v|_{H^1(\Omega)}. \quad (2.42)$$

Pretpostavimo sada da postoji broj η takav da vrijedi:

$$\forall \mathbf{x} \in \Omega, \quad a_0(\mathbf{x}) - \frac{1}{2} \operatorname{div} \mathbf{b}(\mathbf{x}) \geq -\eta. \quad (2.43)$$

Tada imamo ocjenu

$$a(u, u) \geq \alpha |u|_{H^1(\Omega)}^2 - \eta \|u\|_{L^2(\Omega)}^2 \geq (\alpha - \eta C_\Omega) |u|_{H^1(\Omega)}^2.$$

Time dolazimo do sljedećeg dovoljnog uvjeta koercitivnosti: ako postoji broj η takav da je

$$\alpha - \eta C_\Omega > 0, \quad (2.44)$$

i za koji vrijedi (2.43), onda je bilinearna forma (2.40) koercitivna na $H_0^1(\Omega)$.

Teorem 2.8 Neka su zadovoljeni uvjeti (2.36)–(2.39) te neka postoji konstanta $\eta \geq 0$ koja zadovoljava (2.43) i (2.44), gdje je C_Ω konstanta iz Poincaréove nejednakosti (2.42). Tada zadaća (2.34), (2.35) ima jedinstveno slabo rješenje.

Napomena 2.3 Jedan važan specijalan slučaj je $\operatorname{div} \mathbf{b} = 0$ i $a_0 \geq 0$.

Zadatak. Zadan je rubni problem

$$-\operatorname{div}(\mathbf{A}\nabla u + \mathbf{c}u) + \mathbf{b} \cdot \nabla u + a_0 u = f \quad u \text{ na } \Omega \quad (2.45)$$

$$u = 0 \quad \text{na } \partial\Omega, \quad (2.46)$$

pri čemu pretpostavljamo da su ispunjeni uvjeti (2.36)–(2.39), te da vektorska funkcija \mathbf{c} zadovoljava iste uvjete kao i \mathbf{b} (uniformnu ograničenost svih komponenti c_i i divergencije $\operatorname{div} \mathbf{c}$). Pokažite da na zadaću (2.45), (2.46) možemo primijeniti Lax-Milgramovu lemu ako postoji broj η takav da je

$$\alpha - \eta C_\Omega > 0, \quad (2.47)$$

i za koji vrijedi

$$\forall \mathbf{x} \in \Omega, \quad a_0(\mathbf{x}) - \frac{1}{2} \operatorname{div}(\mathbf{b}(\mathbf{x}) + \mathbf{c}(\mathbf{x})) \geq -\eta.$$

C_Ω je kao i ranije Poincaréova konstanta.

Zadatak. Kako se uvjet (2.47) mijenja ako homogeni Dirichletov uvjet zamijenimo Neumannovim uvjetom, a kako ako imamo mješoviti Dirichlet-Neumannov uvjet?

Napomena 2.4 Do sada smo egzistenciju i jedinstvenost slabih rješenja eliptičnih rubnih zadaća dokazivali svođenjem na Lax-Milgramov teorem. Postavlja se pitanje da li su uvjeti tog teorema nužni za egzistenciju i jedinstvenost pripadne varijacijske jednadžbe. Pokazuje se da je odgovor na to pitanje negativan, odnosno da zadaća s bilinearnom formom koja nije koercitivna može imati jedinstveno rješenje. Jedan generalniji teorem egzistencije i jedinstvenosti (korektnosti) rješenja zasniva se na Fredholmovoj alternativi za kompaktne operatore na Banachovim prostorima i može se naći npr. u [3].

2.7 Spektralna zadaća

Spektralne zadaće za parcijalne diferencijalne jednadžbe također mogu biti formulirane varijacijski, posve analogno kao i druge zadaće. Pogledajmo na primjer sljedeći spektralni problem postavljen u ograničenoj domeni $\Omega \subset \mathbb{R}^d$:

$$-\Delta u = \lambda u, \quad u \text{ na } \Omega, \quad (2.48)$$

$$u = 0, \quad \text{na } \partial\Omega. \quad (2.49)$$

Nepoznanice u zadaći (2.48), (2.49) su svojstvena funkcija u i svojstvena vrijednost λ . Standardnim postupkom formiramo varijacijsku zadaću:

$$\begin{cases} \text{Naći } u \in H_0^1(\Omega), u \neq 0, \lambda \in \mathbb{R} \\ \forall v \in H_0^1(\Omega), \int_\Omega \nabla u \cdot \nabla v \, d\mathbf{x} = \lambda \int_\Omega uv \, d\mathbf{x} \end{cases} \quad (2.50)$$

Spektralna zadaća definira svojstvenu funkciju do na skalarni multiplikatora pa je uobičajeno normirati rješenje u , tj. tražiti spektralnu funkciju koja zadovoljava $\|u\|_{H^1(\Omega)} = 1$.

Apstraktnu varijacijsku spektralnu zadaću postavljamo na sljedeći način:

- (A) Zadana su dva realna separabilna Hilbertova prostora V i H takvi da je $V \subset H$ a neprekidnom injekcijom i V je gust podskup od H ;
- (B) Zadana je bilinearna i neprekidna forma $a: V \times V \rightarrow \mathbb{R}$.

Označit ćemo skalarni produkt i normu u prostoru H respektivno s (\cdot, \cdot) i $|\cdot|$, dok će norma prostora V biti označena s $\|\cdot\|$. Spektralna zadaća sada ima sljedeću formu: naći $\lambda \in \mathbb{R}$ i $u \in V$, $u \neq 0$, takve da vrijedi

$$\forall v \in V, \quad a(u, v) = \lambda(u, v). \quad (2.51)$$

Za skalar λ kažemo da je svojstvena vrijednost, a za element $u \in V$ da je svojstveni vektor ili svojstvena funkcija.

U našem prvom primjeru je $H = L^2(\Omega)$, $V = H_0^1(\Omega)$ te

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x}.$$

Lako je provjeriti da su uvjeti A i B ispunjeni.

Napomena 2.5 *Svojstvena zadaća, kako je ovdje postavljena, nije dana u svoj općenitosti budući da su svojstvene vrijednosti i svojstveni vektori općenito kompleksni. Da bismo korektno postavili opću spektralnu zadaću trebali bismo krenuti od kompleksnih Hilbertovih prostora H i V i na $V \times V$ promatrati seskvilinearnu formu $a: V \times V \rightarrow \mathbb{C}$ (linearnu u prvom argumentu i antilinearnu u drugom). Svojstvena zadaća je tada zadana s (2.51) pri čemu tražimo $\lambda \in \mathbb{C}$. Ako je seskvilinearna forma a hermitska, tj. ako vrijedi*

$$\forall u, v \in V, \quad a(u, v) = \overline{a(v, u)},$$

onda se lako pokazuje da je svaka svojstvena vrijednost realna. Na ako je više od toga bilinearna forma realna i simetrična, onda zadaća (2.51) ima i realne svojstvene vektore (u smislu da ako je svojstveni vektor kompleksan, onda su njegov realni i imaginarni dio i sami svojstveni vektori). Stoga tu zadaću možemo promatrati na realnim Hilbertovim prostorima.

Pored uvjeta A i B postavljamo i sljedeće uvjete na bilinearnu formu:

- (C) Bilinearna forma $a: V \times V \rightarrow \mathbb{R}$ je simetrična;
- (D) Bilinearna forma $a: V \times V \rightarrow \mathbb{R}$ je koercitivna na V .

Sada smo u uvjetima Lax-Milgramove leme pa zadaća: naću $u \in V$ takvo da je

$$\forall v \in V, \quad a(u, v) = (f, v), \quad (2.52)$$

ima jedinstveno rješenje za svako $f \in H$. Uočimo da tu koristimo neprekidnost ulaganja prostora V u H zbog koje postoji konstanta c_0 takva da je

$$\forall v \in V, \quad |v| \leq c_0 \|v\|.$$

Uz taj uvjet je $v \mapsto (f, v)$ element prostora V' za svako $f \in H$.

Budući da svaki $f \in H$ određuje jedinstveni element $u \in V$ za koji vrijedi (2.52) dobro je definirani operator $T: H \rightarrow V$ koji djeluje po pravilu $Tf = u$. Operator T je evidentno linearan i neprekidan s ocjenom norme

$$\|Tf\| \leq \frac{c_0}{\alpha} |f|. \quad (2.53)$$

Pomoću operatora T spektralnu zadaću (2.51) možemo zapisati u obliku jednadžbe

$$u = \lambda Tu. \quad (2.54)$$

Rješenje zadaće (2.54) proizlazi iz općih rezultata funkcionalne analize ako je operator T kompaktan. Prisjetimo se pri tome da je neprekidno preslikavanje s jednog metričkog prostora u drugi kompaktan ukoliko ograničene skupove preslikava u relativno kompaktne.

Lema 2.4 Neka su ispunjeni uvjeti A, B i D te neka je kanonsko ulaganje prostora V u prostor H kompaktan. Tada je operator T , promatran kao operator s V u V kompaktan

Dokaz. Neka je $1: V \rightarrow H$ kanonsko ulaganje ($1(v) = v$). Tada je operator T promatran kao operator s V u V ustvari $T \circ 1$. Oba operatora, T i 1 , su neprekidni i prema pretpostavci 1 je kompaktan. Stoga je i kompozicija kompaktan operator. \square

Ponovimo sada neke rezultate funkcionalne analize. Niz vektora $(e_n; n \in \mathbb{N})$ je *topološka baza* prostora X ako se svaki vektor $x \in X$ može na jedinstven način prikazati u obliku

$$x = \sum_{n=1}^{\infty} \alpha_n e_n,$$

gdje su α_i skalari i red konvergira po normi. Lako se pokazuje da je svaki normirani prostor koji ima topološku bazu separabilan.

Ako je prostor X unitaran onda uvodimo pojam ortonormirane baze: Ortonormirani niz vektora $(e_n; n \in \mathbb{N})$ je *ortonormirana baza* ako za svako $x \in X$ vrijedi

$$x = \sum_{n=1}^{\infty} (x, e_n) e_n,$$

pri čemu red konvergira po normi. Korisne su još dvije definicije: Familija vektora $(e_j, j \in \mathcal{J})$ je *fundamentalna* (totalna) u X ako je linearan prostor razapet skupom $\{e_j: j \in \mathcal{J}\}$ gust u X .

Ortonormirana familija vektora $(e_j, j \in \mathcal{J})$ je *maksimalna* u X ako je nulvektor jedini vektor okomit na skup $\{e_j: j \in \mathcal{J}\}$. Za detalje vidi na primjer Kurepa [4]. Nam je potreban sljedeći rezultat:

Teorem 2.9 Neka je V separabilan Hilbertov prostor i (e_n) niz u njemu. Sljedeća su svojstva ekvivalentna.

1. Niz (e_n) je ortonormirana baza u V ;
2. Niz (e_n) je fundamentalan ortonormiran niz u V ;
3. Niz (e_n) je maksimalan ortonormiran niz u V .

Za dokaz vidi Kurepa [4], str. 51.

Teorem 2.10 Neka je V unitaran prostor i $T: V \rightarrow V$ kompaktan simetričan operator beskonačnog ranga. Tada postoji ortonormiran niz $(e_n; n \in \mathbb{N})$ i niz realnih brojeva $(\mu_n; n \in \mathbb{N})$ takvi da je:

$$|\mu_1| \geq |\mu_2| \geq \cdots \geq |\mu_n| \geq \cdots > 0, \quad (2.55)$$

$$Te_n = \mu_n e_n \quad (n \in \mathbb{N}); \quad \lim \mu_n = 0, \quad (2.56)$$

$$u = \sum_{n=1}^{\infty} (u, e_n) e_n \quad (u \in V). \quad (2.57)$$

Niz (μ_n) sadrži sve svojstvene vrijednosti operatora T različite od nule. Svojstveni potprostor svake svojstvene vrijednosti različite od nule je konačnodimenzionalan. Niz (e_n) je maksimalan u V ako i samo ako nula nije svojstvena vrijednost operatora T .

Dokaz se može naći u [4], str 210.

Neka su ispunjene pretpostavke A, B, C i D. Tada bilinearna forma $a(\cdot, \cdot)$ predstavlja jedan skalarni produkt na V koji generira normu ekvivalentnu s normom prostora V . Štoviše, u tom skalarnom produktu je operator T , promatran kao operator s V u V , simetričan i pozitivan. Simetrija:

$$a(Tu, v) = (u, v) = (v, u) = a(Tv, u) = a(u, Tv).$$

Pozitivnost: za $v \neq 0$,

$$a(Tv, v) = |v|^2 > 0.$$

Uz pretpostavku da je kanonsko ulaganje $V \subset H$ kompaktno možemo primijeniti Teorem 2.10.

Teorem 2.11 Neka su ispunjene pretpostavke A, B, C i D te neka je kanonsko ulaganje $V \subset H$ kompaktno. Tada postoji niz brojeva $(\lambda_n; n \in \mathbb{N})$ i niz vektora $(w_n; n \in \mathbb{N})$, $w_n \in V$, koji zadovoljavaju:

$$\forall v \in V, \quad a(w_n, v) = \lambda_n (w_n, v), \quad n \in \mathbb{N}.$$

Nadalje, niz $(\lambda_n; n \in \mathbb{N})$ zadovoljava

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq \cdots; \quad \lim \lambda_n = +\infty, \quad (2.58)$$

niz $(w_n; n \in \mathbb{N})$ je ortonormirana baza u H dok je niz $(\lambda_n^{-1/2} w_n; n \in \mathbb{N})$ ortonormirana baza u prostoru V opskrbljenom skalarnim produktom $a(\cdot, \cdot)$.

Dokaz. Zanimarito ćemo konačnodimenzionalni slučaj buduću da je on jednostavniji i može se tretirati elementarnim sredstvima i primijeniti ćemo Teorem 2.10 na operator T koji generira bilinearna forma. Već smo vidjeli da u prostoru V opskrbljenom skalarnim produktom $a(\cdot, \cdot)$ operator T postaje simetričan, pozitivan i kompaktan. Nadalje, njegov rang ne može biti konačan jer bi tada iz

$$\forall v \in V, \quad a(Tf, v) = (f, v) \quad (2.59)$$

slijedilo da su svi elementi iz V okomiti na $\mathcal{R}(T)$ u novom skalarnom produktu (svi $v \in V$ takvi da je $a(Tf, v) = 0$ za sve $f \in H$) ujedno okomiti na čitav H ($(f, v) = 0$ za sve $f \in H$), no uzimajući $f = v$ dolazimo do $v = 0$, odnosno da nema vektora okomitih na sliku od T . Rang operatora je dakle beskonačan i svi uvjeti Teorema 2.10 su ispunjeni.

Prema Teoremu 2.10 svojstvena zadaća $Te = \mu e$ prebrojivo mnogo rješenja: $(\mu_n), (e_n)$. Zbog pozitivnosti operatora T nula nije svojstvena vrijednost i sve svojstvene vrijednosti su strogo pozitivne, pa stoga prema (2.55), (2.56) one čine niz,

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n \geq \cdots > 0, \quad \lim \mu_n = 0.$$

Zadaća $u = \lambda Tu$, koja predstavlja ekvivalentni zapis zadaće (2.51), ima stoga prebrojivo mnogo rješenja $(\lambda_n), (e_n)$, gdje je $\lambda_n = 1/\mu_n$, pa stoga vrijedi (2.58). Teorem 2.10 nam govori da je niz (e_n) ortonormirana baza u prostoru V opskrbljenom skalarnim produktom induciranim bilinearnom formom. Definirajmo $w_n = \lambda_n^{1/2} e_n$ i pokažimo da niz (w_n) predstavlja ortonormiranu bazu u H . Uočimo prvo da je $w_n = \lambda_n T w_n$, tj. skaliranje ponovo daje rješenje svojstvene zadaće. Odatle, po definiciji operatora T (vidi (2.59)) slijedi

$$(w_n, w_m) = \frac{1}{\lambda_n} a(w_n, w_m) = \sqrt{\frac{\lambda_m}{\lambda_n}} a(e_n, e_m) = \delta_{n,m}.$$

Time smo pokazali da je niz (w_n) ortonormiran u H . Pokažimo sada da je on maksimalan u H , tj. da je samo nulvektor okomit na sve w_n u H . Neka je $f \in H$ takav da je $(f, w_n) = 0$ za sve $n \in \mathbb{N}$. Tada je evidentno $(f, e_n) = 0$ za sve $n \in \mathbb{N}$ a kako niz (e_n) čini ortonormiranu bazu u V to je f okomit na čitav V . Budući da je V gust u H slijedi da je $f = 0$ i maksimalnost je dokazana. Prema Teoremu 2.9 niz je ortonormirana baza u H . \square

Primijenjujući Teorem 2.11 na primjer (2.48), (2.49) možemo naći niz (w_n) u $H_0^1(\Omega)$ koji čini ortogonalnu bazu u $H_0^1(\Omega)$ i ortonormiranu bazu u $L^2(\Omega)$. Promatrajući Neumannovu spektralnu zadaću možemo takav niz naći u $H^1(\Omega)$.

Napomena 2.6 U Teoremu 2.11 možemo oslabiti pretpostavku koercitivnosti. Dovoljno je pretpostaviti da bilinearna forma $a(\cdot, \cdot)$ zadovoljava sljedeći uvjet: postoje brojevi $\alpha > 0$ i $\lambda \in \mathbb{R}$ takvi da je

$$\forall v \in V, \quad a(v, v) + \lambda|v|^2 \geq \alpha\|v\|^2.$$

Tada promatrajući bilinearnu formu $a(u, v) + \lambda(u, v)$ dolazimo u uvjete Teorema 2.11 koji daje niz svojstvenih parova modificirane zadaće. Polazna zadaća ima iste svojstvene funkcije, s istim svojstvima, ali pomaknute svojstvene vrijednosti za λ . Stoga sada svojstvene vrijednosti zadovoljavaju

$$-\lambda < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \leq \dots; \quad \lim \lambda_n = +\infty. \quad \square$$

2.8 Nelinearne zadaće

Teorija egzistencije rješenja nelinearnih diferencijalnih jednadžbi je značajno složenija od linearne teorije i stoga se ovdje u nju nećemo upuštati. Ipak, varijacijska formulacija zadaće se u mnogim slučajevima postavlja na posve isti način kao za linearne zadaće što nam omogućava formiranje varijacijske aproksimacije i konačno upotrebu metode konačnih elemenata. Temeljna je razlika u tome što nas varijacijska aproksimacija sada vodi na nelinearnu jednadžbu koju treba riješiti nekim iterativnim postupkom, na primjer *Newtonovom metodom*.

Ilustrirajmo to jednim promjerom. Promatramo sljedeću rubnu zadaću u ograničenoj domeni $\Omega \subset \mathbb{R}^d$:

$$\operatorname{div}(\alpha(u)\mathbf{A}(\nabla u - \mathbf{b}(u))) = 0, \quad \text{na } \Omega \quad (2.60)$$

$$\alpha(u)\mathbf{A}(\nabla u - \mathbf{b}(u)) \cdot \mathbf{n} = g, \quad \text{na } \Gamma_1 \quad (2.61)$$

$$u = 0, \quad \text{na } \Gamma_2, \quad (2.62)$$

u kojoj je granica domene podijeljena na Dirichletov i Neumannov dio: $\partial\Omega = \Gamma_1 \cup \Gamma_2$, $\Gamma_1 \cap \Gamma_2 = \emptyset$. Matrična funkcija $\mathbf{A} = \mathbf{A}(\mathbf{x})$ simetrična i uniformno pozitivno definitna, s ograničenim koeficijentima (vidi (2.15) i (2.16)). Funkcije $\alpha: \mathbb{R} \rightarrow \mathbb{R}$ i $\mathbf{b}: \mathbb{R} \rightarrow \mathbb{R}^d$ su dovoljno glatke (minimalno neprekidne) i postoji konstanta $\alpha_0 > 0$ takva da je

$$\forall u \in \mathbb{R}, \quad \alpha(u) \geq \alpha_0. \quad (2.63)$$

Ta je pretpostavka nužna kako bi zadaća bila uniformno eliptična. Konačno, o zadanoj funkciji g pretpostavljamo

$$g \in L^2(\Gamma_1). \quad (2.64)$$

Napomena 2.7 Zadaća (2.60)–(2.62) u slučaju $\mathbf{b}(u) = \alpha(u)\mathbf{g}$ opisuje jednofazni stacionarni tok fluida kroz poroznu sredinu. Pri tome rješenje u ima značenje tlaka, $\alpha(u)$ je gustoća fluida, \mathbf{A} je propusnost sredine, a \mathbf{g} je vektor ubrzanja sile teže. Prirodni uvjeti na funkciju $\alpha(u)$ su stroga monotonost (rasuća) i stroga pozitivnost.

Varijacijsku zadaću formiramo kao i do sada, množenjem s dovoljno glatkom test funkcijom v koja se poništava na Γ_2 i parcijalnom integracijom:

$$\int_{\Omega} \alpha(u) \mathbf{A}(\nabla u - \mathbf{b}(u)) \cdot \nabla v \, d\mathbf{x} = \int_{\Gamma_1} gv \, d\sigma.$$

Time smo dobili zadaću naizgled sličnu do sada promatranim zadaćama. Apstraktno zapisana ona prima sljedeći oblik:

$$\begin{cases} \text{Naći } u \in V \\ \forall v \in V, \quad a(u, v) = \langle F, v \rangle, \end{cases} \quad (2.65)$$

gdje je $V = \{\phi \in H^1(\Omega) : \phi = 0 \text{ na } \Gamma_2\}$, te

$$a(u, v) = \int_{\Omega} \alpha(u) \mathbf{A}(\nabla u - \mathbf{b}(u)) \cdot \nabla v \, d\mathbf{x}, \quad \langle F, v \rangle = \int_{\Gamma_1} gv \, d\sigma. \quad (2.66)$$

Osnovna razlika je to što forma $a(\cdot, \cdot)$ više nije bilinearana: ona je linearna u drugom argumentu, ali nelinearna u prvom, i stoga se Lax-Milgramova lema ne može primijeniti.

Istaknimo da kod nelinearnih zadaća uvijek treba pažljivo provjeriti je li forma $a(\cdot, \cdot)$ (i isto tako funkcional desne strane) dobro definirana na odabranom funkcijskom prostoru, odnosno jesu li svi integrali u definiciji varijacijske jednadžbe konačni. U našem slučaju ograničenost nelinearnih funkcija $\alpha(u)$ i $\mathbf{b}(u)$ osigurava da je forma $a(\cdot, \cdot)$ dobro definirana na $H^1(\Omega) \times H^1(\Omega)$.

Varijacijsku aproksimaciju zadaće (2.65) definiramo prirodno, odabirom konačnodimenzionalnog potprostora $V_h \subset V$:

$$\begin{cases} \text{Naći } u_h \in V_h \\ \forall v \in V_h, \quad a(u_h, v) = \langle F, v \rangle \end{cases} \quad (2.67)$$

Odaberimo jednu bazu u prostoru V_h i označimo njene elemente s $\phi_1, \phi_2, \dots, \phi_N$. Funkciju u_h ćemo raspisati po vektorima baze:

$$u_h(\mathbf{x}) = \sum_{j=1}^N U_j \phi_j(\mathbf{x}), \quad \mathbf{U} = (U_j)_{j=1}^N \in \mathbb{R}^N \text{ je vektor koeficijenata.}$$

Time smo dobili sustav nelinearnih algebarskih jednadžbi

$$\mathcal{F}(\mathbf{U}) = 0,$$

gdje je $\mathcal{F}: \mathbb{R}^N \rightarrow \mathbb{R}^N$ funkcija definirana formulama

$$\mathcal{F}_i(\mathbf{U}) = a\left(\sum_{j=1}^N U_j \phi_j, \Phi_i\right) - \langle F, \phi_i \rangle, \quad i = 1, \dots, N.$$

Dobiveni sustav možem rješavati Newtonovom metodom, koja se bazira na linearnoj aproksimaciji. Pretpostavimo da imamo aproksimaciju nultočke \mathbf{U} i da želimo naći korekciju \mathbf{V} takvu da je $\mathcal{F}(\mathbf{U} + \mathbf{V}) = 0$. Tada razvojem u Taylorov red dobivamo

$$0 = \mathcal{F}(\mathbf{U} + \mathbf{V}) = \mathcal{F}(\mathbf{U}) + D\mathcal{F}(\mathbf{U})\mathbf{V} + \dots$$

Zanemarujući više članove u razvoju dolazimo do jednadžbe za \mathbf{V} :

$$0 = \mathcal{F}(\mathbf{U}) + D\mathcal{F}(\mathbf{U})\mathbf{V} \quad (2.68)$$

iz koje slijedi $\mathbf{V} = -D\mathcal{F}(\mathbf{U})^{-1}\mathcal{F}(\mathbf{U})$, a $\mathbf{U} + \mathbf{V}$ je nova aproksimacija. Time dolazimo do sljedećeg iterativnog postupka: za zadanu početnu iteraciju $\mathbf{U}^0 \in \mathbb{R}^N$ i $n = 0, 1, \dots$ imamo

$$\mathbf{U}^{n+1} = \mathbf{U}^n - D\mathcal{F}(\mathbf{U}^n)^{-1}\mathcal{F}(\mathbf{U}^n).$$

To je Newtonov iterativni postupak koji prirodnije zapisujemo u obliku:

$$D\mathcal{F}(\mathbf{U}^n)\mathbf{V} = -\mathcal{F}(\mathbf{U}^n), \quad (2.69)$$

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \mathbf{V}, \quad (2.70)$$

gdje u (2.69) rješavamo linearan sustav s matricom koja je derivacija funkcije \mathcal{F} u prethodnoj iteraciji. Novu iteraciju dobivamo u (2.70) korekcijom prethodne.

Linearni sustav koji treba riješiti u svakom koraku Newtonovog algoritma predstavlja jednu linearnu eliptičku rubnu zadaću. To je najlakše vidjeti ako Newtonovu metodu izvedemo neposredno na varijacijskoj formulaciji nelinearne zadaće, bez uvođenja baze u prostoru konačnih elemenata. S tim ciljem uvedimo funkciju

$$\mathcal{F}_v(u_h) = a(u_h, v) - \langle F, v \rangle = \int_{\Omega} \alpha(u) \mathbf{A}(\nabla u - \mathbf{b}(u)) \cdot \nabla v \, dx - \int_{\Gamma_1} gv \, d\sigma,$$

u kojoj test funkcija v igra ulogu parametra.

Uzmimo da je u_h^n poznata aproksimacija rješenja u_h zadaće (2.67), tj. rješenje jednadžbe $\mathcal{F}_v(u_h) = 0$, za svako $v \in V_h$, i sljedeću aproksimaciju potražimo u obliku $u_h^{n+1} = u_h^n + w_h$. Definiramo realnu funkciju

$$f(t) = \mathcal{F}_v(u_h^n + tw_h) = a(u_h^n + tw_h, v) - \langle F, v \rangle,$$

koju razvijemo Taylorov red,

$$f(1) = f(0) + f'(0) + \frac{1}{2}f''(0) + \dots \quad (2.71)$$

Ako je u_h^{n+1} točno rješenje, onda je $f(1) = 0$. Približnu korekciju w_h stoga računamo, zanemarujući više članove u razvoju (2.71), iz jednadžbe

$$0 = f(0) + f'(0), \quad (2.72)$$

koja je ekvivalentna s (2.68). Raspisivanjem (2.72) dobivamo:

$$0 = a(u_h^n, v) - \langle F, v \rangle + \frac{d}{dt} a(u_h^n + tw_h, v) \Big|_{t=0}$$

i nova iteracija je $u_h^{n+1} = u_h^n + w_h$. Neposrednim računom dobivamo

$$\begin{aligned} \frac{d}{dt} a(u_h^n + tw_h, v) \Big|_{t=0} &= \frac{d}{dt} \int_{\Omega} \alpha(u_h^n + tw_h) \mathbf{A}(\nabla(u_h^n + tw_h) - \mathbf{b}(u_h^n + tw_h)) \cdot \nabla v \, d\mathbf{x} \Big|_{t=0} \\ &= \int_{\Omega} \alpha(u_h^n) \mathbf{A}(\nabla w_h - w_h \mathbf{b}'(u_h^n)) \cdot \nabla v \, d\mathbf{x} \\ &\quad + \int_{\Omega} \alpha'(u_h^n) w_h \mathbf{A}(\nabla u_h^n - \mathbf{b}(u_h^n)) \cdot \nabla v \, d\mathbf{x} \end{aligned}$$

Time smo došli do sljedeće linearne varijacijske zadaće koja definira $w_h \in V_h$:

$$\forall v \in V_h, \quad \mathcal{A}^n(w_h, v) = \langle \mathcal{F}^n, v \rangle \quad (2.73)$$

gdje je

$$\begin{aligned} \mathcal{A}^n(w_h, v) &= \int_{\Omega} \alpha(u_h^n) \mathbf{A}(\nabla w_h - w_h \mathbf{b}'(u_h^n)) \cdot \nabla v \, d\mathbf{x} + \int_{\Omega} \alpha'(u_h^n) w_h \mathbf{A}(\nabla u_h^n - \mathbf{b}(u_h^n)) \cdot \nabla v \, d\mathbf{x} \\ \langle \mathcal{F}^n, v \rangle &= \int_{\Gamma_1} gv \, d\sigma - \int_{\Omega} \alpha(u_h^n) \mathbf{A}(\nabla u_h^n - \mathbf{b}(u_h^n)) \cdot \nabla v \, d\mathbf{x}. \end{aligned}$$

Da bi ta varijacijska jednadžba (2.73) bila dobro definirana moramo postaviti dodatne pretpostavke o nelinearnim koeficijentima i minimalno tražiti da α' u \mathbf{b}' budu ograničene funkcije te da je α' strogo pozitivna.

2.9 Varijacijska aproksimacija

Kada smo jednom rubnu zadaću zapisali u varijacijskoj formi njena se numerička aproksimacija svodi na zamjenu beskonačnodimenzionalnog prostora V , na kome je varijacijska jednadžba postavljena, nekim njegovim konačnodimenzionalnim potprostorom $V_h \subset V$. Taj se postupak naziva varijacijska aproksimacija, a greška aproksimacije zavisi samo o tome koliko dobro V_h aproksimira prostor V , u smislu koji će biti razjašnjen u Teoremu 2.12.

Mogući su različiti načini konstrukcije prostora V_h . Na primjer, ako možemo naći jednu prebrojivu bazu u prostoru V , onda je prirodno za V_h uzeti potprostor razapet s prvih N svojstvenih vektora. Uzimajući sve veći N dobivamo sve bolju aproksimaciju prostora V . Metoda aproksimacije zasnovana na tom izboru naziva se *Galerkinova metoda* i predstavlja moćan postupak za dokazivanje egzistencije rješenja nelinearnih eliptičkih rubnih zadaća. Metoda konačnih elemenata, s druge strane, konstruira V_h kao prostor po dijelovima polinomijalnih funkcija, dovoljno

glatkih kako bi imali $V_h \subset V$. Tu se prvo domena na kojoj je zadaća postavljena aproksimira unijom jednostavnih (najčešće poliedarskih) skupova koji čine *triangulaciju* domene. Funkcije iz V_h se zatim definiraju kao polinomi određenog stupnja na svakom pojedinom elementu iz triangulacije, s određenim uvjetima neprekidnosti na granici susjednih elemenata, koji osiguravaju inkluziju $V_h \subset V$. U toj situaciji parametar h predstavlja dijametar najvećeg elementa triangulacije i pokazuje se da prostor V_h aproksimira V tim bolje što je h manji.

Za analizu varijacijske aproksimacije polazimo od varijacijske jednadžbe koja zadovoljava uvjete Lax-Milgramove leme koje ovdje sumiramo:

Pretpostavke (A):

- V je realan Hilbertov prostor s normom $\|\cdot\|$;
- $a: V \times V \rightarrow \mathbb{R}$ je bilinearna forma;
- $\exists M > 0$ takvo da $\forall u, v \in V \quad |a(u, v)| \leq M\|u\|\|v\|$ (ograničenost);
- $\exists \alpha > 0$ takvo da $\forall v \in V \quad a(v, v) \geq \alpha\|v\|^2$ (koercitivnost);
- $F \in V'$.

Apstraktna varijacijska zadaća ima sljedeći oblik:

Zadaća (P):

$$\begin{cases} \text{Naći } u \in V \\ \forall v \in V, \quad a(u, v) = \langle F, v \rangle \end{cases}$$

Odabirom konačnodimenzionalnog potprostora $V_h \subset V$ dolazimo do *aproksimativne* varijacijske zadaće:

Zadaća (P_h):

$$\begin{cases} \text{Naći } u_h \in V_h \\ \forall v \in V_h, \quad a(u, v) = \langle F, v \rangle \end{cases}$$

Zadaću (P_h) smo promatrali u Lemi 2.1 i tamo smo pokazali da se uvođenjem baze u prostor V_h ona svodi na jedan matrični sustav s pozitivno definitnom matricom (koercitivnost bilinearne forme), koja je simetrična ukoliko je bilinearna forma simetrična. Pozitivna definitnost matrice sustava ima za posljedicu da zadaća (P_h) ima jedinstveno rješenje, no do tog zaključka dolazimo direktno: ako su zadovoljeni svi uvjeti koji osiguravaju da se Lax-Milgramova lema može primijeniti na zadaću (P), onda ti isti uvjeti osiguravaju jedinstvenu rješivost zadaće (P_h), jer je jedina razlika između njih u funkcijskom prostoru. Iz činjenice da je svaki konačnodimenzionalni potprostor Hileberovog prostora i sam Hilbertov (u nasljeđenom skalarnom produktu) vidimo da su sve pretpostavke Lax-Milgramove leme ispunjene i za zadaću (P_h).

Egzistencija i jedinstvenost rješenja diskretne zadaće slijedi dakle direktno iz egzistencije i jedinstvenosti kontinuirane zadaće. Ostaje još pitanje greške aproksimacije, tj. ocjene razlike $\|u - u_h\|$ u različitim normama. To se pitanje, za normu prostora V , rješava sljedećim teoremom koji problem svodi na pitanje *koliko dobro prostor V_h aproksimira prostor V* .

Teorem 2.12 (Céa-ina lema) Neka su ispunjene pretpostavke (A) i neka je $u \in V$ rješenje zadatke (P) te $u_h \in V_h \subset V$ rješenje zadatke $(P)_h$. Tada je

$$\|u - u_h\| \leq \frac{M}{\alpha} \inf_{v \in V_h} \|u - v\|,$$

gdje je $\|\cdot\|$ norma prostora V .

Dokaz. Iz

$$\begin{aligned} a(u, v) &= \langle F, v \rangle, \quad \forall v \in V_h \subset V \\ a(u_h, v) &= \langle F, v \rangle, \quad \forall v \in V_h, \end{aligned}$$

dobivamo *relaciju ortogonalnosti*

$$a(u - u_h, v) = 0, \quad \forall v \in V_h. \quad (2.74)$$

Za proizvoljno $v \in V_h$, koristeći koercitivnost i neprekidnost bilinearne forme te (2.74), dobivamo

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

Time smo dobili ocjenu

$$\forall v \in V_h, \quad \|u - u_h\| \leq \frac{M}{\alpha} \|u - v\|.$$

Uzimanjem infimuma po svim $v \in V_h$ dobivamo tvrdnju. \square

U slučaju simetrične bilinearne forme možemo dobiti bolju konstantu u ocjeni. Uočimo prvo da je u slučaju simetričnosti bilinearne forme $a(u, v)$ jedan dobar skalarni produkt i izraz $\sqrt{a(u, u)}$ predstavlja normu na V , ekvivalentnu s normom $\|\cdot\|$; tu normu nazivamo *energetska norma*.

Zbog ortogonalnosti (2.74) i simetrije imamo za proizvoljno $v \in V_h$

$$\begin{aligned} a(u - v, u - v) &= a(u - u_h + (u_h - v), u - u_h + (u_h - v)) \\ &= a(u - u_h, u - u_h) + 2a(u - u_h, u_h - v) + a(u_h - v, u_h - v) \\ &= a(u - u_h, u - u_h) + a(u_h - v, u_h - v) \end{aligned}$$

Time smo dobili optimalnu ocjenu u energetskej normi

$$a(u - u_h, u - u_h) = \inf_{v \in V_h} a(u - v, u - v). \quad (2.75)$$

Oдавde je

$$\alpha \|u - u_h\|^2 \leq M \|u - v\|^2,$$

izlazi da je za svako $v \in V_h$

$$\|u - u_h\| \leq \sqrt{\frac{M}{\alpha}} \|u - v\|.$$

Time vidimo da Teorem 2.12 u simetričnom slučaju vrijedi s konstantom $C = \sqrt{M/\alpha}$ koja može biti značajno manja od M/α .

Napomena 2.8 *Uzmimo da imamo niz rastućih potprostora $V_{h_i} \subset V$, $i \in \mathbb{N}$; $V_{h_1} \subset V_{h_2} \subset V_{h_3} \subset \dots$. Tada je iz Teorema 2.12 jasno da je minimalna pretpostavka koja osigurava konvergenciju aproksimacijskog procesa sljedeća:*

$$\forall v \in V, \quad \lim_i \left(\inf_{v_h \in V_{h_i}} \|v - v_h\| \right) = 0.$$

Teorem 2.12 želimo primijeniti na varijacijske zadaće koje dobivamo diskretizacijom eliptičkih rubnih zadaća. Pri tome se susrećemo sa sljedećom poteškoćom: Bilinearna je forma u varijacijskoj zadaći dana određenim integralom koji u diskretnoj zadaći općenito moramo aproksimirati nekom integracijskom formulom. Izuzetak je slučaj kada su koeficijenti diferencijalne jednadžbe konstantni, kada najčešće sve integrale možemo izračunati egzaktno.

Ako želimo obuhvatiti situaciju u kojoj integrale u varijacijskoj jednadžbi izračunavamo primjenom neke integracijske formule, onda trebamo diskretnu zadaću $(P)_h$ zamijeniti zadaćom sljedećeg oblika:

Zadaća (P_{hh}) :

$$\begin{cases} \text{Naći } u_h \in V_h \\ \forall v \in V_h, \quad a_h(u_h, v) = F_h(v) \end{cases}$$

gdje je $a_h(\cdot, \cdot): V_h \times V_h \rightarrow \mathbb{R}$ bilinearna forma koja aproksimira formu $a(\cdot, \cdot)$ i $F_h: V_h \rightarrow \mathbb{R}$ linearan funkcional koji aproksimira $F \in V'$. Uočimo dobro da ne pretpostavljamo da je bilinearna forma definirana na čitavom $V \times V$ (i analogno za F_h).

Sljedeći teorem daje apstraktnu ocjenu greške aproksimacije. On pokazuje da se greška sastoji od greške aproksimacije prostora V potprostorom V_h i *greške konzistencije* koja mjeri koliko su dobro aproksimirani bilinearna forma $a(\cdot, \cdot)$ i funkcional $F \in V'$. Budući da koercitivnost diskretne bilinearne forme sada treba posebno dokazivati, ovdje ćemo ju uvrstiti kao pretpostavku.

Teorem 2.13 (Prva Strangova lema) Neka su ispunjene pretpostavke (A) i neka je $u \in V$ rješenje zadaće (P) te $u_h \in V_h \subset V$ rješenje zadaće (P_{hh}) . Pretpostavimo da je diskretna bilinearna forma a_h **uniformno V_h -eliptična** tj. da postoji konstanta $\tilde{\alpha}$, neovisna o potprostoru V_h , takva da je

$$\forall v \in V_h, \quad \tilde{\alpha} \|v\|^2 \leq a_h(v, v). \quad (2.76)$$

Tada vrijedi

$$\|u - u_h\| \leq \inf_{v \in V_h} \left\{ \left(1 + \frac{M}{\tilde{\alpha}}\right) \|u - v\| + \frac{1}{\tilde{\alpha}} \sup_{0 \neq w \in V_h} \frac{|a(v, w) - a_h(v, w)|}{\|w\|} \right\} + \frac{1}{\tilde{\alpha}} \sup_{0 \neq w \in V_h} \frac{|\langle F, w \rangle - F_h(w)|}{\|w\|}. \quad (2.77)$$

Dokaz. Neka je $v \in V_h$ proizvoljan element. Uniformna V_h -eliptičnost daje

$$\begin{aligned}\tilde{\alpha}\|u_h - v\|^2 &\leq a_h(u_h - v, u_h - v) \\ &= a_h(u_h - v, u_h - v) + a(u, u_h - v) - \langle F, u_h - v \rangle \\ &= a(u - v, u_h - v) + \{a(v, u_h - v) - a_h(v, u_h - v)\} \\ &\quad + F_h(u_h - v) - \langle F, u_h - v \rangle.\end{aligned}$$

Koristeći neprekidnost bilinearne forme $a(\cdot, \cdot)$, dobivamo

$$\begin{aligned}\tilde{\alpha}\|u_h - v\| &\leq M\|u - v\| + \frac{|a(v, u_h - v) - a_h(v, u_h - v)|}{\|u_h - v\|} + \frac{|F_h(u_h - v) - \langle F, u_h - v \rangle|}{\|u_h - v\|} \\ &\leq M\|u - v\| + \sup_{w \in V_h} \frac{|a(v, w) - a_h(v, w)|}{\|w\|} + \sup_{w \in V_h} \frac{|F_h(w) - \langle F, w \rangle|}{\|w\|}.\end{aligned}$$

Koristeći nejednakost trokuta

$$\|u - u_h\| \leq \|u - v\| + \|u_h - v\|,$$

te zbog proizvoljnosti $v \in V_h$, izlazi tvrdnja. \square

Napomena 2.9 Uočimo da se u slučaju $a(\cdot, \cdot) = a_h(\cdot, \cdot)$, i $F_h(\cdot) = \langle F, \cdot \rangle$, Prva Strangova lema svodi na Céa-inu lemu, no konstanta više nije optimalna. To je posljedica toga što u dokazu nismo pretpostavljali da je bilinearne forma a_h dobro definirana na čitavom V .

U dosadašnjim razmatranjima pretpostavljali smo da je aproksimativna metoda **konformna**, što znači da je $V_h \subset V$, odnosno da aproksimativno rješenje tražimo u prostoru u kojem se nalazi i egzaktno rješenje. Za metodu kažemo na je **nekonformna** ako uvjet $V_h \subset V$ nije zadovoljen. Nekonformnu aproksimaciju koristimo tipično onda kada funkcije iz prostora aproksimativnih rješenja nemaju glatkoću koja se zahtjeva od funkcija iz V .

Diskretna zadaća u nekonformnoj aproksimaciji općenito će imati oblik (P_{hh}) . Sljedeće pretpostavke će osigurati da je zadaća (P_{hh}) dobro postavljena:

Pretpostavke (A_h) :

- V_h je konačnodimenzionalni prostor s normom $\|\cdot\|_h$;
- $a_h: (V + V_h) \times (V + V_h) \rightarrow \mathbb{R}$ je bilinearne forma;
- $\exists M^* > 0$, neovisno o h , takvo da $\forall u, v \in V + V_h \quad |a_h(u, v)| \leq M\|u\|_h\|v\|_h$ (ograničenost);
- $\exists \alpha^* > 0$, neovisno o h , takvo da $\forall v \in V_h \quad a_h(v, v) \geq \alpha\|v\|_h^2$ (koercitivnost);
- F_h je linearan funkcional na V_h .

Prostor $V + V_h$ definira se na uobičajen način:

$$V + V_h = \{w = v_1 + v_2 : v_1 \in V, v_2 \in V_h\}.$$

Pretpostavke (A_h) evidentno osiguravaju jedinstvenu rješivost diskretne zadaće (P_{hh}) , čak i kad prostor $V + V_h$ zamijenimo s V_h . Pretpostavka da je bilinearna forma definirana i ograničena na širem prostoru $V + V_h$ bit će korištena u sljedećoj ocjeni greške aproksimacije.

Teorem 2.14 (Druga Strangova lema) Neka su ispunjene pretpostavke (A) i (A_h) te neka je $u \in V$ rješenje zadaće (P) i $u_h \in V_h$ rješenje zadaće (P_{hh}) . Tada vrijedi

$$\|u - u_h\|_h \leq \left(1 + \frac{M^*}{\alpha^*}\right) \inf_{v \in V_h} \|u - v\|_h + \frac{1}{\alpha^*} \sup_{0 \neq w \in V_h} \frac{|a_h(u, w) - F_h(w)|}{\|w\|_h}. \quad (2.78)$$

Dokaz. Neka je $v \in V_h$ proizvoljan element. Prema pretpostavkama (A_h) imamo

$$\begin{aligned} \alpha^* \|u_h - v\|_h^2 &\leq a_h(u_h - v, u_h - v) \\ &= a_h(u - v, u_h - v) + (F_h(u_h - v) - a_h(u, u_h - v)) \\ &\leq M^* \|u - v\|_h \|u_h - v\|_h + |a_h(u, u_h - v) - F_h(u_h - v)|, \end{aligned}$$

odnosno,

$$\begin{aligned} \alpha^* \|u_h - v\|_h &\leq M^* \|u - v\|_h + \frac{|a_h(u, u_h - v) - F_h(u_h - v)|}{\|u_h - v\|_h} \\ &\leq M^* \|u - v\|_h + \sup_{0 \neq w \in V_h} \frac{|a_h(u, w) - F_h(w)|}{\|w\|_h}. \end{aligned}$$

Rezultat sada slijedi iz nejednakosti trokuta

$$\|u - u_h\|_h \leq \|u - v\|_h + \|u_h - v\|_h. \quad \square$$

Napomena 2.10 U Drugoj Strangovoj lemi kao greška konzistencije pojavljuje se residual – greška koju dobivamo kada u aproksimativnu varijacijsku formulaciju uvrstimo točno rješenje u . Analognu ocjenu greške aproksimacije možemo izvesti, na posve isti način, i u konformnom slučaju, ako pretpostavimo da je aproksimativna bilinearna forma definirana i ograničena na širem prostoru $V + V_h$.

2.10 Dodatak. Lax-Milgramova lema

Teorem poznat pod nazivom *Lax-Milgramova lema* predstavlja osnovni egzistencijalni rezultat na kojem se bazira teorija slabih rješenja eliptičkih parcijalnih diferencijalnih jednadžbi. U osnovi se radi o generalizaciji Rieszovog teorema o reprezentaciji linearnog funkcionala, no zbog važnosti

rezultata donosimo nekoliko njegovih varijanti. Premda se posve analogna teorija može zasnovati na kompleksnim prostorima mi ćemo, radi jednostavnosti, promatrati isključivo realne prostore funkcija. Čitatelj zainteresiran za *kompleksnu* teoriju može pogledati npr. [6].

U ovoj sekciji V označava realan Hilbertov prostor (vidi npr. Kurepa [4]). V je dakle linearan (vektorski) prostor u kome postoji skalarni produkt koji ćemo jednostavno označavati s (\cdot, \cdot) . On generira normu po formuli

$$\|u\| = \sqrt{(u, u)},$$

i u toj normi je V potpun prostor (u njemu je svaki Cauchyjev niz konvergentan, [4]). Uz svaki normirani prostor V redovito promatramo i linearan prostor V' koji se sastoji od svih linearnih i neprekidnih funkcionala nad V . Budući da je neprekidnost linearnog funkcionala ekvivalentna njegovoj ograničenosti ([4]), prostor V' je normiran prostor s normom

$$\|F\| = \sup_{u \neq 0} \frac{F(u)}{\|u\|}.$$

Pri tome ćemo djelovanje funkcionala $F \in V'$ na element $u \in V$ označavati funkcijski s $F(u)$ ili kao $\langle F, u \rangle$, pri čemu je druga oznaka uobičajenija u funkcionalnoj analizi. Prostor V' je potpun (Banachov) čak i kada normirani prostor V to nije (vidi [4], Teorem 2, str. 60).

Definicija 2.4 Preslikavanje $a: V \times V \rightarrow \mathbb{R}$ je bilinearna forma (ili bilinearni funkcional) ako je linearan funkcional u svakoj varijabli zasebno. Preciznije,

$$\begin{aligned} \forall \alpha, \beta \in \mathbb{R}, \forall u, v, w \in V, \quad a(\alpha u + \beta v, w) &= \alpha a(u, w) + \beta a(v, w), \\ \forall \alpha, \beta \in \mathbb{R}, \forall u, v, w \in V, \quad a(u, \alpha v + \beta w) &= \alpha a(u, v) + \beta a(u, w). \end{aligned}$$

Definicija 2.5 Bilinearna forma $a: V \times V \rightarrow \mathbb{R}$ je ograničena ako postoji konstanta M , takva da vrijedi

$$\forall u, v \in V, \quad |a(u, v)| \leq M \|u\|_V \|v\|_V. \quad (2.79)$$

Neprekidnost bilinearne forme je ekvivalentna s ograničenosti ([4], Propozicija 8, str 73). Sljedeće svojstvo je fundamentalno za Lax-Milgramovu lemu.

Definicija 2.6 Bilinearna forma $a: V \times V \rightarrow \mathbb{R}$ je koercitivna ako postoji konstanta $\alpha > 0$ takva da je

$$\forall v \in V, \quad \alpha \|v\|_V^2 \leq a(v, v). \quad (2.80)$$

Definicija 2.7 Bilinearna forma $a: V \times V \rightarrow \mathbb{R}$ je simetrična ako vrijedi

$$\forall u, v \in V, \quad a(u, v) = a(v, u). \quad (2.81)$$

Napomena 2.11 Neka je V Hilbertov prostor s normom $\|\cdot\|$. Tada svaka simetrična bilinearna, ograničena i koercitivna forma $a: V \times V \rightarrow \mathbb{R}$ predstavlja jedan skalarni produkt na V i formulom

$$\|v\|_a = \sqrt{a(v, v)}$$

je dobro definirana norma ekvivalentna s $\|\cdot\|$. Tu normu nazivamo **energetska norma**.

Vidjeli smo da se eliptička rubna zadaća može preformulirati u obliku varijacijska zadaće sljedećeg oblika:

$$\begin{cases} \text{naći } u \in V \\ a(u, v) = F(v), \quad \forall v \in V. \end{cases} \quad (2.82)$$

Funkciju v iz varijacijske jednadžbe nazivamo **test funkcija**.

Sljedeća lema pokazuje da u slučaju simetrične bilinearne forme varijacijskoj jednadžbi (2.82) možemo pridružiti jedan problem minimizacije kome je ona ekvivalentna. Pokazuje se da je (2.82) Eulerova jednadžba za pripadni problem minimizacije. Takva karakterizacija je vrlo korisna jer je jednostavno pokazati da problem minimizacije ima jedinstveno rješenje.

Eliptičke parcijalne diferencijalne jednadžbe, pa onda i pripadne varijacijske zadaće, modeliraju ravnotežne fizikalne procese. Ravnotežna konfiguracija je u principu ona u kojoj je pripadna energija minimalna, tako da pridruženi funkcional ima redovito interpretaciju *energije sustava*.

Lema 2.5 Neka je V linearan prostor i $a: V \times V \rightarrow \mathbb{R}$ simetrična, pozitivna bilinearna forma, tj. $a(v, v) > 0$ za svako $0 \neq v \in V$. Zadan je još linearan funkcional $F: V \rightarrow \mathbb{R}$. Tada funkcional

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

dostiže svoj minimum na V u točki $u \in V$ ako i samo ako vrijedi

$$\forall v \in V, \quad a(u, v) = \langle F, v \rangle. \quad (2.83)$$

Štoviše, (2.83) ima najviše jedno rješenje.

Dokaz. Za $u, v \in V$ i $t \in \mathbb{R}$ imamo

$$\begin{aligned} J(u + tv) &= \frac{1}{2}a(u + tv, u + tv) - \langle F, u + tv \rangle \\ &= J(u) + t[a(u, v) - \langle F, v \rangle] + \frac{1}{2}t^2 a(v, v). \end{aligned}$$

Ako $u \in V$ zadovoljava (2.83), onda za $t = 1$ imamo

$$J(u + v) = J(u) + \frac{1}{2}a(v, v) > J(u), \quad \forall v \in V, v \neq 0.$$

Prema tome, $u \in V$ je jedinstvena točka minimuma. Obratno, neka J ima minimum u točki $u \in V$. Tada derivacija funkcije $t \mapsto J(u + tv)$, za svako $v \in V$, mora biti nula u $t = 0$ (radi se o kvadratnoj funkciji). Iz

$$\left. \frac{d}{dt} J(u + tv) \right|_{t=0} = a(u, v) - \langle F, v \rangle$$

izlazi tvrdnja. \square

Sada dokazujemo verziju Lax-Milgramove leme koja se bazira na simetriji bilinearne forme. Štoviše, možemo dokazati jednu generalizaciju koja se primjenjuje na *varijacijske nejednakosti*.

Teorem 2.15 (Lax-Milgramova lema za konveksan skup). Neka je $K \subset V$ neprazan zatvoren i konveksan podskup Hilbertovog prostora V te neka je $a: V \times V \rightarrow \mathbb{R}$ simetrična bilinearna, ograničena i koercitivna forma. Tada za svako $F \in V'$ minimizacijska zadaća

$$J(u) = \min_{v \in K} J(v), \quad J(v) = \frac{1}{2}a(v, v) - \langle F, v \rangle, \quad (2.84)$$

ima jedinstveno rješenje $u \in K$,

Dokaz. 1) Funkcional J je ograničen odozdo:

$$\begin{aligned} J(v) &\geq \frac{1}{2}\alpha\|v\|^2 - \|F\|\|v\| \\ &= \frac{1}{2\alpha}(\alpha\|v\| - \|F\|)^2 - \frac{1}{2\alpha}\|F\|^2 \geq -\frac{1}{2\alpha}\|F\|^2. \end{aligned}$$

2) Neka je $c_1 = \inf\{J(v) : v \in K\}$. Po definiciji infimuma možemo naći niz (v_n) (tzv. minimizirajući niz) takav $J(v_n) \rightarrow c_1$ kada $n \rightarrow \infty$. Koristeći bilinearne forme dobivamo

$$\begin{aligned} \alpha\|v_n - v_m\|^2 &\leq a(v_n - v_m, v_n - v_m) \\ &= 2a(v_n, v_n) + 2a(v_m, v_m) - a(v_n + v_m, v_n + v_m) \\ &= 4J(v_n) + 4J(v_m) - 8J\left(\frac{v_n + v_m}{2}\right) \\ &\leq 4J(v_n) + 4J(v_m) - 8c_1. \end{aligned}$$

U zadnjoj ocjeni smo koristili konveksnost skupa K koja daje $(v_n + v_m)/2 \in V$. Sada evidentno desna strana teži u nulu kada $n, m \rightarrow \infty$ pa izlazi da je (v_n) Cauchyjev niz. Zbog potpunosti prostora, postoji $v \in V$ takav da je $u = \lim v_n$. Zatvorenost skupa K povlači da je $u \in K$, a neprekidnost funkcionala J daje

$$J(u) = \lim J(v_n) = \inf_{v \in K} J(v).$$

3) Jedinstvenost. Pretpostavimo da su u_1 i u_2 rješenja problema minimizacije. Niz $u_1, u_2, u_1, u_2, \dots$ je evidentno minimizirajući niz. Prema dokazanom, on je Cauchyjev, što je moguće jedino ako je $u_1 = u_2$. \square

Napomena 2.12 U dokazu smo koristili jednakost paralelograma

$$2a(v_n, v_n) + 2a(v_m, v_m) = a(v_n - v_m, v_n - v_m) + a(v_n + v_m, v_n + v_m)$$

koja slijedi iz bilinearosti forme.

Napomena 2.13 U konačnodimenzionalnom prostoru, umjesto koercitivnosti u Teoremu 2.15 je dovoljno zahtijevati da forma zadovoljava strogu pozitivnost

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

Koercitivnost slijedi iz kompaktnosti jedinične kugle u konačno dimenzionalnom prostoru.

U specijalnom slučaju $K = V$ problem minimizacije (2.84) je prema Lemi 2.5 ekvivalentan varijacijskoj zadaći (2.83) i stoga imamo ovaj rezultat:

Teorem 2.16 (Lax-Milgramova lema. Simetričan slučaj) Neka je V Hilbertov prostor i $a: V \times V \rightarrow \mathbb{R}$ bilinearna, ograničena, simetrična i koercitivna forma. Tada za svako $F \in V'$ problem

$$\begin{cases} \text{Naći } u \in V \\ \forall v \in V, \quad a(u, v) = \langle F, v \rangle \end{cases} \quad (2.85)$$

ima jedinstveno rješenje $u \in V$ koje je ujedno rješenje minimizacijske zadaće (2.84) uz $K = V$.

Teorem 2.16 možemo primijeniti na skalarni produkt u V i time dobivamo:

Teorem 2.17 (Rieszov teorem o reprezentaciji linearnog funkcionala) Za svako $F \in V'$ postoji jedinstveni element $u_F \in V$ sa svojstvom

$$\forall v \in V, \quad (u_F, v) = \langle F, v \rangle.$$

Pri tome je $\|F\|_{V'} = \|u_F\|_V$.

Za kompleksan Hilbertov prostor vidi Teorem 4, str. 112 u [4].

Teorem 2.18 (Projekcija na zatvoren konveksan skup) Neka je $K \subset V$ neprazan zatvoren i konveksan podskup Hilbertovog prostora V . Za svako $u \in V$ postoji jedinstven element $u_K \in K$ takav da je

$$\|u - u_K\| = \min_{v \in K} \|u - v\|. \quad (2.86)$$

Dokaz. Primijenimo Teorem 2.15 na

$$J(v) = \frac{1}{2}(v, v) - (u, v)$$

i definirajmo $u_K \in K$ kao jedinstveno rješenje problema minimizacije

$$J(u_K) = \min_{v \in K} J(v).$$

Budući da je

$$\frac{1}{2}\|u - v\|^2 = J(v) + \frac{1}{2}\|u\|^2$$

imamo

$$\frac{1}{2}(\min_{v \in K} \|u - v\|)^2 = \min_{v \in K} J(v) + \frac{1}{2}\|u\|^2 = J(u_K) + \frac{1}{2}\|u\|^2 = \frac{1}{2}\|u - u_K\|^2,$$

i time dobivamo da je u_K rješenje zadaće (2.86). Jedinственost slijedi iz jedinstvenosti minimuma funkcionala J . \square

Teorem 2.18 pokazuje da je dobro definirano preslikavanje $P_K: V \rightarrow K$ po formuli $P_K(u) = u_K$. Uočimo da općenito P_K nije linearan operator.

Lema 2.5 daje Eulerovu jednadžbu za problem minimizacije na čitavom prostoru V . Analogan rezultat možemo dokazati i za problem minimizacije na zatvorenom konveksnom skupu $K \subset V$. U tom je slučaju minimum karakteriziran pomoću varijacijske nejednakosti.

Lema 2.6 Neka je V linearan prostor i $a: V \times V \rightarrow \mathbb{R}$ simetrična, pozitivna bilinearna forma, tj. $a(v, v) > 0$ za svako $0 \neq v \in V$. Zadan je još linearan funkcional $F: V \rightarrow \mathbb{R}$ i neprazan zatvoren i konveksan podskup $K \subset V$. Tada funkcional

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

dostiže svoj minimum na K u točki $u \in K$ ako i samo ako vrijedi

$$\forall v \in K, \quad a(u, v - u) \geq \langle F, v - u \rangle. \quad (2.87)$$

Štoviše, (2.87) ima najviše jedno rješenje.

Dokaz. Za $u, v \in K$ i $t \in [0, 1]$ znamo da je $tv + (1 - t)u = u + t(v - u) \in K$ te vrijedi

$$J(tv + (1 - t)u) = J(u) + t[a(u, v - u) - \langle F, v - u \rangle] + \frac{1}{2}t^2a(v - u, v - u). \quad (2.88)$$

1. Ako $u \in K$ zadovoljava (2.87), onda se iz (2.88) vidi da je

$$J(tv + (1 - t)u) > J(u), \quad \forall v \in K, v \neq u, \quad \forall t \in (0, 1],$$

što pokazuje da je u točka strogog minimuma funkcionala. Budući da je minimum strogi, on je i jedinstven, a onda i zadaća (2.87) može imati najviše jedno rješenje.

2. Neka J ima minimum na K u točki $u \in K$. Tada je za proizvoljno $v \in K$ i $t \in [0, 1]$, $J(tv + (1-t)u) \geq J(u)$ i prema (2.88) za svako $t \in (0, 1]$ vrijedi

$$[a(u, v - u) - \langle F, v - u \rangle] + \frac{t}{2}a(v - u, v - u) \geq 0.$$

Prijelazom na limes kada $t \rightarrow 0+$ dobivamo (2.87). \square

Napomena 2.14 (*Stampacchia*) Kombinacijom Teorema 2.15 i Leme 2.6 dolazimo do zaključka da u uvjetima Teorema 2.15 varijacijska nejednakost (2.87) ima jedinstveno rješenje i da je to rješenje minimum funkcionala J na skupu K .

Posljedica 2.1 Neka je K neprazan zatvoren i konveksan podskup Hilbertovog prostora V i $P_K: V \rightarrow K$ pripadni projektor, definiran prema Teoremu 2.18. Tada je $u_K = P_K(u)$ ako i samo ako vrijedi

$$\begin{cases} u_K \in K \\ \forall v \in K, \quad (u - u_K, v - u_K) \leq 0. \end{cases} \quad (2.89)$$

Štoviše, za svako $u, v \in V$ vrijedi

$$\|P_K(u) - P_K(v)\| \leq \|u - v\|. \quad (2.90)$$

Dokaz. U Teoremu 2.18 je dokazano da je $u_K = P_K(u)$ ako i samo ako je u_K minimum funkcionala $J(v) = a(v, v)/2 - \langle F, v \rangle$ na skupu K , gdje je $a(w, v) = (w, v)$ i $\langle F, v \rangle = (u, v)$. Prema Lemi 2.6 u_K je traženi minimum ako i samo ako zadovoljava

$$\forall v \in K, \quad (u_K, v - u_K) \geq (u, v - u_K),$$

odakle slijedi (2.89).

Da bismo dokazali (2.90) uzmimo $u, v \in V$ i napišimo karakterizacije projektora: za sve $w \in K$

$$(u - P_K(u), w - P_K(u)) \leq 0, \quad (v - P_K(v), w - P_K(v)) \leq 0.$$

Pogodnim izborom funkcije $w \in K$ u obje jednadžbe dobivamo:

$$(u - P_K(u), P_K(v) - P_K(u)) \leq 0, \quad (v - P_K(v), P_K(u) - P_K(v)) \leq 0.$$

Zbrajanjem i reorganizacijom izraza dobivamo

$$(P_K(u) - P_K(v), P_K(u) - P_K(v)) \leq (P_K(u) - P_K(v), u - v)$$

odakle slijedi tvrdnja. \square

Posljedica 2.2 Neka je K zatvoren potprostor Hilbertovog prostora V i $P_K: V \rightarrow K$ pripadni projektor, definiran prema Teoremu 2.18. Tada je $u_K = P_K(u)$ ako i samo ako vrijedi

$$\begin{cases} u_K \in K \\ \forall v \in K, \quad (u - u_K, v) = 0. \end{cases} \quad (2.91)$$

Štoviše, operator P_K je linearan.

Dokaz je evidentan i izlazi iz činjenice da test funkciju u (2.89) možemo uzeti u obliku $v = u_K \pm w$, gdje je $w \in K$ proizvoljan element.

Da bismo dokazali Lax-Milgramovu lemu za nesimetrične ograničene i koercitivne bilinearne forme moramo se poslužiti Rieszovim teoremom o reprezentaciji funkcionala na Hilbertovom prostoru i njegovom generalizacijom na bilinearne forme.

Prvo se prisjetimo da svakom bilinearnoj i ograničenoj formi možemo pridružiti jedinstven neprekidan linearan operator.

Lema 2.7 Neka je V Hilbertov prostor sa skalarnim produktom (\cdot, \cdot) i $a: V \times V \rightarrow \mathbb{R}$ bilinearna ograničena forma. Tada postoji jedinstveni ograničeni linearni operator $A: V \rightarrow V$ takav da je

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

Dokaz. Primjenom Rieszovog teorema o reprezentaciji funkcionala (Teorem 2.17) za svako $u \in V$ možemo naći jedinstveni element $A(u) \in V$ takav da je $a(u, v) = (A(u), v)$ za svako $v \in V$. Time je definiran operator $u \mapsto A(u)$ koji je linearan zbog bilinearnosti forme. Iz ograničenosti bilinearne forme slijedi

$$|(Au, v)| = |a(u, v)| \leq M\|u\|\|v\|,$$

i time $\|A\| \leq M$. \square

Nadalje, iz svojstva koercitivnosti bilinearne forme dobivamo

$$\forall u \in V, \quad (Au, u) \geq \alpha\|u\|^2, \quad (2.92)$$

pa primjenom Schwarzove nejednakosti slijedi

$$\forall u \in V, \quad \|Au\| \geq \alpha\|u\|. \quad (2.93)$$

Odavde je jasno da je operator A injektivan i ako pokažemo da je *surjektivan* iz (2.93) će slijediti da je ograničen te $\|A^{-1}\| \leq 1/\alpha$.

Iz (2.93) slijedi da je $\mathcal{R}(A)$ zatvoren prostor. Zaista, neka je $f \in V$ takav da postoji niz $v_n \in V$ takav da $Av_n \rightarrow f$ kada $n \rightarrow \infty$. Niz (Av_n) je konvergentan pa je i Cauchyjev, a zbog (2.93) je i niz (v_n) Cauchyjev. Potpunost prostora V osigurava da niz (v_n) konvergira prema nekom $v \in V$, a neprekidnost (ograničenost) operatora A daje da niz (Av_n) konvergira prema Av . Zbog jedinstvenosti limesa je $f = Av$ i time je zatvorenost slike $\mathcal{R}(A)$ dokazana.

Pokažimo da je $\mathcal{R}(A) = V$. U suprotnom bismo prema Posljedici 2.2 mogli naći netrivialan vektor $v \in V$ okomit na $\mathcal{R}(A)$, tj.

$$(Au, v) = 0 \quad \forall u \in V.$$

No zbog (2.92) izlazi $v = 0$, što je kontradikcija te zaključujemo da je operator A surjektivan. Sada se lako dokazuje

Teorem 2.19 (Lax-Milgramova lema) Neka je V Hilbertov prostor i $a: V \times V \rightarrow \mathbb{R}$ bilinearna, ograničena i koercitivna forma. Tada za svako $F \in V'$ problem

$$\begin{cases} \text{naći } u \in V \\ \forall v \in V, \quad a(u, v) = \langle F, v \rangle \end{cases} \quad (2.94)$$

ima jedinstveno rješenje $u \in V$ i preslikavanje $F \mapsto u$ je neprekidno sa V' u V .

Dokaz. Neka je $A \in \mathcal{L}(V)$ linearan operator pridružen bilinearnoj formi a . Zadaća (2.94) se svodi na jednadžbu

$$Au = f,$$

gdje je $f \in V$ element definiran po Rieszovom teoremu s $\langle F, v \rangle = (f, v)$. Prema dokazanom operator A je neprekidna bijekcija, pa stoga zadaća ima jedinstveno rješenje. Nadalje, iz (2.93) slijedi $\|u\| \leq \|f\|/\alpha$. \square

Ovaj se dokaz lako generalizira na varijacijske zadaće u kojima se prostor u kome se traži rješenje i prostor test funkcija razlikuju. Neka su V i W dva Hilbertova prostora i $a: W \times V \rightarrow \mathbb{R}$ bilinearna forma sa sljedećim svojstvima:

$$\exists M > 0, \forall w \in W, \forall v \in V, \quad |a(w, v)| \leq M \|w\|_W \|v\|_V \quad (2.95)$$

$$\exists \alpha > 0, \forall w \in W, \quad \sup_{v \in V, v \neq 0} \frac{a(w, v)}{\|v\|_V} \geq \alpha \|w\|_W, \quad (2.96)$$

$$\forall v \in V, v \neq 0, \quad \sup_{w \in W} a(w, v) > 0. \quad (2.97)$$

Za proizvoljno $F \in V'$ imamo varijacijsku zadaću:

$$\begin{cases} \text{naći } u \in W \\ \forall v \in V, \quad a(u, v) = \langle F, v \rangle \end{cases} \quad (2.98)$$

Teorem 2.20 (Nečas) Neka su V i W dva Hilbertova prostora i $a: W \times V \rightarrow \mathbb{R}$ bilinearna forma koja zadovoljava (2.95), (2.96) i (2.97). Tada za svako $F \in V'$ problem (2.98) ima jedinstveno rješenje $u \in W$ za koje vrijedi

$$\|u\|_W \leq \frac{1}{\alpha} \|F\|_{V'}. \quad (2.99)$$

Dokaz. Postoji linearan i neprekidan operator $A: W \rightarrow V$ takav da je $a(w, v) = (Aw, v)$ za sve $w \in W$ i $v \in V$, te zbog (2.95) vrijedi $\|Aw\|_V \leq M\|w\|_W$. Analogno, postoji jedinstveni element $f \in V$ takav da je $\langle F, v \rangle = (f, v)$. Time zadaća (2.98) postaje ekvivalentna jednadžbi $Aw = f$ u prostoru V .

1. Operator A je injektivan jer iz (2.96) slijedi $\|Aw\| \geq \alpha\|w\|$ za svako $w \in W$.
 2. Iz te relacije izlazi, kao ranije, da je $\mathcal{R}(A)$ zatvoren prostor.
 3. Iz (2.97) slijedi da je jedini element okomit na $\mathcal{R}(A)$ nulvektor, pa je stoga $\mathcal{R}(A) = V$.
- Time su egzistencija i jedinstvenost dokazani, a ocjena (2.99) slijedi lako iz točke 1. \square

Da bismo dokazali nesimetričnu verziju Lax-Milgramove leme za konveksan skup preba nam Banachov teorem o fiksnoj točki:

Teorem 2.21 (Banach) Neka je (X, d) potpun metrički prostor i $S: X \rightarrow X$ preslikavanje sa svojom da postoji konstanta $L < 1$ takva da je

$$\forall u, v \in X, \quad d(S(u), S(v)) \leq Ld(u, v).$$

Tada S ima jedinstvenu fiksnu točku, $u = S(u)$.

Za dokaz vidi [5], Teorem 8, str 131.

Teorem 2.22 (Lax-Milgramova lema za konveksan skup - nesimetrični slučaj) Neka je V Hilbertov prostor i $a: V \times V \rightarrow \mathbb{R}$ bilinearna, ograničena i koercitivna forma, te neka je $K \subset V$ neprazan, zatvoren i konveksan skup. Tada za svako $F \in V'$ problem

$$\begin{cases} \text{naći } u \in K \\ \forall v \in K, \quad a(u, v - u) \geq \langle F, v - u \rangle \end{cases} \quad (2.100)$$

ima jedinstveno rješenje $u \in K$.

Dokaz.

1. *Jedinstvenost.* Neka su $u_1, u_2 \in K$ dva rješenja varijacijske nejednakosti (2.100). Imamo:

$$a(u_1, v - u_1) \geq \langle F, v - u_1 \rangle, \quad a(u_2, v - u_2) \geq \langle F, v - u_2 \rangle.$$

Odabirom test funkcije slijedi

$$a(u_1, u_2 - u_1) \geq \langle F, u_2 - u_1 \rangle, \quad a(u_2, u_1 - u_2) \geq \langle F, u_1 - u_2 \rangle.$$

Zbrajanjem se desne strane krata pa imamo

$$0 \leq a(u_1, u_2 - u_1) + a(u_2, u_1 - u_2) = -a(u_1 - u_2, u_1 - u_2) \leq -\alpha\|u_1 - u_2\|^2$$

gdje smo iskoristili koercitivnost. Odavdje slijedi $u_1 = u_2$.

2. *Egzistencija.* Prijeđimo na operatorsku formulaciju zadaće. Neka je $A \in \mathcal{L}(V)$ operator pridružen bilinearnoj formi a : $a(u, v) = (Au, v)$ za sve $u, v \in V$; neka je $f \in F$ reprezentacija funkcionala $F \in V'$: $\langle F, v \rangle = (f, v)$. Sada se zadaća (2.100) svodi na

$$\forall v \in K, \quad (Au - f, v - u) \geq 0. \quad (2.101)$$

Odaberimo proizvoljan $\rho > 0$ i zapišimo (2.101) u obliku

$$\forall v \in K, \quad (u - \rho(Au - f) - u, v - u) \leq 0. \quad (2.102)$$

Usporedimo li (2.102) s karakterizacijom (2.89) projekcije na konveksan skup K dobivamo da se (2.102) ekvivalentno može zapisati u obliku:

$$u = P_K(u - \rho(Au - f)). \quad (2.103)$$

Definirajmo (nelinearni) operator $W_\rho: V \rightarrow V$ formulom $W_\rho(u) = P_K(u - \rho(Au - f))$ i pokažimo da je za dovoljno malo ρ on kontrakcija. Zaista, za $u_1, u_2 \in V$, koristeći (2.90) (P_K je kontrakcija) imamo

$$\begin{aligned} \|W_\rho(u_1) - W_\rho(u_2)\|^2 &= \|P_K(u_1 - \rho(Au_1 - f)) - P_K(u_2 - \rho(Au_2 - f))\|^2 \\ &\leq \|(u_1 - \rho(Au_1 - f)) - (u_2 - \rho(Au_2 - f))\|^2 \\ &= \|(u_1 - u_2) - \rho A(u_1 - u_2)\|^2 \\ &= \|u_1 - u_2\|^2 + \rho^2 \|A(u_1 - u_2)\|^2 - 2\rho(u_1 - u_2, A(u_1 - u_2)) \\ &\leq \|u_1 - u_2\|^2(1 + \rho^2 \|A\|^2 - 2\rho\alpha). \end{aligned}$$

U zadnjem koraku smo iskoristili koercitivnost. Sad je jasno da za dovoljno malo $\rho > 0$ možemo postići

$$1 + \rho^2 \|A\|^2 - 2\rho\alpha < 1;$$

dovoljno je uzeti $0 < \rho < 2\alpha/\|A\|^2$. Time je prslkavanje W_ρ kontrakcija i po Teoremu 2.21 ima jedinstvenu fiksnu točku u : $u = W_\rho(u) = P_K(u - \rho(Au - f))$. Time je egzistencija rješenja dokazana. \square

Pogledajmo sada primjer zadaće:

$$-\Delta u + u = f \quad \text{u } \Omega \quad (2.104)$$

$$u \geq 0, \quad \frac{\partial u}{\partial \mathbf{n}} \geq g \quad \text{na } \partial\Omega \quad (2.105)$$

$$u\left(\frac{\partial u}{\partial \mathbf{n}} - g\right) = 0 \quad \text{s.s. na } \Omega. \quad (2.106)$$

Uvedim skup

$$K = \{v \in H^1(\Omega) : v \geq 0 \text{ na } \partial\Omega\}.$$

Skup K je zatvoren (neprekidnost operatora traga) i konveksan potskup od $V = H^1(\Omega)$. Za svako $v \in K$ imamo

$$\int_{\Omega} (-\Delta u + u)v \, dx = \int_{\Omega} f v \, dx,$$

te parcijalnom integracijom

$$-\int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} v \, dS + \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx = \int_{\Omega} f v \, dx.$$

Posebno za $v = u \in K$ imamo

$$-\int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} u \, dS + \int_{\Omega} (\nabla u \cdot \nabla u + u^2) \, dx = \int_{\Omega} f u \, dx$$

i oduzimanjem

$$\int_{\Omega} (\nabla u \cdot \nabla(v - u) + u(v - u)) \, dx = \int_{\Omega} f(v - u) \, dx + \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}}(v - u) \, dS.$$

Koristeći (2.105) i (2.106) dobivamo

$$\int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}}(v - u) \, dS = \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}} v - g u \, dS \geq \int_{\partial\Omega} g(v - u) \, dS.$$

Time dolazimo do varijacijske nejednadžbe: Naći $u \in K$ takvo da je

$$\forall v \in K, \quad a(u, v - u) \geq \langle F, v - u \rangle, \quad (2.107)$$

gdje je

$$a(u, v) = \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx, \quad \langle F, v \rangle = \int_{\Omega} f v \, dx + \int_{\partial\Omega} g v \, dS.$$

Uz pretpostavke $f \in L^2(\Omega)$ i $g \in L^2(\partial\Omega)$ taj problem evidentno zadovoljava uvjete Teorema 2.22 i ima jedinstveno rješenje. Pokažimo da je to slabo rješenje dobra generalizacija klasičnog rješenja.

Odaberimo u varijacijskoj nejednadžbi $v = u + w$, gdje je $w \in \mathcal{D}(\Omega)$. Tada je

$$\forall w \in \mathcal{D}(\Omega), \quad a(u, w) \geq \int_{\Omega} f w \, dx.$$

Oдавдје се лако доказује да је диференцијална једнакдба (2.104) задовољена скоро свуда. Сада парцијалном интеграцијом из диференцијалне једнакдбе поново мождмо извести

$$a(u, v - u) = \int_{\Omega} f(v - u) \, dx + \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}}(v - u) \, dS.$$

Искористимо ли варијацијску неједнакост, добивамо

$$\int_{\Omega} f(v - u) \, dx + \int_{\partial\Omega} \frac{\partial u}{\partial \mathbf{n}}(v - u) \, dS \geq \langle F, v - u \rangle$$

odnosno

$$\int_{\partial\Omega} \left(\frac{\partial u}{\partial \mathbf{n}} - g \right) (v - u) dS \geq 0.$$

Ovdje, uzimajući $v = u + w$, $w \geq 0$ na $\partial\Omega$ prvo dobivamo

$$\frac{\partial u}{\partial \mathbf{n}} - g \geq 0 \quad \text{na } \partial\Omega.$$

Odavde i zbog $u \geq 0$ na $\partial\Omega$, uz izbor $v = u/2$, dobivamo

$$u \left(\frac{\partial u}{\partial \mathbf{n}} - g \right) = 0 \quad \text{s.s. na } \Omega.$$

Time je u potpunosti dokazano da je zadaća (2.107) generalizacija zadaće (2.104)–(2.106).

2.11 Zadaci

1. Dokažite da za svakih n brojeva a_1, \dots, a_n vrijedi

$$\left(\sum_{i=1}^n |a_i| \right)^2 \leq n \sum_{i=1}^n |a_i|^2. \quad (2.108)$$

Uputa: Treba iskoristiti konveksnost kvadratne funkcije i indukcijom po n dokazati

$$\left(\frac{1}{n} \sum_{i=1}^n |a_i| \right)^2 \leq \frac{1}{n} \sum_{i=1}^n |a_i|^2.$$

Ta se nejednakost zove *Jensenova* i vrijedi za svaku konveksnu funkciju $f(x)$, dok smo je mi iskazali samo za $f(x) = x^2$. Generalizirajte.

2. Dokažite da je s (2.19) dana jedna polunorma na $H^1(\Omega)$.
3. Neka je prostor V definiran s (2.32). Pokažite da je $|\cdot|_{H^1(\Omega)}$ norma na V , ekvivalentana s $H^1(\Omega)$ normom.
4. Pokažite da je rješenje varijacijske zadaće (2.33), koje je klase $C^2(\bar{\Omega})$, ujedno klasično rješenje zadaće (2.30), ako je zadovoljen uvjet kompatibilnosti (2.31). Koja je interpretacija slabog rješenja ako uvjet kompatibilnosti nije zadovoljen?
5. Dokažite da je dovoljno glatko rješenje varijacijske zadaće (2.33) ujedno i klasično rješenje zadaće (2.30) samo ako je zadovoljen uvjet kompatibilnosti (2.31).
6. Definirajte slabo rješenje zadaće s Robinovim rubnim uvjetom

$$\begin{cases} -\Delta u = f & \text{u } \Omega \\ \frac{\partial u}{\partial \mathbf{n}} + \alpha u = g & \text{na } \partial\Omega. \end{cases}$$

Uz koje uvjete na $\alpha(\mathbf{x})$ možemo primijeniti Lax-Milgramovu lemu?

Bibliografija

- [1] Ibrahim Aganović. *Uvod u rubne zadaće mehanike kontinuuma*. Element, Zagreb, 2003.
- [2] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 1998.
- [3] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 1998.
- [4] Svetozar Kurepa. *Funkcionalna analiza. Elementi teorije operatora*. Školska knjiga, Zagreb, 1990.
- [5] Sibe Mardešić. *Matematička analiza. U n -dimenzionalnom realnom prostoru. Prvi dio*. Školska knjiga, Zagreb, 1991.
- [6] Jindřich Nečas. *Les méthodes directes en théorie des équations elliptiques*. Academia, Praha, and Masson et Cie, Editeurs, Paris, 1967.

3

Konstrukcija prostora konačnih elemenata u $H^1(\Omega)$

Promatramo linearnu varijacijsku zadaću: naći $u \in V$ takav da je

$$\forall v \in V, \quad a(u, v) = \langle F, v \rangle, \quad (3.1)$$

gdje je V Hilbertov prostor, a bilinearna forma $a(\cdot, \cdot)$ i funkcional F zadovoljavaju uvjete Lax-Milgramove leme (Teorem 2.1). Aproximacijska zadaća glasi: naći $u_h \in V_h$ takav da je

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

gdje je $V_h \subset V$ konačnodimenzionalan potprostor. U ovom se poglavlju bavimo konstrukcijom prostora konačnih elemenata $V_h \subset H^1(\Omega)$.

Prvi korak u konstrukciji prostora V_h je dekompozicija (triangulacija) domene na jednostavne podskupove, uglavnom poliedre. Funkcije iz V_h su na svakom podskupu triangulacije polinomi određenog stupnja. Zatim se pokazuje da je za inkluziju $V_h \subset H^1(\Omega)$ potrebno da su funkcije iz V_h neprekidne. Neprekidnost se postiže pažljivim odabirom stupnjeva slobode za koje se uzimaju vrijednosti funkcije u simetrično odabranim točkama. Konačno, u prostoru V_h moramo odabrati jednu bazu koju je jednostavno numerički konstruirati i koja vodi na linearni sustav s dobro uvjetovanom matricom.

3.1 Osnovna svojstva prostora konačnih elemenata

Varijacijska zadaća (3.1) postavljena je u domeni $\Omega \subset \mathbb{R}^d$, $d = 1, 2$ ili 3 , za koju radi jednostavnosti pretpostavljamo da je ograničen, poliedarski skup (poligonalan skup u \mathbb{R}^2 ili otvoreni interval u \mathbb{R}).

1. Prvi korak u konstrukciji prostora V_h je triangulacija domene Ω . Triangulacija domene Ω je svaka konačna familija \mathcal{T}_h podskupova od $\bar{\Omega}$ koja ima ova svojstva:

(T1) $\bar{\Omega} = \cup_{K \in \mathcal{T}_h} K$;

(T2) Svaki $K \in \mathcal{T}_h$ je zatvoren poliedarski skup i $\text{Int}(K) \neq \emptyset$;

(T3) Za svaka dva različita $K_1, K_2 \in \mathcal{T}_h$ vrijedi $\text{Int}(K_1) \cap \text{Int}(K_2) = \emptyset$;

To su minimalne pretpostavke o triangulaciji. Skupove iz \mathcal{T}_h nazivamo **elementima**.

Triangulacija je dakle razbijanje domene na uniju disjunktih poliedarskih skupova (elemenata) koji ju posve prekrivaju. Elementi su najčešće trokuti ili četverokuti u \mathbb{R}^2 , te tetraedri, paralelepipedi i prizme u \mathbb{R}^3 , no vidjet ćemo da se koriste i elementi koji nisu poligonalni. Da bismo imali dobra aproksimacijska svojstva triangulacija mora zadovoljavati, pored T1–T3, još neka svojstva koja ćemo navesti kasnije.

2. Svakom elementu $K \in \mathcal{T}_h$ pridružujemo jedan konačnodimenzionalni linearni prostor funkcija koji označavamo s P_K . Zatim definiramo prostor

$$X_h = \{v: \bar{\Omega} \rightarrow \mathbb{R}: \forall K \in \mathcal{T}_h, \quad v|_K \in P_K\}.$$

Funkcije¹ iz X_h nisu neprekidne jer općenito imaju skokove na granicama susjednih elemenata. Stoga ćemo promatrati podskup $V_h \subset X_h$ koji se sastoji od neprekidnih funkcija i pokazat ćemo da je on podskup prostora $H^1(\Omega)$. To govori sljedeća lema:

Lema 3.1 Neka je $P_K \subset H^1(K)$ za svako $K \in \mathcal{T}_h$ i neka je $V_h = X_h \cap C(\bar{\Omega})$. Tada je

$$V_h \subset H^1(\Omega), \quad V_{oh} = \{v \in V_h: v = 0 \text{ na } \partial\Omega\} \subset H_0^1(\Omega).$$

Dokaz. Evidentno je $X_h \subset L^2(\Omega)$. Treba samo dokazati da svako $v \in V_h$ ima slabu derivaciju u $L^2(\Omega)$. To znači da za $i = 1, 2, \dots, n$ postoje $v_i \in L^2(\Omega)$ takve da je

$$\forall \phi \in \mathcal{D}(\Omega), \quad \int_{\Omega} v \partial_i \phi \, dx = - \int_{\Omega} v_i \phi \, dx.$$

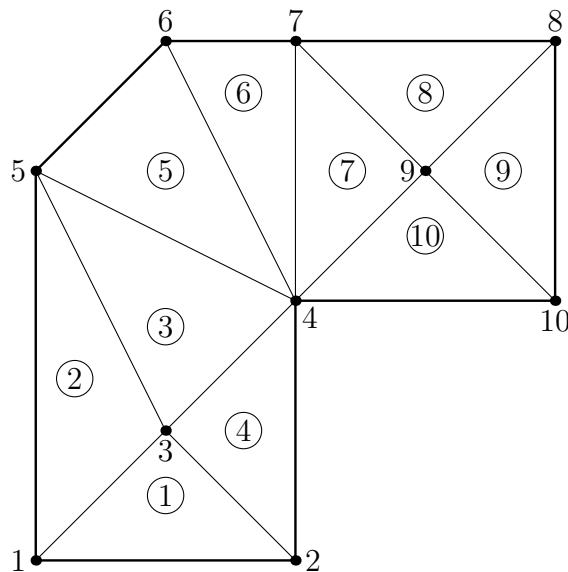
Parcijalnom integracijom na svakom elementu dobivamo

$$\int_K v \partial_i \phi \, dx = - \int_K \partial_i(v|_K) \phi \, dx + \int_{\partial K} v|_K \phi n_{i,K} \, d\sigma,$$

gdje je $n_{i,K}$ komponenta jedinične vanjske normale na ∂K . Sumiranjem po svim elementima dobivamo

$$\int_{\Omega} v \partial_i \phi \, dx = - \int_{\Omega} v_i \phi \, dx + \sum_{K \in \mathcal{T}_h} \int_{\partial K} v|_K \phi n_{i,K} \, d\sigma,$$

¹Elementi prostora X_h , strogo govoreći, nisu funkcije jer su višeznačne na granicama elemenata. Bilo bi ispravnije pisati $X_h = \prod_{K \in \mathcal{T}_h} P_K$.

Slika 3.1: Primjer triangulacije trokutima poligonalne domene u \mathbb{R}^2 .

gdje smo v_i definirali formulom $v_i|_K = \partial_i(v|_K) \in L^2(K)$, pa je očito $v_i \in L^2(\Omega)$. Da bi v_i bila slaba derivacija, mora suma integrala po rubovima elemenata iščezavati. No to je istina jer je ϕ nula na $\partial\Omega$, a doprinosi od dva susjedna elementa se poništavaju zbog neprekidnosti funkcije v . \square

Uvjet neprekidnosti moguće je postići samo ako postoji određena usklađenost prostora P_K , što će voditi na nova ograničenja na elemente K i prostore P_K .

3. Sljedeći bitan element metode konačnih elementa je zahtjev da prostori P_K sadrže polinome (ili funkcije bliske polinomima). To je važno iz praktičnog razloga – jednostavnosti računanja s polinomima, ali se pokazuje i kao ključan element u dokazu teorema konvergencije.

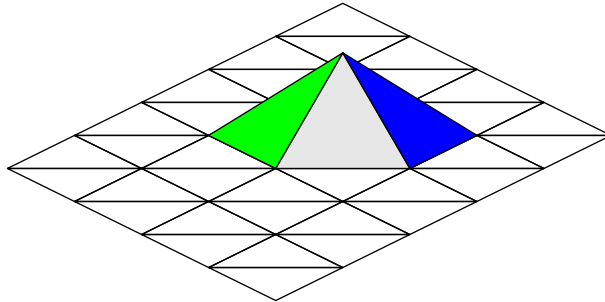
4. Konačno, u prostoru V_h mora postojati jedna *kanonska baza* koja se sastoji od funkcija s malim nosačem. To će svojstvo voditi na matricu krutosti s vrlo malo elemenata različitih od nule, što je odlučujući element u numeričkom rješavanju dobivenog sustava.

3.2 Primjer: \mathbb{P}_1 elementi na trokutima

Da bismo objasnili opisanu shemu konstrukcije prostora konačnih elemenata uzet ćemo jedan jednostavan primjer. Pretpostavimo da je $\Omega \subset \mathbb{R}^2$ ograničena poligonalna domena i da je \mathcal{T}_h jedna njena triangulacija trokutima, kao na Slici 3.1.

Uočimo da se triangulacija sastoji od 10 trokuta i da je broj vrhova triangulacije isto jednak 10. Svaki trokut sa susjednim trokutom ima zajedničku čitavu stranicu.

Funkcije iz prostora V_h , koji konstruiramo nad triangulacijom \mathcal{T}_h , su po dijelovima afine: preciznije, na svakom trokutu $K \in \mathcal{T}_h$ funkcija $v \in V_h$ je oblika $v(\mathbf{x}) = a_K + b_K x_1 + c_K x_2$. Nadalje, na granici susjednih elemenata funkcija $v(\mathbf{x})$ mora biti neprekidna. Ta će neprekidnost

Slika 3.2: Primjer nodalne bazne funkcije za $k = 1$.

smanjiti broj stupnjeva slobode koji nam stoje na raspolaganju. Naime, bez uvjeta neprekidnosti na svakom trokutu K imamo 3 stupnja slobode, odnosno tri koeficijenta koje možemo slobodno birati (a_K , b_K i c_K). Ukupan broj stupnjeva slobode je stoga $3N$, gdje je N broj trokuta u triangulaciji ($\dim(X_h) = 3N$). Uvjet neprekidnosti smanjuje broj stupnjeva slobode, jer se neki moraju iskoristiti za zadovoljenje uvjeta neprekidnosti, i prisiljava nas da potražimo adekvatije stupnjeve slobode od koeficijenata a_K , b_K i c_K .

Uzmimo dva susjedna trokuta iz \mathcal{T}_h , na primjer K_1 i K_4 , i označimo njihovu zajedničku stranicu s γ (to je stranica s vrhovima 2 i 3, vidi Sliku 3.1). Pogledajmo restrikciju funkcije $v \in V_h$ na γ : s v_1 označimo restrikciju uzetu iz K_1 , a s v_4 onu uzetu iz K_4 . Obje su te funkcije affine na γ i općenito međusobno različite. Da bismo ih učinili jednakima dovoljno je zahtijevati da su jednake u vrhovima stranice γ , i uz taj uvjet neprekidnost funkcije je zadovoljena.²

Sada se vrijednosti funkcije u vrhovima trokuta nameću kao adekvatni stupnjevi slobode. Naime, ako zadamo vrijednosti funkcije $v \in V_h$ u vrhovima svih trokuta, onda je ona time potpuno određena i neprekidna: Na svakom trokutu afina funkcija je posve određena svojim vrijednostima u vrhovima trokuta, a na stranicam koje povezuju susjedne trokute je neprekidna. Dimenzija prostora V_h je stoga jednaka broju vrhova triangulacije (u našem primjeru sa Slike 3.1 to je 10, odnosno za neprekidnost je trebalo 20 stupnjeva slobode).

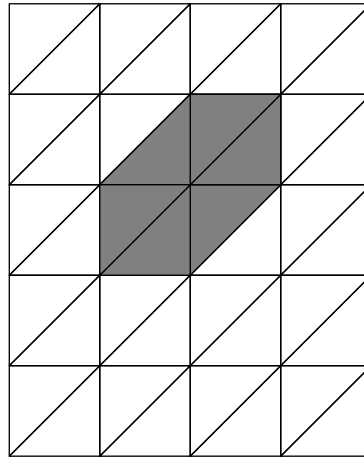
Konačno, da bismo efikasno konstruirali prostor V_h moramo naći jednu bazu u njemu. Tu se nameće tzv. *nodalna baza*, koja je pridružena vrhovima triangulacije. Uvedimo stoga skup svih vrhova triangulacije

$$\mathcal{V}_h = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_M\},$$

gdje je M broj svih vrhova trokuta u \mathcal{T}_h . Svakom vrhu \mathbf{a}_i pridružujemo baznu funkciju ϕ_i definiranu na sljedeći način:

$$\phi_i \in V_h, \quad \phi_i(\mathbf{a}_j) = \begin{cases} 1 & \text{za } i = j, \\ 0 & \text{za } i \neq j. \end{cases}$$

²Obje funkcije v_1 i v_4 su affine funkcije *jedne varijable* na γ pa su posve određene vrijednostima u dvije točke.



Slika 3.3: Trokuti koji čine nosač bazne funkcije.

Bazna funkcija ϕ_i je afina na svakom trokutu i određena je svojim vrijednostima u vrhovima trokuta. Oblik takve funkcije prikazan je Slici 3.2.

Ona je različita od nule samo na onim trokutima koji imaju \mathbf{a}_i kao jedan svoj vrh i stoga ima *mali nosač*, vidi Sliku 3.3.

Uočimo da smo neprekidnost funkcija iz V_h postigli zahvaljujući tome što svaka dva susjedna trokuta imaju ili zajedničku čitavu stranicu, ili zajednički samo jedan vrh. Situacija prikazana na Slici 3.4 nije dozvoljena, jer bi bila narušena neprekidnost funkcija iz V_h i ne bismo imali konformnu aproksimaciju, odnosno $V_h \subset H^1(\Omega)$.

Stoga možemo definirati dodatan zahtjev koji mora zadovoljavati svaka triangulacija poligonalne domene. Formuliramo ga samo za slučaj $d \leq 3$:

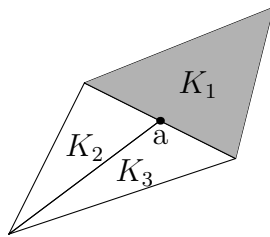
(T4) Svaka dva elementa K_1 i K_2 iz \mathcal{T}_h za koje je $K_1 \cap K_2 \neq \emptyset$ imaju ili zajedničku stranicu, ili zajednički brid ili zajednički vrh.

3.3 Simplicijalni elementi

Konstrukcija iz prethodne točke lako se generalizira na 3 i više dimenzija te na polinome višeg stupnja. Promatrat ćemo prostore dimenzije $d = 1, 2$ i 3 . Kanonsku bazu u \mathbb{R}^d označavamo s $(\mathbf{e}_i)_{i=1}^d$, a s \mathbb{P}_k označavamo prostor svih polinoma stupnja manjeg ili jednakog k , u \mathbb{R}^d , tj. u varijablama x_1, \dots, x_d . Koristit ćemo i oznaku

$$\mathbb{P}_k(K) = \{p|_K : p \in \mathbb{P}_k\}.$$

Neka su $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{d+1}$ točke u \mathbb{R}^d koje ne leže u istoj hiperravnini. d -simpleks K , razapet



Slika 3.4: Nedoželjena situacija

točkama $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{d+1}$, je konveksna ljuska točaka $\mathbf{a}_1, \dots, \mathbf{a}_{d+1}$:

$$K = \left\{ \mathbf{x} = \sum_{i=1}^{d+1} \lambda_i \mathbf{a}_i : 0 \leq \lambda_i \leq 1, \quad 1 \leq i \leq d+1, \quad \sum_{i=1}^{d+1} \lambda_i = 1 \right\}.$$

Točke \mathbf{a}_i nazivamo vrhovima simpleksa. Evidentno je 2-simpleks trokut, dok je 3-simpleks tetraedar.

Pretpostavimo da imamo triangulaciju domene koja se sastoji od simpleksa: trokuta u dvije dimenzije i tetraedara u tri. Na takvim elementima ćemo promatrati polinome stupnja $k \geq 1$ i reprezentirat ćemo ih na način pogodan za konstrukciju funkcija neprekidnih na čitavoj domeni Ω .

Neka je $\mathbf{a}_i = (a_{ji})_{j=1}^d$. Točke $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{d+1}$ ne leže u istoj hiperravnini ako i samo ako je matrica

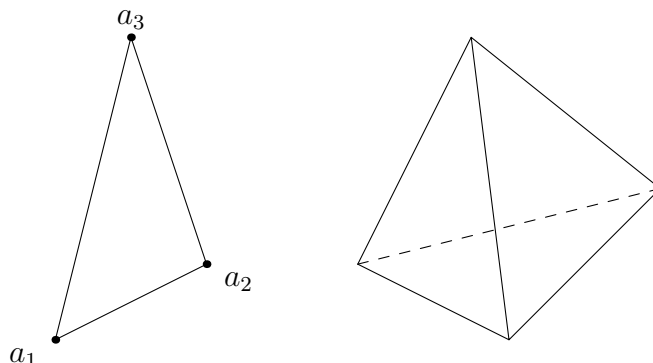
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n+1} \\ a_{21} & a_{22} & \cdots & a_{2n+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dd+1} \\ 1 & 1 & \cdots & 1 \end{bmatrix} \quad (3.2)$$

regularna.

Baricentričke koordinate $\lambda_i = \lambda_i(\mathbf{x})$, $1 \leq i \leq d+1$ bilo koje točke $\mathbf{x} \in \mathbb{R}^d$, u odnosu na točke $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{d+1}$, koje ne leže u istoj hiperravnini, je jedinstveno rješenje linearnog sustava

$$\begin{aligned} \sum_{i=1}^{d+1} a_{ji} \lambda_i &= x_j, \quad 1 \leq j \leq d \\ \sum_{i=1}^{d+1} \lambda_i &= 1. \end{aligned}$$

Matrica sustava je (3.2) pa je egzistencija i jedinstvenost koordinata osigurana uvjetom nekomplanarnosti.

Slika 3.5: Simpleksi tipa (1) u \mathbb{R}^2 i \mathbb{R}^3

Lako se provjerava da su baricentričke koordinate afine funkcije varijabli x_1, \dots, x_d (dakle, nalaze se u \mathbb{P}_1):

$$\lambda_i = \sum_{j=1}^d b_{ij}x_j + b_{id+1}, \quad 1 \leq i \leq d+1,$$

gdje je $B = (b_{ij})$ inverz matrice A iz (3.2).

Iz definicije baricentričkih koordinata se lako vidi da je i -ta koordinata λ_i afina funkcija definirana svojstvom da je jednaka 1 u vrhu \mathbf{a}_i , a nula u svim ostalim vrhovima simpleksa. Dakle, λ_i ima svojstva

$$\lambda_i \in \mathbb{P}_1, \quad \lambda_i(\mathbf{a}_j) = \begin{cases} 1 & \text{za } i = j \\ 0 & \text{za } i \neq j, \end{cases}$$

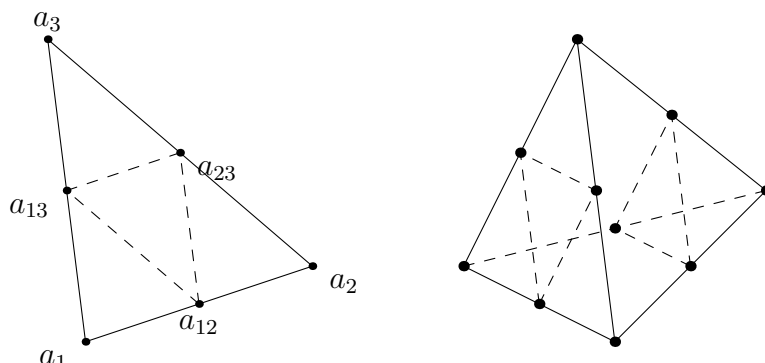
i ta ju svojstva potpuno određuju. Pored toga vidimo da svaki polinom $p \in \mathbb{P}_1$ možemo na jedinstven način prikazati u obliku

$$p(\mathbf{x}) = \sum_{i=1}^d p(\mathbf{a}_i)\lambda_i(\mathbf{x}). \quad (3.3)$$

Najjednostavniji konačni element konstruiran nad jednim d -simpleksom je onaj u kojem je $P_K = \mathbb{P}_1(K)$. Sve funkcije iz tog prostora možemo jednostavno reprezentirati pomoću formule (3.3). Vrijednosti $p(\mathbf{a}_i)$ nazivaju se **stupnjevi slobode** konačnog elementa (eng. *degrees of freedom*, kraće *dofs*), a funkcije λ_i su pripadne **lokalne bazne funkcije** pomoću kojih ćemo dobiti kanonsku bazu. Za skup svih stupnjeva slobode elementa K koristit ćemo oznaku Σ_K ; vrhove simpleksa \mathbf{a}_i nazivamo **nodalnim točkama** elementa.

Ovako konstruirani element nazivamo *d -simpleksom tipa (1)*. U $d = 2$ govorimo o trokutu tipa (1) ili *Courant-ovom trokutu*, a za $d = 3$ govorimo o tetraedru tipa (1).

d -simpleks tipa (1)
$P_K = \mathbb{P}_1(K), \quad \dim P_K = d + 1$
$\Sigma_K = \{p(\mathbf{a}_i) : 1 \leq i \leq d + 1\}.$

Slika 3.6: Simpleksi tipa (2) u \mathbb{R}^2 i \mathbb{R}^3

d -simpleks tipa (2) sadrži sve polinome stupnja manjeg ili jednakog dva: $P_K = \mathbb{P}_2(K)$. Da bismo definirali stupnjeve slobode uvedimo dodatne nodalne točke $\mathbf{a}_{ij} = (\mathbf{a}_i + \mathbf{a}_j)/2$, za $1 \leq i, j \leq d+1$. To su polovišta bridova d -simpleksa. Uočimo da je broj koeficijenata u polinomu drugog stupnja jednak broju nodalnih točaka, tako da vrijednosti funkcije u nodalnim točkama možemo uzeti kao stupnjeve slobode. To ćemo pokazati eksplicitnom konstrukcijom baznih funkcija.

Budući da su baricentričke koordinate afine funkcije koje u vrhovima simpleksa primaju vrijednost 0 ili 1, vidimo da je

$$\lambda_k(\mathbf{a}_{ij}) = \frac{1}{2}(\delta_{ki} + \delta_{kj}).$$

Sada je lako uočiti da funkcije

$$\lambda_i(2\lambda_i - 1), \quad i = 1, \dots, d+1,$$

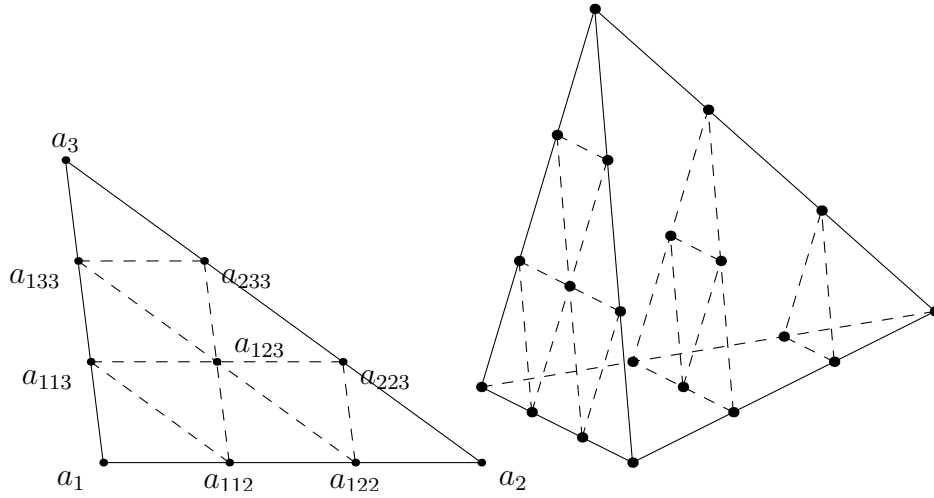
imaju svojstvo da su jednake jedan u \mathbf{a}_i , a nula u svim ostalim \mathbf{a}_k te \mathbf{a}_{ij} . Nadalje funkcije

$$\lambda_i\lambda_j, \quad i < j, \quad i, j = 1, \dots, d+1,$$

jednake su $1/4$ u \mathbf{a}_{ij} , a nula u svim drugim točkama. Na temelju tih činjenica lako konstruiramo polinom drugog stupnja koji ima zadane vrijednosti u točkama \mathbf{a}_i , \mathbf{a}_{ij} . Dobivamo

$$\forall p \in \mathbb{P}_2, \quad p = \sum_{i=1}^{d+1} \lambda_i(2\lambda_i - 1)p(\mathbf{a}_i) + 4 \sum_{\substack{i,j=1 \\ i < j}}^{d+1} \lambda_i\lambda_j p(\mathbf{a}_{ij}).$$

d -simpleks tipa (2)
$P_K = \mathbb{P}_2(K), \quad \dim P_K = (d+1)(d+2)/2$
$\Sigma_K = \{p(\mathbf{a}_i), 1 \leq i \leq d+1, p(\mathbf{a}_{ij}), 1 \leq i < j \leq d+1\}$.

Slika 3.7: Simpleksi tipa (3) u \mathbb{R}^2 i \mathbb{R}^3

Kako na svakoj stranici imamo 3 (u \mathbb{R}^2) odnosno 6 (u \mathbb{R}^3) simetrično raspoređenih dodatnih točaka, one se kod susjednih elemenata poklapaju i na jedinstven način određuju restrikciju polinoma drugog stupnja na zajedničku stranicu. Na taj način postizemo globalnu neprekidnost funkcije definirane svojim vrijednostima u nodalnim točkama.

d -simpleksi tipa (3) definiraju se analogno. Pored vrhova simpleksa uvedu se dodatne nodalne točke:

$$\mathbf{a}_{ij} = \frac{1}{3}(2\mathbf{a}_i + \mathbf{a}_j), \quad \text{za } i \neq j, \quad \mathbf{a}_{ijk} = \frac{1}{3}(\mathbf{a}_i + \mathbf{a}_j + \mathbf{a}_k), \quad \text{za } i < j < k,$$

i dobije se da za svako $p \in \mathbb{P}_3$ vrijedi

$$p = \frac{1}{2} \sum_{i=1}^{d+1} \lambda_i (3\lambda_i - 1)(3\lambda_i - 2) p(\mathbf{a}_i) + \frac{9}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{d+1} \lambda_i \lambda_j (3\lambda_i - 1) p(\mathbf{a}_{ij}) + 27 \sum_{i < j < k} \lambda_i \lambda_j \lambda_k p(\mathbf{a}_{ijk}). \quad (3.4)$$

d -simpleks tipa (3)
$P_K = \mathbb{P}_3(K), \quad \dim P_K = (d+1)(d+2)(d+3)/6$
$\Sigma_K = \{p(\mathbf{a}_i), 1 \leq i \leq d+1, p(\mathbf{a}_{ij}), 1 \leq i, j \leq d+1, i \neq j, p(\mathbf{a}_{ijk}), 1 \leq i < j < k \leq d+1\}$.

Ova se konstrukcija može nastaviti za proizvoljno $k \in \mathbb{N}$. Globalna neprekidnost se postiže na isti način kao za polinome prvog i drugog stupnja.

Napomena 3.1 *Konačni element se strogo definira kao trojka (K, P_K, Σ_K) , gdje je K skup, P_K je konačnodimenzionalni prostor funkcija na K i Σ_K skup stupnjeva slobode. Stupnjevi slobode su općenito funkcionali na prostoru funkcija P_K . Kod simplicijalnih elemenata K je simpleks, P_K je odgovarajući prostor polinoma, a Σ_K je prostor funkcionala oblika $\phi(p) = p(\mathbf{a}_i)$, gdje su \mathbf{a}_i nodalne točke simpleksa.*

Kao stupnjevi slobode u konačnom elementu mogu se pojaviti i vrijednosti derivacija u nodalnim točkama (Hermiteovi elementi) i vrijednosti integralnog tipa kao $\int_K p(x) dx$. Kada su jedini stupnjevi slobode vrijednosti funkcije u nodalnim točkama, onda kažemo da je element Lagrangeov. Simplicijalni elementi su, dakle, predstavnici Lagrangeovih elemenata.

Osnovno svojstvo koje konačni element mora zadovoljavati je unisolventnost: Stupnjevi slobode u Σ_K moraju na jedinstven način određivati funkciju iz P_K .

Neka je zadana triangulacija pomoću d -simpleksa tipa (k) , $k \geq 1$, koja zadovoljava svojstva (T1)–(T4). Simetričan raspored nodalnih točaka u svakom elementu i uvjet poklapanja stranica susjednih elemenata ima za posljedicu da se nodalne točke dva susjedna elementa podudaraju na zajedničkoj stranici.

Označimo skup svih nodalnih točaka u cijeloj triangulaciji s \mathcal{N}_h . On se sastoji od svih nodalnih točaka svih elemenata triangulacije. Prostor X_h na triangulaciji s d -simpleksima tipa (k) definiran je kao prostor svih funkcija na $\bar{\Omega}$ koje su na svakom elementu polinomi stupnja najviše k . Te funkcije općenito imaju skokove na granicama elemenata. U prostoru X_h promatramo samo one funkcije koje su jednoznačno definirane u točkama skupa \mathcal{N}_h . Drugim riječima, to su funkcije koje u nodalnim točkama koje se nalaze na granici dva ili više elementa primaju istu vrijednost u svim elementima koji ih sadrže. Skup svih takvih funkcija je naš prostor konačnih elemenata $V_h \subset X_h$. Svaka funkcija iz V_h posve je određena svojim vrijednostima u skupu \mathcal{N}_h , pa vrijednosti funkcije u tim točkama predstavljaju globalne stupnjeve slobode. Skup globalnih stupnjeva slobode označavamo s

$$\Sigma_h = \{v_h(\mathbf{b}) : \mathbf{b} \in \mathcal{N}_h\}.$$

Evidentno je prostor V_h linearan i $\dim(V_h) = \text{card}(\Sigma_h)$. Time smo dokazali:

Teorem 3.1 Neka je V_h prostor konačnih elemenata konstruiran pomoću d -simpleksa tipa (k) , za $k \geq 1$. Tada vrijedi

$$V_h \subset C(\bar{\Omega}) \cap H^1(\Omega).$$

Konačno, kanonsku bazu u V_h definiramo na sljedeći način: Indeksiramo skup \mathcal{N}_h tako da bude $\mathcal{N}_h = \{\mathbf{b}_1, \mathbf{b}_2, \dots\}$ i definiramo

$$w_i \in V_h, \quad w_i(\mathbf{b}_j) = \delta_{ij}, \quad i, j = 1, \dots, \text{card}(\Sigma_h).$$

Funkcije $(w_1, \dots, w_{\text{card}(\Sigma_h)})$ očito čine jednu bazu koju nazivamo **nodalna baza**. Lako se provjerava da bazne funkcije imaju malen nosač: funkcija w_i ima za nosač uniju svih elemenata koji sadrže nodalnu točku \mathbf{b}_i .

Napomena 3.2 Ako imamo Dirichletov rubni uvjet na dijelu granice $\Gamma_D \subset \partial\Omega$, onda se funkcije iz prostora V moraju poništavati na Γ_D . Da bismo imali $V_h \subset V$ i funkcije iz V_h se moraju poništavati na Γ_D . To se postiže tako da se Dirichletova granica egzaktno prekrije s konačnim elementima te da se u svim nodalnim točkama koje leže na Γ_D stavi vrijednost funkcije $v \in V_h$ na nulu. Tada će se v poništavati na čitavom Γ_D i aproksimacija će biti konformna. Stavljanje na nulu funkcije u jednom broju nodalnih točaka znači smanjenje broja stupnjeva slobode.

3.4 Interpolacijski operatori

Interpolacijski operator ima centralnu ulogu u ocjeni greške metode konačnih elemenata. Za sve Lagrangeove konačne elemente definira se na isti način. Prvo definiramo lokalni interpolacijski operator.

Definicija 3.1 Neka je (K, P_K, Σ_K) simplicijalni element tipa (k) , $k \geq 1$, te neka su p_1, \dots, p_N lokalne bazne funkcije, a $\mathbf{a}_1, \dots, \mathbf{a}_N$ nodalne točke. Tada preslikavanje $\Pi_K: C(K) \rightarrow P_K$, definirano formulom

$$(\Pi_K v)(\mathbf{x}) = \sum_{i=1}^N v(\mathbf{a}_i) p_i(\mathbf{x}) \quad (3.5)$$

nazivamo interpolacijski operator pridružen elementu (K, P_K, Σ_K) .

Interpolacijski operator funkciji $v(\mathbf{x})$ pridružuje funkciju iz prostora P_K (polinom stupnja $\leq k$) koja u nodalnim točkama prima iste vrijednosti kao i $v(\mathbf{x})$. Takva je funkcija jedinstvena, pa je operator dobro definiran.

Definicija 3.2 Neka je V_h prostor konačnih elemenata konstruiran pomoću simplicijalni element tipa (k) , $k \geq 1$, te neka su w_1, \dots, w_M globalne bazne funkcije, a $\mathbf{b}_1, \dots, \mathbf{b}_M$ nodalne točke triangulacije. Tada preslikavanje $\Pi_h: C(\bar{\Omega}) \rightarrow V_h$, definirano formulom

$$(\Pi_h v)(\mathbf{x}) = \sum_{i=1}^M v(\mathbf{b}_i) w_i(\mathbf{x}) \quad (3.6)$$

nazivamo (globalni) interpolacijski operator u prostoru V_h .

Globalni interpolacijski operator funkciji $v(\mathbf{x})$ pridružuje funkciju iz prostora V_h koja prima u nodalnim točkama triangulacije iste vrijednosti kao i $v(\mathbf{x})$. Globalni i lokalni interpolacijski operatori su u jednostavnoj vezi: za svako $v \in C(\bar{\Omega})$ i za svaki element K triangulacije \mathcal{T}_h vrijedi

$$(\Pi_h v)|_K = \Pi_K v. \quad (3.7)$$

U dimenziji $d \geq 2$ interpolacijski operator nije dobro definiran na prostoru $H^1(\Omega)$, $\Omega \subset \mathbb{R}^d$, stoga što ne vrijedi $H^1(\Omega) \subset C(\bar{\Omega})$ (ta je inkluzija istinita samo za $d = 1$). Zbog toga interpolacijski operator promatramo na potprostoru $H^1(\Omega) \cap C(\bar{\Omega})$ ili na nekom još užem potprostoru. U dvije i tri prostorne dimenzije vrijedi $H^2(\Omega) \subset C(\bar{\Omega})$, pa stoga najčešće operator Π_h promatramo na funkcijama iz $H^2(\Omega)$.

3.5 Afina familija elemenata

Simplicijalne je elemente moguće generirati polazeći od jednog elementa pomoću afinog preslikavanja. Ta se činjenica koristi u teoriji za dokazivanje greške interpolacije, ali i kod kodiranja metode. Obično se fiksira jedan element, takozvani **referentni element**, a svi drugi elementi triangulacije dobivaju se afinim preslikavanjem referentnog elementa.

Objasnimo to na primjeru trokuta tipa 2. Fiksirajmo jedan trokut tipa 2, \hat{K} , s vrhovima $\hat{\mathbf{a}}_i$ te polovištima stranica $\hat{\mathbf{a}}_{ij} = (\hat{\mathbf{a}}_i + \hat{\mathbf{a}}_j)/2$, $1 \leq i, j \leq 3$. Stupnjevi slobode su

$$\hat{\Sigma} = \{p(\hat{\mathbf{a}}_i), i = 1, 2, 3, p(\hat{\mathbf{a}}_{ij}), 1 \leq i < j \leq 3\},$$

a prostor funkcija je $\hat{P} = \mathbb{P}_2(\hat{K})$. Za svaki element K iz triangulacije postoji jedinstveno afino preslikavanje

$$F_K(\hat{\mathbf{x}}) = \mathbf{B}_K \hat{\mathbf{x}} + \mathbf{b}_K$$

koje je invertibilno i za koje je $F_K(\hat{\mathbf{a}}_i) = \mathbf{a}_i$, $i = 1, 2, 3$, gdje su \mathbf{a}_i vrhovi trokuta K . Takvo se preslikavanje lagano konstruira pomoću baricentričkih koordinata formulom:

$$F_K(\hat{\mathbf{x}}) = \sum_{i=1}^3 \lambda_i(\hat{\mathbf{x}}) \mathbf{a}_i.$$

Evidentno je

$$F_K(\hat{\mathbf{a}}_{ij}) = \mathbf{a}_{ij}, \quad 1 \leq i < j \leq 3,$$

pa je prirodno na K definirati prostor funkcija P_K na sljedeći način:

$$P_K = \{p: K \rightarrow \mathbb{R}: p = \hat{p} \circ F_K^{-1}, \hat{p} \in \hat{P}\}.$$

Budući da je preslikavanje afino, evidentno imamo

$$P_K = \mathbb{P}_2(K).$$

Zaključujemo da je da trokut tipa 2, (K, P_K, Σ_K) , možemo u potpunosti opisati pomoću referentnog trokuta tipa 2, $(\hat{K}, \hat{P}, \hat{\Sigma})$, i afinog preslikavanja F_K na sljedeći način:

$$\begin{aligned} K &= F_K(\hat{K}), \\ P_K &= \{p: K \rightarrow \mathbb{R}: p = \hat{p} \circ F_K^{-1}, \hat{p} \in \hat{P}\} \\ \Sigma_K &= \{p(F_K(\hat{\mathbf{a}}_i)), i = 1, 2, 3, p(F_K(\hat{\mathbf{a}}_{ij})), 1 \leq i < j \leq 3\}. \end{aligned}$$

Za elemente (K, P, Σ) i $(\hat{K}, \hat{P}, \hat{\Sigma})$ tada kažemo da su **afino ekvivalentni** (pokažite da se radi o relaciji ekvivalencije). Pripadni lokalni interpolacijski operatori su u sljedećem odnosu:

$$(\Pi_K v) \circ F_K = \Pi_{\hat{K}}(v \circ F_K).$$

Ista tvrdnja vrijedi za sve simplicijane elemente, pa govorimo da simplicijalni elementi čine afinu familiju elemenata. S praktične strane ta se činjenica koristi tako da se konstruira samo referentni element, a za svaki element triangulacije pamti se samo njegovo afino preslikavanje. Bazne se funkcije konstruiraju samo na referentnom elementu i formule numeričke integracije koriste se samo na referentnom elementu. Sva računanja na proizvoljnom elementu triangulacije K prenose se na referentni element putem preslikavanja F_K .

3.6 Zadaci

1. $\dim(\mathbb{P}_k) = \binom{n+k}{k}$. Ako K ima neprazan interior, onda \mathbb{P}_k i $\mathbb{P}_k(K)$ imaju istu dimenziju.
2. Dokažite da točke $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{d+1}$ ne leže u istoj hiperravnini ako i samo ako je matrica (3.2) regularna.

4

Greška interpolacije

Ocjena greške interpolacije u prostorima Soboljeva tehnički je dosta složena te ju stoga nećemo dokazivati u najvećoj općenitosti. Zadovoljit ćemo se promatranjem afine familije Lagrangeovih konačnih elemenata.

Pri konstrukciji prostora konačnih elemenata $V_h \subset H^1(\Omega)$ polazimo od triangulacije \mathcal{T}_h domene Ω i referentnog konačnog elementa $(\hat{K}, \hat{P}, \hat{\Sigma})$. Za konstrukciju konačnog elementa na proizvoljnom elementu $K \in \mathcal{T}_h$ koristimo difeomorfizam $F_K: \hat{K} \rightarrow K$, i definiramo konačni element (K, P, Σ) pomoću funkcija $v \in P$ oblika $v = \hat{v} \circ F_K^{-1}$, gdje je $\hat{v} \in \hat{P}$, te nodalnih točaka na K koje su slike nodalnih točaka na referentnom elementu \hat{K} .

Nakon što je prostor konačnih elemenata V_h konstruiran, potreban nam je operator koji funkciji iz $H^1(\Omega)$, ili iz nekog njegovog podskupa, pridružuje element iz V_h . Jedna mogućnost je uzeti operator ortogonalne projekcije, no pokazuje se prirodnije uzeti interpolacijski operator Π_h koji je definiran na $H^1(\Omega) \cap C(\bar{\Omega})$. On ima poželjno svojstvo da je definiran lokalno, pomoću lokalnih interpolacijskih operatora Π_K , za $K \in \mathcal{T}_h$.

Greške interpolacije $\|u - \Pi_h u\|$ u različitim Soboljevlevim normama ocjenjuje se pomoću finoće triangulacije h , koja je definirana kao maksimalni dijаметar elementa u triangulaciji. Ta se ocjena svodi na lokalnu ocjenu greške interpolacije $\|u - \Pi_K u\|$, na svakom pojedinim elementu $K \in \mathcal{T}_h$, a izražava se pomoću geometrijskih svojstava elementa i stupnja polinoma u prostoru konačnih elemenata. Da bi ovisnost o geometriji elementa bila eksplicitna koristi se preslikavanje F_K kojim se problem transformira na referentni element čija je geometrija zadana, pa stoga geometrijska svojstva elementa K dolaze do izražaja samo kroz preslikavanje F_K . Ta tehnika olakšava ocjenu greške interpolacije, pogotovo u slučaju afinih elemenata, kada je svako prelikavanje F_K afino, pa njegov gradijent konstantan.

4.1 Lokalne ocjene

Budući da ćemo koristiti tehniku referentnog elementa, prvo moramo vidjeti kako se vrši zamjena varijabli u Soboljevlevim normama. Uzmimo stoga da su \hat{K} i K dva ograničena zatvorena skupa i da postoji afina bijekcija $F_K: \hat{K} \rightarrow K$. Za takve skupove ćemo govoriti da su **afino**

ekvivalentni. Neka je

$$F_K(\hat{\mathbf{x}}) = B_K \hat{\mathbf{x}} + \mathbf{b}_K, \quad (4.1)$$

gdje je $B_K \in \mathbb{R}^{d \times d}$ regularna matrica, $\mathbf{b}_K \in \mathbb{R}^d$ vektor, te $K = F_K(\hat{K})$. Ako je $v(\mathbf{x}) = (\hat{v} \circ F_K^{-1})(\mathbf{x})$ funkcija definirana na K , generirana funkcijom $\hat{v} \in L^2(\hat{K})$, onda se lako vidi da je $v \in L^2(K)$ i zamjenom varijabli u integralu dobivamo,

$$\begin{aligned} \int_K |v(\mathbf{x})|^2 d\mathbf{x} &= \int_{F_K(\hat{K})} |(\hat{v} \circ F_K^{-1})(\mathbf{x})|^2 d\mathbf{x} = \int_{\hat{K}} |\hat{v}(\hat{\mathbf{x}})|^2 |\det(\nabla F_K(\hat{\mathbf{x}}))| d\hat{\mathbf{x}} \\ &= |\det(B_K)| \int_{\hat{K}} |\hat{v}(\hat{\mathbf{x}})|^2 d\hat{\mathbf{x}}. \end{aligned}$$

Time smo dobili prvu od traženih transformacija normi između promatranog i referentnog elementa:

$$\|v\|_{L^2(K)} = |\det(B_K)|^{1/2} \|\hat{v}\|_{L^2(\hat{K})}. \quad (4.2)$$

Pokažimo sada kako se transformiraju polunorme višeg reda. Iskoristit ćemo vezu $\hat{\mathbf{x}} = F_K^{-1}(\mathbf{x})$ i $\nabla F_K^{-1}(\mathbf{x}) = B_K^{-1}$. Za prve derivacije dobivamo

$$\frac{\partial v}{\partial x_i} = \sum_{k=1}^d \frac{\partial \hat{v}}{\partial \hat{x}_k} \frac{\partial \hat{x}_k}{\partial x_i} = \sum_{k=1}^d \frac{\partial \hat{v}}{\partial \hat{x}_k} (B_K^{-1})_{k,i} = \nabla \hat{v} \cdot B_K^{-1} \mathbf{e}_i$$

(gdje je \mathbf{e}_i vektor kanonske baze). Kako je matrica B_K konstantna sličnu formulu dobivamo za druge i više derivacije. Na primjer,

$$\frac{\partial^2 v}{\partial x_j \partial x_i} = \sum_{l=1}^d \sum_{k=1}^d \frac{\partial^2 \hat{v}}{\partial \hat{x}_l \partial \hat{x}_k} \frac{\partial \hat{x}_l}{\partial x_j} (B_K^{-1})_{k,i} = \sum_{l=1}^d \sum_{k=1}^d \frac{\partial^2 \hat{v}}{\partial \hat{x}_l \partial \hat{x}_k} (B_K^{-1})_{l,j} (B_K^{-1})_{k,i}.$$

Formula se na evidentan način generalizira na treće i više derivacije. Kompaktno ju zapisujemo u obliku

$$D^2 v(\mathbf{x}) \mathbf{e}_i \cdot \mathbf{e}_j = D^2 \hat{v}(\hat{\mathbf{x}}) B_K^{-1} \mathbf{e}_i \cdot B_K^{-1} \mathbf{e}_j,$$

gdje smo matricu drugih derivacija označili s $D^2 v$, odnosno $D^2 \hat{v}$.

Koristeći Cauchy - Schwartzovu nejednakost, sada imamo

$$\begin{aligned} \int_K \left| \frac{\partial v}{\partial x_i}(\mathbf{x}) \right|^2 d\mathbf{x} &= \int_{F_K(\hat{K})} \left| \sum_{k=1}^d \left(\frac{\partial \hat{v}}{\partial \hat{x}_k} \circ F_K^{-1} \right)(\mathbf{x}) (B_K^{-1})_{k,i} \right|^2 d\mathbf{x} \\ &= |\det(B_K)| \cdot \int_{\hat{K}} \left| \sum_{k=1}^d \frac{\partial \hat{v}}{\partial \hat{x}_k}(\hat{\mathbf{x}}) (B_K^{-1})_{k,i} \right|^2 d\hat{\mathbf{x}} \\ &\leq |\det(B_K)| \cdot \sum_{k=1}^d |(B_K^{-1})_{k,i}|^2 \int_{\hat{K}} \sum_{k=1}^d \left| \frac{\partial \hat{v}}{\partial \hat{x}_k}(\hat{\mathbf{x}}) \right|^2 d\hat{\mathbf{x}} \\ &= |\det(B_K)| \cdot \|B_K^{-1} \mathbf{e}_i\|^2 \|\hat{v}\|_{H^1(\hat{K})}^2. \end{aligned}$$

Po definiciji operatorske norme matrice,

$$\|B_K^{-1}\| = \sup_{\mathbf{e} \neq 0} \frac{\|B_K^{-1}\mathbf{e}\|}{\|\mathbf{e}\|}$$

možemo ocijeniti $\|B_K^{-1}\mathbf{e}_i\| \leq \|B_K^{-1}\|$, pa stoga imamo

$$|v|_{H^1(K)} \leq C|\det(B_K)|^{1/2} \cdot \|B_K^{-1}\| \cdot |\hat{v}|_{H^1(\hat{K})}. \quad (4.3)$$

Analogno imamo,

$$\begin{aligned} \int_K \left| \frac{\partial^2 v(\mathbf{x})}{\partial x_j \partial x_i} \right|^2 d\mathbf{x} &= \int_{F_K(\hat{K})} \left| \sum_{k=1}^d \sum_{l=1}^d \left(\frac{\partial^2 \hat{v}}{\partial \hat{x}_l \partial \hat{x}_k} \circ F_K^{-1} \right)(\mathbf{x}) (B_K^{-1})_{l,j} (B_K^{-1})_{k,i} \right|^2 d\mathbf{x} \\ &= |\det(B_K)| \cdot \int_{\hat{K}} \left| \sum_{k=1}^d \sum_{l=1}^d \frac{\partial^2 \hat{v}}{\partial \hat{x}_l \partial \hat{x}_k}(\hat{\mathbf{x}}) (B_K^{-1})_{l,j} (B_K^{-1})_{k,i} \right|^2 d\hat{\mathbf{x}} \\ &\leq |\det(B_K)| \cdot \sum_{k=1}^d |(B_K^{-1})_{k,i}|^2 \int_{\hat{K}} \sum_{k=1}^d \left| \sum_{l=1}^d \frac{\partial^2 \hat{v}}{\partial \hat{x}_l \partial \hat{x}_k}(\hat{\mathbf{x}}) (B_K^{-1})_{l,j} \right|^2 d\hat{\mathbf{x}} \\ &\leq |\det(B_K)| \cdot \sum_{k=1}^d |(B_K^{-1})_{k,i}|^2 \sum_{l=1}^d |(B_K^{-1})_{l,j}|^2 \int_{\hat{K}} \sum_{k,l=1}^d \left| \frac{\partial^2 \hat{v}}{\partial \hat{x}_l \partial \hat{x}_k}(\hat{\mathbf{x}}) \right|^2 d\hat{\mathbf{x}} \\ &\leq C|\det(B_K)| \cdot \|B_K^{-1}\mathbf{e}_i\|^2 \cdot \|B_K^{-1}\mathbf{e}_j\|^2 \cdot |\hat{v}|_{H^2(\hat{K})}^2. \end{aligned}$$

Konstanta C dolazi iz ekvivalencije posljednjeg integrala i Soboljevljeve polunorme drugog reda te ovisi samo o d . Time dolazimo do ocjene

$$|v|_{H^2(K)} \leq C|\det(B_K)|^{1/2} \cdot \|B_K^{-1}\|^2 \cdot |\hat{v}|_{H^2(\hat{K})}. \quad (4.4)$$

Ocjena (4.4) se lako generalizira, tako da za proizvoljno $m \geq 0$ vrijedi,

$$|v|_{H^m(K)} \leq C|\det(B_K)|^{1/2} \cdot \|B_K^{-1}\|^m \cdot |\hat{v}|_{H^m(\hat{K})}, \quad (4.5)$$

gdje C ovisi samo o d i m .

Uočimo da se inverzne ocjene dobivaju na posve isti način, zamjenom uloga skupova K i \hat{K} , te zamjenom preslikavanja F_K^{-1} s F_K . Time smo dokazali sljedeći teorem.

Teorem 4.1 Neka su K i \hat{K} ograničeni afino ekvivalentni skupovi u \mathbb{R}^d . Ako je $v \in H^m(K)$ za neki $m \geq 0$, onda je $\hat{v} = v \circ F_K \in H^m(\hat{K})$ te postoji konstanta $C = C(m, d)$ takva da je

$$\forall v \in H^m(K), \quad |\hat{v}|_{H^m(\hat{K})} \leq C\|B_K\|^m |\det(B_K)|^{-1/2} |v|_{H^m(K)}, \quad (4.6)$$

gdje je B_K matrica afinog preslikavanja F_K iz (4.1). Analogno, vrijedi

$$\forall \hat{v} \in H^m(\hat{K}), \quad |v|_{H^m(K)} \leq C\|B_K^{-1}\|^m |\det(B_K)|^{1/2} |\hat{v}|_{H^m(\hat{K})}. \quad (4.7)$$

U ocjenama iz Teorema 4.1 htjeli bismo izraziti norme matrica B_K i B_K^{-1} pomoću geometrijskih karakteristika skupova K i \hat{K} . Stoga uvodimo ove geometrijske parametre:

$$\begin{aligned} h_K &= \text{diam}(K), & \hat{h} &= \text{diam}(\hat{K}) \\ \rho_K &= \sup\{\text{diam}(S) : S \subset K \text{ je kugla sadržana u } K\}, \\ \hat{\rho} &= \sup\{\text{diam}(S) : S \subset \hat{K} \text{ je kugla sadržana u } \hat{K}\}. \end{aligned}$$

Lema 4.1 Neka su \hat{K} i K dva ograničena afino ekvivalentna skupa u \mathbb{R}^d i neka je pripadana afina bijekcija dana s (4.1). Tada je

$$\|B_K\| \leq \frac{h_K}{\hat{\rho}}, \quad \|B_K^{-1}\| \leq \frac{\hat{h}}{\rho_K}. \quad (4.8)$$

Dokaz. Po definiciji operatorske norme matrice imamo

$$\|B_K\| = \sup_{\|\hat{\mathbf{x}}\|=1} \|B_K \hat{\mathbf{x}}\| = \frac{1}{\hat{\rho}} \sup_{\|\hat{\mathbf{x}}\|=\hat{\rho}} \|B_K \hat{\mathbf{x}}\|.$$

Za svaki vektor $\hat{\mathbf{x}}$ koji zadovoljava $\|\hat{\mathbf{x}}\| = \hat{\rho}$ možemo naći dvije točke $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2 \in \hat{K}$ takve da je $\hat{\mathbf{x}}_2 - \hat{\mathbf{x}}_1 = \hat{\mathbf{x}}$. Kako je $B_K \hat{\mathbf{x}} = F_K(\hat{\mathbf{x}}_2) - F_K(\hat{\mathbf{x}}_1)$ vidimo da je

$$\|B_K\| \leq \frac{1}{\hat{\rho}} \sup_{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2 \in \hat{K}} \|F_K(\hat{\mathbf{x}}_2) - F_K(\hat{\mathbf{x}}_1)\| \leq \frac{h_K}{\hat{\rho}}.$$

Druga nejednakost slijedi iz prve zamjenom uloga skupova K i \hat{K} . \square

Sljedeći apstraktan rezultat koristimo na referentnom elementu \hat{K} i primijenjujemo ga na operator $\mathbb{I} - \Pi_{\hat{K}}$, gdje je $\Pi_{\hat{K}}$ interpolacijski operator.

Propozicija 4.1 (Bramble-Hilbertova lema) Neka su $m \geq 0$ i $s \geq 0$ te neka je $\hat{\lambda}: H^s(\hat{K}) \rightarrow H^m(\hat{K})$ linearan i neprekidan operator sa svojstvom

$$\forall \hat{p} \in \mathbb{P}_l, \quad \hat{\lambda}(\hat{p}) = 0, \quad (4.9)$$

gdje je $l \geq 0$. Tada za svako $\hat{v} \in H^s(\hat{K})$ vrijedi

$$|\hat{\lambda}(\hat{v})|_{H^m(\hat{K})} \leq \|\hat{\lambda}\|_{\mathcal{L}(H^s(\hat{K}), H^m(\hat{K}))} \inf_{\hat{p} \in \mathbb{P}_l} \|\hat{v} + \hat{p}\|_{H^s(\hat{K})}. \quad (4.10)$$

Dokaz. Koristeći neprekidnost operatora $\hat{\lambda}$ i (4.9) imamo za svako $\hat{p} \in \mathbb{P}_l$

$$|\hat{\lambda}(\hat{v})|_{H^m(\hat{K})} = |\hat{\lambda}(\hat{v} + \hat{p})|_{H^m(\hat{K})} \leq \|\hat{\lambda}\|_{\mathcal{L}(H^s(\hat{K}), H^m(\hat{K}))} \|\hat{v} + \hat{p}\|_{H^s(\hat{K})}.$$

Zbog proizvoljnosti polinoma $\hat{p} \in \mathbb{P}_l$ slijedi rezultat. \square

Uočimo da se na desnoj strani u (4.10) javlja izraz $\inf \|\hat{v} + \hat{p}\|_{H^s(\hat{K})}$ koji je norma na kvocijentalnom prostoru $H^s(\hat{K})/\mathbb{P}_l$.

Željeli bismo primijeniti Bramble-Hilbertovu lemu na linearan operator $\hat{\lambda} = \mathbb{I} - \Pi_{\hat{K}}$, pri čemu, radi određenosti, promatramo prostor \mathbb{P}_k elemenata. Tada je svojstvo (4.9) ispunjeno za $l = k$ te nam ostaje odabrati indekse s i m takve da operator $\hat{\lambda}$ bude neprekidan kao operator s $H^s(\hat{K})$ u $H^m(\hat{K})$. Tu se susrećemo sa sljedećim problemom: Interpolacijski operator je definiran na neprekidnim funkcijama budući da uzima vrijednosti funkcije u nodalnim točkama. Stoga polazni prostor $H^s(\hat{K})$ mora zadovoljavati $H^s(\hat{K}) \subset C(\hat{K})$, a to nije ispunjeno za $s = 1$, osim u jednoj prostornoj dimenziji. Dakle u dimenzijama $d = 2$ i 3 , koje nas najviše zanimaju (fizički prostor dimenzije veće od 3 nećemo promatrati premda to ima smisla u nekim situacijama) ne možemo uzeti $s = 1$. Prema teoremu ulaganja, moramo uzeti minimalno $s = 2$ kako bi imali $H^s(\hat{K}) \subset C(\hat{K})$. Dolazni prostor, $H^m(\hat{K})$, ne može imati indeks veći od s jer tada ne bismo imali neprekidnost operatora. Da bismo pokazali neprekidnost operatora $\hat{\lambda} = \mathbb{I} - \Pi_{\hat{K}}: H^s(\hat{K}) \rightarrow H^m(\hat{K})$ ($m \leq s$) dovoljno je pokazati da je operator $\Pi_{\hat{K}}: H^s(\hat{K}) \rightarrow H^m(\hat{K})$ neprekidan. To vidimo uzimajući u obzir da interpolacijski operator djeluje po formuli

$$(\Pi_{\hat{K}}\hat{v})(\hat{\mathbf{x}}) = \sum_{i=1}^N \hat{v}(\hat{a}_i)\hat{p}(\hat{\mathbf{x}}),$$

odakle slijedi

$$\|\Pi_{\hat{K}}\hat{v}\|_{H^m(\hat{K})} \leq \sum_{i=1}^N |\hat{v}(\hat{a}_i)| \|\hat{p}\|_{H^m(\hat{K})} \leq C_1 \|\hat{v}\|_{C(\hat{K})} \leq C_2 \|\hat{v}\|_{H^s(\hat{K})},$$

gdje smo u zadnjem koraku iskoristili neprekidnost ulaganja $H^s(\hat{K}) \subset C(\hat{K})$, odnosno $s \geq 2$. Time dolazimo do sljedećeg rezultata:

Propozicija 4.2 Neka je $(\hat{K}, P_{\hat{K}}, \Sigma_{\hat{K}})$ simpleks tipa k , $k \geq 0$, te neka je $s \geq 2$ i $0 \leq m \leq s$. Tada za svako $\hat{v} \in H^s(\hat{K})$ vrijedi

$$|\hat{v} - \Pi_{\hat{K}}\hat{v}|_{H^m(\hat{K})} \leq \|\mathbb{I} - \Pi_{\hat{K}}\|_{\mathcal{L}(H^s(\hat{K}), H^m(\hat{K}))} \inf_{\hat{p} \in \mathbb{P}_k} \|\hat{v} + \hat{p}\|_{H^s(\hat{K})}. \quad (4.11)$$

Sljedeći rezultat nam omogućuje da ocijenimo kvocijentalnu normu odgovarajućom polunormom.

Propozicija 4.3 (Deny-Lionsova lema) Za svako $k \geq 0$ postoji konstanta $C = C(d, k, \hat{K})$ takva da je

$$\forall \hat{v} \in H^{k+1}(\hat{K}), \quad \inf_{\hat{p} \in \mathbb{P}_k} \|\hat{v} + \hat{p}\|_{H^{k+1}(\hat{K})} \leq C |\hat{v}|_{H^{k+1}(\hat{K})}. \quad (4.12)$$

Dokaz Deny-Lionsove leme ćemo dati u dodatku budući da je u njemu potrebno koristiti kompaktnost ulaganja Soboljevljevih prostora.

Sada smo u poziciji dokazati osnovni rezultat ove točke.

Teorem 4.2 Neka je (K, P_K, Σ_K) simpleks tipa k , $k \geq 1$, i neka je $0 \leq m \leq l + 1$, gdje je $l = \min(k, s - 1) \geq 1$. Tada postoji konstanta $C = C(\hat{K}, \Pi_{\hat{K}}, d, m, s)$ takva da vrijedi

$$\forall v \in H^s(K), \quad |v - \Pi_K v|_{H^m(K)} \leq C \frac{h_K^{l+1}}{\rho_K^m} |v|_{H^{l+1}(K)}. \quad (4.13)$$

Dokaz. Uočimo da je prema našim pretpostavkama $s \geq 2$ i $m \leq s$. Stoga za $v \in H^s(K) \subset H^m(K)$ imamo prema (4.7) i (4.8)

$$|v - \Pi_K v|_{H^m(K)} \leq C \frac{1}{\rho_K^m} |\det(B_K)|^{1/2} |\hat{v} - \Pi_{\hat{K}} \hat{v}|_{H^m(\hat{K})}, \quad (4.14)$$

gdje smo iskoristili činjenicu da je $(\Pi_K v) \circ F_K = \Pi_{\hat{K}} \hat{v}$. (Sve veličine koje ovise o referentnom elementu uključujemo u konstantu.) Prema Propoziciji 4.2 imamo

$$|\hat{v} - \Pi_{\hat{K}} \hat{v}|_{H^m(\hat{K})} \leq C \inf_{\hat{p} \in \mathbb{P}_k} \|\hat{v} + \hat{p}\|_{H^s(\hat{K})}. \quad (4.15)$$

Pogledajmo sada dva slučaja.

1) $s - 1 < k$. U tom slučaju je $s = l + 1$ i $l < k$. Ocjenu (4.15) sada možemo iskoristiti tako da u njoj smanjimo skup po kome se uzima infimum i primijenimo Deny-Lionsovu lemu s $k = l$:

$$|\hat{v} - \Pi_{\hat{K}} \hat{v}|_{H^m(\hat{K})} \leq C \inf_{\hat{p} \in \mathbb{P}_k} \|\hat{v} + \hat{p}\|_{H^{l+1}(\hat{K})} \leq C \inf_{\hat{p} \in \mathbb{P}_l} \|\hat{v} + \hat{p}\|_{H^{l+1}(\hat{K})} \leq C_1 |\hat{v}|_{H^{l+1}(\hat{K})}.$$

2) $s - 1 \geq k$. U tom slučaju je $l = k$ i $s \geq k + 1$. U tom smo slučaju Propoziciju 4.2 mogli primijeniti s indeksom $k + 1$ umjesto indeksa s i dobiti

$$|\hat{v} - \Pi_{\hat{K}} \hat{v}|_{H^m(\hat{K})} \leq C \inf_{\hat{p} \in \mathbb{P}_k} \|\hat{v} + \hat{p}\|_{H^{k+1}(\hat{K})}. \quad (4.16)$$

umjesto (4.15). Time smo ponovo u poziciji da primijenimo Deny-Lionsovu lemu i dobivamo (uoči da je sada $k = l$)

$$|\hat{v} - \Pi_{\hat{K}} \hat{v}|_{H^m(\hat{K})} \leq C |\hat{v}|_{H^{l+1}(\hat{K})}. \quad (4.17)$$

Dakle, u oba slučaja dobivamo (4.17). Primjenom (4.6) (s $m = l + 1$) i (4.8) dobivamo

$$|\hat{v}|_{H^{l+1}(\hat{K})} \leq C h_K^{l+1} |\det(B_K)|^{-1/2} |v|_{H^{l+1}(K)}. \quad (4.18)$$

Kombiniranjem (4.14), (4.17) i (4.18) dobivamo

$$|v - \Pi_K v|_{H^m(K)} \leq C \frac{h_K^{l+1}}{\rho_K^m} |\det(B_K)|^{1/2} |\det(B_K)|^{-1/2} |v|_{H^{l+1}(K)} = C \frac{h_K^{l+1}}{\rho_K^m} |v|_{H^{l+1}(K)}.$$

Time je teorem dokazan. \square

Napomena 4.1 Uočimo da u slučaju $k > s - 1$ dobivamo istu ocjenu greške interpolacije kao i za $k = s - 1$. Stoga za danu glatkoću funkcije s ne treba uzimati stupanj elementa k veći od $s - 1$; vrijednost $k = s - 1$ je optimalna.

Napomena 4.2 Teorem 4.2 ćemo najčešće koristiti u slučaju $s = k + 1 \geq 2$ i $m = 0, 1$. Tada imamo

$$\|v - \Pi_K v\|_{L^2(K)} \leq Ch_K^{k+1} |v|_{H^{k+1}(K)}, \quad (4.19)$$

$$|v - \Pi_K v|_{H^1(K)} \leq C \frac{h_K^{k+1}}{\rho_K} |v|_{H^{k+1}(K)}. \quad (4.20)$$

Teorem 4.2 pokazuje da ako na nizu sve finijih triangulacija domene želimo postići optimalnu konvergenciju interpolacijskog procesa, trebamo postaviti dodatno ograničenje na konstrukciju triangulacije koje se svodi na to da elementi triangulacije *ne degeneriraju*.

Pri proučavanju konvergencije metode konačnih elemenata uvijek uzimamo da je zadan niz triangulacija (\mathcal{T}_{h_i}) , $i \in \mathbb{N}$, sa svojstvom $\lim h_i = 0$. Pri tome jednostavno govorimo o familiji triangulacija $(\mathcal{T}_h)_{h>0}$.

Imamo sljedeću definiciju:

Definicija 4.1 Familija triangulacija $(\mathcal{T}_h)_{h>0}$ je regularna ako postoji broj $\sigma > 0$ takav da je

$$\forall h > 0, \quad \max_{K \in \mathcal{T}_h} \frac{h_K}{\rho_K} \leq \sigma.$$

U regularnoj familiji triangulacija ocjena (4.13) može se zapisati u obliku

$$\forall v \in H^s(K), \quad |v - \Pi_K v|_{H^m(K)} \leq Ch_K^{l+1-m} |v|_{H^{l+1}(K)}. \quad (4.21)$$

U slučaju triangulacije trokutima, regularnost znači da prilikom profinjenja triangulacije minimalni kut u svakom pojedinom trokutu ostaje ograničen odozdo. Drugim riječima, nigdje se neće desiti da trokut pri profinjenju degenerira u segment.

4.2 Globalne ocjene

Globalne ocjene greške interpolacije se lako izvode iz lokalnih. Pri tome moramo pretpostaviti, kao što smo i dosad činili, da je domena poliedarska.

Definirajmo

$$h = \max\{h_K : K \in \mathcal{T}_h\}.$$

Sumiranjem po svim elementima triangulacije, koristeći $(\Pi_h v)|_K = \Pi_K v$, za svako $K \in \mathcal{T}_h$, te $h_K \leq h$, dobivamo

$$\begin{aligned} \|v - \Pi_h v\|_{L^2(\Omega)} &= \left(\sum_{K \in \mathcal{T}_h} \|v - \Pi_K v\|_{L^2(K)}^2 \right)^{1/2} \leq Ch^{k+1} \left(\sum_{K \in \mathcal{T}_h} |v|_{H^{k+1}(K)}^2 \right)^{1/2} \\ &= Ch^{k+1} |v|_{H^{k+1}(\Omega)} \end{aligned}$$

i isto tako

$$\begin{aligned} |v - \Pi_h v|_{H^1(\Omega)} &= \left(\sum_{K \in \mathcal{T}_h} |v - \Pi_K v|_{H^1(K)}^2 \right)^{1/2} \leq Ch^k \left(\sum_{K \in \mathcal{T}_h} |v|_{H^{k+1}(K)}^2 \right)^{1/2} \\ &= Ch^k |v|_{H^{k+1}(\Omega)}. \end{aligned}$$

Time smo dokazali:

Teorem 4.3 Neka je Ω ograničena poliedarska domena na kojoj je zadana regularna familija triangulacija $(\mathcal{T}_h)_{h>0}$ na simpleksima tipa k , $k \geq 1$. Tada postoji konstanta C , neovisna o h , takva da za svako $v \in H^{k+1}(\Omega)$ vrijedi

$$\|v - \Pi_h v\|_{L^2(\Omega)} \leq Ch^{k+1} |v|_{H^{k+1}(\Omega)}, \quad (4.22)$$

$$\|v - \Pi_h v\|_{H^1(\Omega)} \leq Ch^k |v|_{H^{k+1}(\Omega)}, \quad (4.23)$$

$$(4.24)$$

gdje je Π_h V_h -interpolacijski operator.

Vratimo se sada varijacijskoj zadaći (P)

$$\begin{cases} \text{Naći } u \in V \\ \forall v \in V, \quad a(u, v) = \langle F, v \rangle \end{cases}$$

i njenoj varijacijskoj aproksimaciji $(P)_h$

$$\begin{cases} \text{Naći } u_h \in V_h \\ \forall v \in V_h, \quad a(u_h, v) = \langle F, v \rangle \end{cases}$$

gdje je V_h prostor konačnih elemenata konstruiran od simpleksa tipa k , $k \geq 1$, na regularnoj triangulaciji. Prema Céainoj lemi imamo

$$\|u - u_h\|_{H^1(\Omega)} \leq C \inf_{v \in V_h} \|u - v\|_{H^1(\Omega)} \leq \|u - \Pi_h u\|_{H^1(\Omega)},$$

gdje smo uvrstili $v = \Pi_h u$. Ako je egzaktno rješenje dovoljno regularno možemo primijeniti prethodni teorem. Time dobivamo:

Teorem 4.4 Neka su ispunjene pretpostavke Teorema 4.3 i pretpostavimo da je $u \in V \cap H^{k+1}(\Omega)$ rješenje eliptičke varijacijske zadaće (P) , koja zadovoljava uvjete Lax-Milgramove leme. Tada postoji konstanta C , neovisna o h , takva da je

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch^k |u|_{H^{k+1}(\Omega)}, \quad (4.25)$$

gdje je $u_h \in V_h$ rješenje diskretne zadaće (P_h) .

Prethodni teorem daje ocjenu greške metode uz uvjet da je egzaktno rješenje dovoljno glatko. Na primjer, triangulacija sastavljena od simpleksa tipa (1) dat će linearno konvergentnu metodu ako je egzaktno rješenje klase $H^2(\Omega)$. Ako primijenimo simplekse tipa 2, dobit ćemo kvadratičnu konvergenciju, ali samo ako je točno rješenje iz $H^3(\Omega)$.

4.3 Aubin-Nitscheova lema

Teorem 4.4 daje red konvergencije metode u H^1 -normi. Sada nam je cilj pokazati da uz dodatnu regularnost u L^2 -normi imamo za jedan viši red konvergencije.

Teorem 4.5 (Aubin-Nitscheova lema) Neka je $u \in V$ rješenje zadaće (P) , a $u_h \in V_h$ rješenje zadaće (P_h) te neka su ispunjeni uvjeti Lax-Milgramove leme. Tada je

$$\|u - u_h\|_{L^2(\Omega)} \leq M \|u - u_h\|_{H^1(\Omega)} \left(\sup_{g \in L^2(\Omega)} \left\{ \frac{1}{\|g\|_{L^2(\Omega)}} \inf_{\phi \in V_h} \|\phi_g - \phi\|_{H^1(\Omega)} \right\} \right), \quad (4.26)$$

gdje smo za $g \in L^2(\Omega)$ sa $\phi_g \in V$ označili jedinstveno rješenje varijacijske zadaće

$$\forall v \in V, \quad a(v, \phi_g) = (g, v)_{L^2(\Omega)}. \quad (4.27)$$

Dokaz. Egzistencija i jedinstvenost rješenja $\phi_g \in V$, zadaće (4.27), dobiva se primjenom Lax-Milgramove leme. Prema karakterizaciji norme u $L^2(\Omega)$ možemo pisati

$$\|u - u_h\|_{L^2(\Omega)} = \sup_{g \in L^2(\Omega)} \frac{|(g, u - u_h)_{L^2(\Omega)}|}{\|g\|_{L^2(\Omega)}}. \quad (4.28)$$

Stvaljajući $u - u_h$ u (4.27) dobivamo

$$a(u - u_h, \phi_g) = (g, u - u_h)_{L^2(\Omega)},$$

pa zbog ortogonalnosti

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

dobivamo

$$\forall \phi \in V_h, \quad a(u - u_h, \phi_g - \phi) = (g, u - u_h)_{L^2(\Omega)},$$

i stoga

$$|(g, u - u_h)_{L^2(\Omega)}| \leq M \|u - u_h\|_{H^1(\Omega)} \inf_{\phi \in V_h} \|\phi_g - \phi\|_{H^1(\Omega)}.$$

Tvrđnja slijedi primjenom formule (4.28). \square

Zadaća (4.27) je adjungirana zadaća u odnosu na (P) . Razlika je u tome što su prva i druga varijabla u bilinearnoj formi zamijenile mjesta. Ukoliko je forma simetrična, onda se polazna i adjungirana zadaća podudaraju.

Kažemo da je zadaća (4.27) regularna, ako za svako $g \in L^2(\Omega)$ ima jedinstveno rješenje $\phi_g \in V \cap H^2(\Omega)$, te postoji konstanta C takva da je

$$\forall g \in L^2(\Omega), \quad \|\phi_g\|_{H^2(\Omega)} \leq C \|g\|_{L^2(\Omega)}. \quad (4.29)$$

Teorem 4.6 Neka su ispunjene pretpostavke Teorema 4.3 i Teorema 4.5 te neka je adjungirani problem (4.27) regularan. Tada postoji konstanta C , neovisna o h , takva da je

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^{k+1} |u|_{H^{k+1}(\Omega)}. \quad (4.30)$$

Dokaz. Primjenom Teorem 4.4 dobivamo

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch^k |u|_{H^{k+1}(\Omega)}. \quad (4.31)$$

Nadalje, ako primijenimo Teorem 4.3 na $v = \phi_g$, uz $k = 1$, dobivamo

$$\|\phi_g - \Pi_h \phi_g\|_{H^1(\Omega)} \leq Ch |\phi_g|_{H^2(\Omega)}$$

i stoga

$$\inf_{\phi \in V_h} \|\phi_g - \phi\|_{H^1(\Omega)} \leq \|\phi_g - \Pi_h \phi_g\|_{H^1(\Omega)} \leq Ch |\phi_g|_{H^2(\Omega)} \leq Ch \|g\|_{L^2(\Omega)},$$

gdje smo u zadnjem koraku iskoristili (4.29). Aubin-Nitscheova lema i (4.31) daju

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch \|u - u_h\|_{H^1(\Omega)} \leq Ch^{k+1} |u|_{H^{k+1}(\Omega)}. \quad \square$$

4.4 O regularnosti točnog rješenja

Vidjeli smo u prethodnim sekcijama da ocjena greške metode konačnih elemenata ovisi o regularnosti točnog rješenja. Simplicijalni elementi reda $k \geq 1$ daju optimalni red konvergencije jedino ako točno rješenje pripada prostoru H^{k+1} , a analogni rezultati vrijede i za druge tipove konačnih elemenata. Rješenje varijacijske zadaće se traži u prostoru H^1 i njegova pripadnost višim Soboljevlijevim prostorima ovisi o regularnosti podataka koji ulaze u formulaciju rubne zadaće. To su:

1. Glatkoća koeficijenata diferencijalne jednadžbe i rubnog uvjete, te glatkoća funkcija koje se nalaze na desnoj strani jednadžbe i rubnog uvjeta.
2. Tip rubnog uvjeta.
3. Oblik domene i posebno glatkoća granice domene.

Ilustrirat ćemo rezultate teorije regularnosti rješenja eliptičnih rubnih zadaća drugog reda na Dirichletovoj zadaći oblika:

$$-\sum_{i,j=1}^d \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) + \sum_{i=1}^d b_i \frac{\partial u}{\partial x_i} + a_0 u = f \quad \text{u } \Omega \quad (4.32)$$

$$u = u_0 \quad \text{na } \partial\Omega \quad (4.33)$$

za koju pretpostavljamo da zadovoljava uvjete (2.36)–(2.39). Tada vrijedi sljedeći rezultat (vidi [2], sekcija 8.4, Teorem 8.12.):

Teorem 4.7 Neka koeficijenti zadaće (4.32), (4.33) zadovoljavaju (2.36)–(2.39). Pretpostavimo još da su funkcije a_{ij} Lipschitzove na $\bar{\Omega}$, da rubni uvjet ima proširenje na čitavu domenu iz prostora $H^2(\Omega)$, te da je Ω ograničena domena klase C^2 . Tada je slabo rješenje u zadaće (4.32), (4.33) iz $H^2(\Omega)$ i vrijedi ocjena

$$\|u\|_{H^2(\Omega)} \leq C(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} + \|u_0\|_{H^2(\Omega)})$$

gdje konstanta C ovisi samo o domeni, konstanti eliptičnosti i normama koeficijenata.

Ukoliko imamo veću glatkoću podataka dobit ćemo još veću glatkoću rješenja. U tom smislu gornji se teorem može *iterirati*. Navedimo sljedeći rezultat (Teorem 8.13 u [2]).

Teorem 4.8 Neka koeficijenti zadaće (4.32), (4.33) zadovoljavaju (2.36)–(2.39). Pretpostavimo još da su k -te derivacije funkcije a_{ij} Lipschitzove na $\bar{\Omega}$, da su $k-1$ -ve derivacije funkcija b_i i a_0 Lipschitzove na $\bar{\Omega}$, $u_0 \in H^{k+2}(\Omega)$, $f \in H^k(\Omega)$, te da je Ω ograničena domena klase C^{k+2} . Tada je slabo rješenje u zadaće (4.32), (4.33) iz $H^{k+2}(\Omega)$ i vrijedi ocjena

$$\|u\|_{H^{k+2}(\Omega)} \leq C(\|u\|_{L^2(\Omega)} + \|f\|_{H^k(\Omega)} + \|u_0\|_{H^{k+2}(\Omega)})$$

gdje konstanta C ovisi samo o domeni, konstanti eliptičnosti i normama koeficijenata.

Ovi rezultati kažu da će rješenje eliptičke (Dirichletove) rubne zadaće biti po volji glatko, ako su svi ulazni podaci dovoljno glatki. Analogan teorem vrijedi za čistu Neumannovu zadaću, ali ne vrijedi nužno za mješovitu Dirichlet-Neumannovu zadaću

Za domenu Ω kažemo da je klase C^k ukoliko se u okolini svake točke $\mathbf{x} \in \partial\Omega$ granica $\partial\Omega$ može lokalno prikazati kao graf funkcije klase C^k . Iz te definicije je jasno da poligonalna domena u \mathbb{R}^2 (poliedarska u \mathbb{R}^3) nije čak niti klase C^1 . To predstavlja poteškoću pri primjeni gornjih rezultata

regularnosti na metodu konačnih elemenata. U slučaju poligonalne domene pokazuje se da verzija Teorema 4.7 vrijedi ako je domena konveksna (vidi [3]). Ilustrirajmo gubitak regularnosti s dva primjera (vidi [1]).

Primjer 1. Kao prvi primjer pogledajmo domenu $\Omega \subset \mathbb{R}^2$ koja predstavlja sektor jedinične kružnice. U polarnim koordinatama imamo

$$\Omega = \{(r, \theta) : 0 < r < 1, 0 < \theta < \pi/\beta\}, \quad 1/2 < \beta < 1.$$

Funkcija

$$v(r, \theta) = \Im(z^\beta) = r^\beta \sin(\beta\theta),$$

je harmonijska kao imaginarni dio analitičke funkcije. Sada se lako provjerava da funkcije

$$u(r, \theta) = (1 - r^2)v(r, \theta)$$

zadovoljava homogeni Dirichletov rubni uvjet na $\partial\Omega$. Nadalje, po formuli

$$\Delta u = (1 - r^2)\Delta v + 2\nabla(1 - r^2) \cdot \nabla v + v\Delta(1 - r^2) = -4r \frac{\partial v}{\partial r} - 4v = -4r^\beta \sin(\beta\theta).$$

gdje smo iskoristili izraze za diferencijalne operatore u polarnom sustavu:

$$\nabla v = \frac{\partial v}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial v}{\partial \theta} \mathbf{e}_\theta, \quad \Delta v = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial v}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2}.$$

Time smo došli do rješenja zadaće

$$\begin{aligned} -\Delta u &= f & \text{u } \Omega, \\ u &= 0 & \text{na } \partial\Omega, \end{aligned}$$

s $f = 4r^\beta \sin(\beta\theta)$ koja je ograničena te evidentno pripada prostoru $L^2(\Omega)$. Na isti način i rješenje pripada prostoru $L^2(\Omega)$. Kvadrat gradijenta rješenja je jednak

$$|\nabla u|^2 = [-(2 + \beta)r^{\beta+1} + \beta r^{\beta-1}]^2 \sin^2(\beta\theta) + \beta^2(1 - r^2)^2 r^{2\beta-2} \cos^2(\beta\theta).$$

Stoga je

$$\int_{\Omega} |\nabla u(\mathbf{x})|^2 d\mathbf{x} = \int_0^{\pi/\beta} \int_0^1 |\nabla u(\mathbf{x})|^2 r dr d\theta < +\infty$$

jer je integral od $r^{2\beta-1}$ konačan. Imamo, dakle, $u \in H_0^1(\Omega)$. Da bismo pokazali da ova funkcija ne leži u $H^2(\Omega)$ dovoljno je pokazati da druga derivacija po r nije u $L^2(\Omega)$:

$$\frac{\partial^2 u}{\partial r^2} = [-(2 + \beta)(\beta + 1)r^\beta + \beta(\beta - 1)r^{\beta-2}] \sin(\beta\theta).$$

Imamo

$$\int_0^1 (r^{\beta-2})^2 r dr = \int_0^1 r^{2\beta-3} dr = \frac{1}{2\beta-2} r^{2\beta-2} \Big|_0^1 = +\infty.$$

Time je dokazano $u \notin H^2(\Omega)$. Uočimo da su u ovom primjeru svi podaci dovoljno glatki osim domene. Štoviše, regularnost nije izgubljena toliko zbog neregularnosti domene već zbog toga što ona nije konveksna. Za $\beta > 1$ domena postaje konveksna i rješenje postaje klase $H^2(\Omega)$.

Primjer 2. Drugi primjer pokazuje da mješoviti rubni uvjet može pokvariti regularnost rješenja. Domena $\Omega \subset \mathbb{R}^2$ ponovo je zadana u polarnim koordinatama

$$\Omega = \{(r, \theta) : 0 < r < 1, 0 < \theta < \pi\}.$$

Granica je podijeljena u dva disjunktna podskupa: $\partial\Omega = \Gamma_1 \cup \Gamma_2$, gdje je $\Gamma_1 = [-1, 0] \times \{0\}$ i $\Gamma_2 = \partial\Omega \setminus \Gamma_1$. Promatramo zadaću

$$\begin{aligned} -\Delta u &= f && \text{u } \Omega, \\ \frac{\partial u}{\partial \mathbf{n}} &= 0 && \text{na } \Gamma_1 \\ u &= 0 && \text{na } \Gamma_2. \end{aligned}$$

Rješenje tražimo u obliku

$$u(r, \theta) = (1 - r^2)r^{1/2} \sin(\theta/2).$$

Dirichletov rubni uvjet na Γ_2 je evidentno zadovoljen. Jednostavnim računom dobivamo

$$\nabla u = \left(\frac{1}{2}r^{-1/2} - \frac{5}{2}r^{3/2}\right) \sin(\theta/2) \mathbf{e}_r + \frac{1-r^2}{2r^{1/2}} \cos(\theta/2) \mathbf{e}_\theta.$$

Sada se lako vidi da je $u \in H^1(\Omega)$ i da je zadovoljen rubni uvjet na Γ_1 . Osim toga jednostavnim računom izlazi $f = 6\sqrt{r} \sin(\theta/2)$, što je ograničena funkcija i stoga iz $L^2(\Omega)$. Budući da je

$$\frac{\partial^2 u}{\partial r^2} \sim r^{-3/2}$$

ponovo dobivamo

$$\int_0^1 \left(\frac{\partial^2 u}{\partial r^2}\right)^2 r dr = +\infty,$$

i stoga $u \notin H^2(\Omega)$. U ovom slučaju gubitak glatkoće je posljedica kombinacije Neumannovog i Dirichletovog rubnog uvjeta.

4.5 Adaptivne metode

Idealni numerički postupak za aproksimaciju rješenja varijacijske zadaće daje aproksimativno rješenje koje u odabranoj normi ima grešku manju od zadane tolerancije i pri tome koristi optimalan broj stupnjeva slobode. Tome se idealu približujemo pomoću adaptivnih metoda koje prilagođavaju triangulaciju domene karakteru rješenja kako bi omogućile dobru aproksimaciju na relativno gruboj mreži. Tamo gdje je greška aproksimacije veća, mreža će biti finija, i obratno.

Osnovni sastojak svake adaptivne metode je *procjenitelj greške*. To je svaka veličina η koja zadovoljava

$$\|u - u^h\|_{H^1(\Omega)} \leq \eta$$

pa će stoga norma greške biti manja od tolerancije ε ako je zadovoljeno $\eta < \varepsilon$. Procjenitelj greške pri tome mora biti lokalnog karaktera, tj. za svaki element triangulacije T mora postojati lokalni procjenitelj η_T , takav da je

$$\|u - u^h\|_{H^1(T)} \leq \eta_T.$$

Tada je evidentno

$$\eta = \left(\sum_{T \in \mathcal{T}_h} \eta_T^2 \right)^{1/2}.$$

Sada možemo primijeniti, na primjer, *algoritam jednake razdiobe greške*. Ako trebamo postići $\eta < \varepsilon$, onda zahtijevamo da bude ispunjeno $\eta_T < \varepsilon/\sqrt{N}$, gdje je N broj elemenata triangulacije. Sve one elemente na kojima to nije zadovoljeno profinimo, na primjer razbijanjem svakog trokuta na četiri nova trokuta. Takva strategija cilja na jednoliku raspodjelu greške u cijeloj domeni. Uočimo da za postizanje zadane tolerancije na profinjenoj mreži, N treba biti broj elemenata *profinjene* mreže no on nam unaprijed nije poznat.

Adaptivni algoritam imao bi stoga ovu strukturu:

```
do{
  Izračunaj aproksimaciju na zadanoj mreži
  Ispitaj da li je procjenitelj greške dovoljno mali na svakom elementu
  Ako je izađi iz petlje
  Ako nije označi elemente koje treba profiniti
  Konstriraj novu mrežu
}
while(greška > tolerancije)
```

Do jednog procjenitelja greške dolazimo na osnovu ocjene greške interpolacije. Naime, ocjenu iz Teorema 4.4, u slučaju $k = 1$, možemo preciznije zapisati kao:

$$\|u - u_h\|_{H^1(\Omega)} \leq C \left(\sum_{T \in \mathcal{T}_h} h_T^2 |u|_{H^2(T)}^2 \right)^{1/2}. \quad (4.34)$$

Odavde se kao lokalni procjenitelj greške nameće

$$\eta_T = h_T |u|_{H^2(T)},$$

te je onda

$$\eta = C \left(\sum_{T \in \mathcal{T}_h} \eta_T^2 \right)^{1/2}.$$

Prema ovom procjenitelju radimo subdiviziju onih elemenata na kojima je H^2 -norma točnog rješenja suviše velika.

Ovakav izbor procjenitelja ima dva nedostatka. Prvi, manji, je u tome što nam konstanta C , premda ne ovisi o h , nije poznata. Drugi, ozbiljniji, je taj što nam točno rješenje u nije poznato.

Jedan evidentan način da se izbjegne ova poteškoća je da se norma $|u|_{H^2(T)}$ aproksimira korištenjem približnog rješenja u_h i konačnih diferencija za aproksimaciju druge derivacije. Tada dobivamo procjenitelj koji je *izračunljiv*.

U nekim situacijama, kao što su singulariteti u šiljcima domene, mi često znamo asimptotsko ponašanje druge derivacije točnog rješenja kada se približavamo šiljku. Funkcija koja opisuje to ponašanje može poslužiti umjesto točnog rješenja jer, premda nećemo dobiti ocjenu greške, dobit ćemo profinjenje mreže tamo gdje su druge derivacije rješenja najveće.

Napomenimo konačno da se svaki procjenitelj greške koji se može izračunati iz aproksimativnog rješenja u_h naziva *aposteriorni procjenitelj*. Postoji više načina kako da se oni konstruiraju o čemu ćemo govoriti u Poglavlju 7.

4.6 Zadaci

1. *Dokažite sljedeću generalizaciju Teorema 4.1:*

Teorem 4.9 Neka su K i \hat{K} ograničeni afino ekvivalentni skupovi u \mathbb{R}^d . Ako je $v \in W^{m,p}(K)$ za neki $m \geq 1$ i $p \in [1, \infty]$, onda je $\hat{v} = v \circ F \in W^{m,p}(\hat{K})$ te postoji konstanta $C = C(m, d)$ takva da je

$$\forall v \in W^{m,p}(K), \quad |\hat{v}|_{m,p;\hat{K}} \leq C \|B\|^m |\det(B)|^{-1/p} |v|_{m,p;K}, \quad (4.35)$$

gdje je B matrica afinog preslikavanja F iz (4.1).

Analogno, vrijedi

$$\forall \hat{v} \in W^{m,p}(\hat{K}), \quad |v|_{m,p;K} \leq C \|B^{-1}\|^m |\det(B)|^{1/p} |\hat{v}|_{m,p;\hat{K}}. \quad (4.36)$$

Uputa: Slučaj $p = +\infty$ svodi se na slučaj $p < +\infty$ korištenjem Lema A.4 i A.5.

2. *Iskažite i dokažite Bramble-Hilbertovu lemu za linearan i neprekidan operator $\hat{\lambda}: W^{k+1,p}(\hat{\Omega}) \rightarrow W^{m,q}(\hat{\Omega})$ koji se poništava na \mathbb{P}_l .*

3. *Dokažite sljedeću verziju Deny-Lionsove leme:*

Teorem 4.10 Za svako $k \geq 0$ postoji konstanta $C = C(d, k, p, \hat{K})$ takva da je

$$\forall \hat{v} \in W^{k+1,p}(\hat{K}), \quad \inf_{\hat{p} \in \mathbb{P}_k} \|\hat{v} + \hat{p}\|_{W^{k+1,p}(\hat{K})} \leq C |\hat{v}|_{W^{k+1,p}(\hat{K})}, \quad (4.37)$$

4. Dokažite sljedeću generaliziranu verziju ocjene lokalne greške interpolacije.

Teorem 4.11 Za odabrane cijele brojeve $k \geq 0$ i $m \geq 0$ te $p, q \in [1, \infty]$, neka su $W^{k+1,p}(\hat{K})$ i $W^{m,q}(\hat{K})$ dva prostora Soboljeva koji zadovoljavaju

$$W^{k+1,p}(\hat{K}) \subset W^{m,q}(\hat{K}), \quad (4.38)$$

i neka je $\hat{\Pi} \in \mathcal{L}(W^{k+1,p}(\hat{K}), W^{m,q}(\hat{K}))$ linearno i neprekidno preslikavanje sa svojstvom

$$\forall \hat{p} \in \mathbb{P}_k, \quad \hat{\Pi} \hat{p} = \hat{p}. \quad (4.39)$$

Za ograničeni otvoreni skup K koji je afino-ekvivalentan s \hat{K} definiramo preslikavanje Π_K formulom

$$(\Pi_K v) \circ F = \hat{\Pi}(v \circ F), \quad (4.40)$$

za sve $v \in W^{k+1,p}(K)$, gdje je $F: \hat{K} \rightarrow K$ pripadno afino preslikavanje. Tada postoji konstanta $C = C(\hat{\Pi}, \hat{K})$ takva da¹

$$\forall v \in W^{k+1,p}(K), \quad |v - \Pi_K v|_{W^{m,q}(K)} \leq C |K|^{(1/q)-(1/p)} \frac{h_K^{k+1}}{\rho_K^m} |v|_{W^{k+1,p}(K)}. \quad (4.41)$$

Primijenite Teorem 4.11 na simplicijalne elemente i izvedite globalnu ocjenu greške u slučaju regularne familije triangulacija.

4.7 Dodatak 1. Dokaz Deny-Lionsove leme

Ovdje donosimo dokaz Deny-Lionsov leme koji je izostavljen u četvrtom poglavlju.

Pretpostavljat ćemo da je Ω ograničena domena s Lipschitzovom granicom, tako da vrijede teoremi o kompaktnim ulaganjima. U primjenama ćemo uzimati da je $\Omega = \text{Int}(K)$, gdje je K konačan element i u tom ćemo slučaju umjesto $W^{k,p}(\text{Int}(K))$ pisati jednostavnije $W^{k,p}(K)$.

Svaki linearan prostor $E \subset W^{m,p}(\Omega)$ omogućava da se u $W^{m,p}(\Omega)$ uvede relacija ekvivalencije

$$\forall v_1, v_2 \in W^{m,p}(\Omega), \quad v_1 \mathcal{R} v_2 \Leftrightarrow v_1 - v_2 \in E.$$

¹S $|K|$ označavamo Lebesgueovu mjeru skupa K .

Pomoću te relacije uvodi se kvocijentni skup $W^{m,p}(\Omega)/E$ čiji su elementi klase ekvivalencije, koje ćemo označavati s \dot{v} . Tako je

$$\dot{v} = \{w \in W^{m,p}(\Omega) : w - v \in E\},$$

klasa svih elemenata ekvivalentnih s $v \in W^{m,p}(\Omega)$. Ako je potprostor E zatvoren, onda je kvocijentni prostor $W^{m,p}(\Omega)/E$ Banachov s normom

$$\|\dot{v}\|_{W^{m,p}(\Omega)/E} = \inf_{w \in \dot{v}} \|w\|_{m,p;\Omega}$$

(vidi [4]).

Nas će zanimati kvocijentni prostor $W^{k+1,p}(\Omega)/\mathbb{P}_k$. Na svakom elementu tog prostora dobro je definirana polunorma

$$|\dot{v}|_{k+1,p} = |v|_{k+1,p}, \quad \forall v \in \dot{v},$$

budući da za svako $q \in \mathbb{P}_k$ vrijedi $|q|_{k+1,p} = 0$. Preslikavanje $\dot{v} \mapsto |\dot{v}|_{k+1,p}$ očito ima svojstva polunorme i vrijedi

$$|\dot{v}|_{k+1,p} \leq \|\dot{v}\|_{W^{k+1,p}(\Omega)/\mathbb{P}_k}.$$

Sljedeći teorem tvrdi da se radi o normi ekvivalentnoj kvocijentnoj normi na $W^{k+1,p}(\Omega)/\mathbb{P}_k$.

Teorem 4.12 Neka je Ω ograničena Lipschitzova domena. Postoji konstanta $C = C(\Omega)$ takva da je

$$\forall v \in W^{k+1,p}(\Omega), \quad \inf_{p \in \mathbb{P}_k} \|v + p\|_{k+1,p;\Omega} \leq C|v|_{k+1,p;\Omega}, \quad (4.42)$$

Dokaz. Neka je $N = \dim \mathbb{P}_k$ i neka linearni funkcionali f_i , $i = 1, 2, \dots, N$ čine bazu u dualu prostora \mathbb{P}_k . Koristeći Hahn-Banachov teorem (vidi [4]) te funkcionale možemo proširiti do linearnih i neprekidnih funkcionala na $W^{k+1,p}(\Omega)$, i ta ćemo proširenja ponovo označavati s f_i , $i = 1, 2, \dots, N$. (Ukoliko se želi izbjeći upotreba Hahn-Banachovog teorema ovi se funkcionali mogu konstruirati eksplicitno, na primjer, uzimajući funkcionale $v \mapsto \int_{\Omega} \partial^{\alpha} v(x) dx$.)

Uočimo da za funkcije $q \in \mathbb{P}_k$ iz $f_i(q) = 0$, za $i = 1, 2, \dots, N$, slijedi $q \equiv 0$.

Pokazat ćemo da postoji konstanta $C = C(\Omega)$ za koju vrijedi

$$\forall v \in W^{k+1,p}(\Omega), \quad \|v\|_{k+1,p;\Omega} \leq C \left(|v|_{k+1,p;\Omega} + \sum_{i=1}^N |f_i(v)| \right). \quad (4.43)$$

Iz te nejednakosti slijedi (4.42); Da bismo se uvjerali odaberimo za zadano $v \in W^{k+1,p}(\Omega)$ funkciju $q \in \mathbb{P}_k$ takvu da je $f_i(v + q) = 0$, za $i = 1, 2, \dots, N$. Tada je

$$\begin{aligned} \inf_{p \in \mathbb{P}_k} \|v + p\|_{k+1,p;\Omega} &\leq \|v + q\|_{k+1,p;\Omega} \leq C \left(|v + q|_{k+1,p;\Omega} + \sum_{i=1}^N |f_i(v + q)| \right) \\ &= C|v|_{k+1,p;\Omega}. \end{aligned}$$

Dokažimo (4.43) metodom kontradikcije. Pretpostavimo da tražena konstanta ne postoji. Tada bi postojao niz funkcija (v_l) iz $W^{k+1,p}(\Omega)$, sa svojstvom

$$\forall l \in \mathbb{N}, \quad \|v_l\|_{k+1,p;\Omega} = 1, \quad \lim_{l \rightarrow \infty} \left(\|v_l\|_{k+1,p;\Omega} + \sum_{i=1}^N |f_i(v_l)| \right) = 0.$$

Prostor $W^{k+1,p}(\Omega)$ je kompaktno uložen u $W^{k,p}(\Omega)$, te stoga niz (v_l) ima podniz koji konvergira u $W^{k,p}(\Omega)$. Taj ćemo podniz ponovo označavati s (v_l) , a njegov limes označavamo s v . Zbog

$$\lim_{l \rightarrow \infty} \|v_l - v\|_{k,p;\Omega} = 0, \quad \lim_{l \rightarrow \infty} \|v_l\|_{k+1,p;\Omega} = 0,$$

zaključujemo da niz (v_l) konvergira prema v u $W^{k+1,p}(\Omega)$ i da je $\|v\|_{k+1,p;\Omega} = 0$. Stoga je $v \in \mathbb{P}_k$, a zbog neprekidnosti linearnih formi slijedi $|f_i(v)| = 0$, za $i = 1, 2, \dots, N$. Time dobivamo $v \equiv 0$, što je kontradikcija s $\|v_l\|_{k+1,p;\Omega} = 1$, jer odatle izlazi $\|v\|_{k+1,p;\Omega} = 1$. \square

4.8 Dodatak 2. Inverzne ocjene

U definiciji regularne familije triangulacija (Definicija 4.1) uveli smo neka ograničenja na izbor triangulacije. Ta ćemo ograničenja dopuniti još jednim uvjetom.

Definicija 4.2 Familija triangulacija $(\mathcal{T}_h)_{h>0}$ je kvaziuniformna ako je regularna i ako postoji broj $\nu > 0$ takav da

$$\forall K \in \bigcup_h \mathcal{T}_h, \quad \frac{h}{h_K} \leq \nu.$$

Teorem 4.13 Neka su zadovoljeni uvjet H1–H3 i neka je familija triangulacija kvaziuniformna. Neka su zadani parovi brojeva (l, r) i (m, q) , s $0 \leq l \leq m$ te $r, q \in [1, \infty]$, takvi da je

$$\hat{P} \subset W^{l,r}(\hat{K}) \cap W^{m,q}(\hat{K}).$$

Tada postoji konstanta $C = C(\sigma, \nu, l, r, m, q)$ takva da je

$$\forall v \in X_h, \quad \left(\sum_{K \in \mathcal{T}_h} |v|_{m,q;K}^q \right)^{1/q} \leq \frac{C}{h^{\max\{0, (n/r) - (n/q)\}} h^{m-l}} \left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^r \right)^{1/r},$$

ako je $p, q < \infty$, s time da

$$\begin{aligned} \left(\sum_{K \in \mathcal{T}_h} |v|_{m,q;K}^q \right)^{1/q} & \text{ zamjenjujemo s } \max_{K \in \mathcal{T}_h} |v|_{m,\infty;K} \quad \text{ako je } q = \infty \\ \left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^r \right)^{1/r} & \text{ zamjenjujemo s } \max_{K \in \mathcal{T}_h} |v|_{l,\infty;K} \quad \text{ako je } r = \infty. \end{aligned}$$

Dokaz. Za svaku funkciju $v \in X_h$ i svaki element $K \in \mathcal{T}_h$ prema Teoremu 4.1 vrijedi

$$|\hat{v}_K|_{l,r;\hat{K}} \leq C_1 \|B_K\|^l |\det(B_K)|^{-1/r} |v|_{l,r;K}, \quad (4.44)$$

$$|v|_{m,q;K} \leq C_2 \|B_K^{-1}\|^m |\det(B_K)|^{1/q} |\hat{v}_K|_{m,q;\hat{K}}, \quad (4.45)$$

gdje je $\hat{v}_K = v \circ F_K$ i $F_K(\hat{x}) = B_K \hat{x} + b_K$ pripadno afino preslikavanje, a konstante C_1 i C_2 ovise o l, r, m i q .

Uvedimo prostor

$$\hat{N} = \{\hat{p} \in \hat{P} : |\hat{p}|_{l,r;\hat{K}} = 0\} = \begin{cases} \{0\} & \text{za } l = 0 \\ \hat{P} \cap \mathbb{P}_{l-1} & \text{za } l \geq 1. \end{cases}$$

Uočimo da je preslikavanje

$$\hat{p} \mapsto |\hat{p}|_{l,r;\hat{K}}$$

jedna norma na kvocijentnom prostoru \hat{P}/\hat{N} . S druge strane i

$$\hat{p} \mapsto \inf_{\hat{s} \in \hat{N}} \|\hat{p} + \hat{s}\|_{m,q;\hat{K}}$$

je norma na \hat{P}/\hat{N} i budući da se radi o konačnodimenzionalnom prostoru te su norme ekvivalentne. Konstante ekvivalencije ovise o l, r, m, q i elementu \hat{K} ; ovisnost o \hat{K} nadalje zanemarujemo budući da je referentni element fiksiran.

Zbog uvjeta $m \geq l$ imamo da je $|\hat{p}|_{m,q;\hat{K}} = 0$ za sve $\hat{p} \in \hat{N}$, pa stoga vrijedi

$$|\hat{p}|_{m,q;\hat{K}} \leq \inf_{\hat{s} \in \hat{N}} \|\hat{p} + \hat{s}\|_{m,q;\hat{K}} \leq \hat{C} |\hat{p}|_{l,r;\hat{K}}, \quad (4.46)$$

gdje je \hat{C} konstanta iz ekvivalencije normi. Sada koristeći (4.45), (4.44), (4.46), (??), Lemu 4.1

$$\begin{aligned} |v|_{m,q;K} &\leq C_2 \|B_K^{-1}\|^m |\det(B_K)|^{1/q} |\hat{v}_K|_{m,q;\hat{K}} \\ &\leq \hat{C} C_2 \|B_K^{-1}\|^m |\det(B_K)|^{1/q} |\hat{v}_K|_{l,r;\hat{K}} \\ &\leq \hat{C} C_1 C_2 \|B_K^{-1}\|^m \|B_K\|^l |\det(B_K)|^{(1/q)-(1/r)} |v|_{l,r;K} \\ &\leq \hat{C} C_1 C_2 \left(\frac{\hat{h}}{\rho_K}\right)^m \left(\frac{h_K}{\hat{\rho}}\right)^l \left(\frac{|K|}{|\hat{K}|}\right)^{(1/q)-(1/r)} |v|_{l,r;K} \\ &= \hat{C} C_1 C_2 \frac{\hat{h}^m}{\hat{\rho}^l} |\hat{K}|^{(1/r)-(1/q)} \frac{h_K^l}{\rho_K^m} |K|^{(1/q)-(1/r)} |v|_{l,r;K}. \end{aligned}$$

Nadalje, volumen elementa možemo ocijeniti kao $|K| \leq \omega_n h_K^n \leq \omega_n h^n$, gdje je ω_n površina jedinične sfere u \mathbb{R}^n , te koristeći regularnost i kvaziuniformnost triangulacije, dobivamo

$$\frac{h_K^l}{\rho_K^m} \leq \sigma^m \frac{h_K^l}{h_K^m} = \sigma^m \frac{1}{h_K^{m-l}} \leq \sigma^m \nu^{m-l} \frac{1}{h^{m-l}},$$

gdje smo uzeli u obzir da je $m - l \geq 0$. Time dobivamo ocjenu

$$|v|_{m,q;K} \leq C \frac{(h^n)^{(1/q)-(1/r)}}{h^{m-l}} |v|_{l,r;K}, \quad (4.47)$$

gdje je

$$C = \hat{C} C_1 C_2 \frac{\hat{h}^m}{\hat{\rho}^l} |\hat{K}|^{(1/r)-(1/q)} \sigma^m \nu^{m-l} \omega_n^{(1/q)-(1/r)}.$$

Da bismo polazeći od (4.47) dobili globalne nejednakosti, promatrat ćemo različite slučajeve u ovisnosti o p i q .

1. Slučaj $q = \infty$. Tada možemo naći element $K_0 \in \mathcal{T}_h$ takav da je

$$\max_{K \in \mathcal{T}_h} |v|_{m,\infty;K} = |v|_{m,\infty;K_0} \leq \frac{h^{-(n/r)}}{h^{m-l}} |v|_{l,r;K_0}.$$

Desnu stranu možemo ocijeniti, u ovisnosti o tome je li $r < \infty$ ili $r = \infty$, izrazima

$$\frac{1}{h^{n/r} h^{m-l}} \left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^r \right)^{1/r} \quad \text{ili} \quad \frac{1}{h^{m-l}} \max_{K \in \mathcal{T}_h} |v|_{l,\infty;K}.$$

2. Slučaj $q < \infty$. Sumiranjem (4.47) po svim elementima dobivamo

$$\left(\sum_{K \in \mathcal{T}_h} |v|_{m,q;K}^q \right)^{1/q} \leq \frac{(h^n)^{(1/q)-(1/r)}}{h^{m-l}} \left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^q \right)^{1/q}. \quad (4.48)$$

2.1 $r \leq q$. U ovom slučaju vrijedi (provjerite)

$$\left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^q \right)^{1/q} \leq \left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^r \right)^{1/r},$$

pa rezultat slijedi.

2.2 $q < r < \infty$. U ovom slučaju primijenjujemo Hölderovu nejednakost:

$$\begin{aligned} \sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^q &\leq \left(\sum_{K \in \mathcal{T}_h} 1^{r/(r-q)} \right)^{1-q/r} \left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^r \right)^{q/r} \\ &= \text{card}(\mathcal{T}_h)^{1-q/r} \left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^r \right)^{q/r} \end{aligned}$$

Broj elemenata u triangulaciji možemo ocijeniti na sljedeći način:

$$\text{card}(\mathcal{T}_h) \leq \frac{1}{\omega_n} \frac{|\Omega|}{(\min_{K \in \mathcal{T}_h} \rho_K)^n} \leq \frac{\sigma^n}{\omega_n} \frac{|\Omega|}{(\min_{K \in \mathcal{T}_h} h_K)^n} \leq \frac{\sigma^n \nu^n |\Omega|}{\omega_n h^n} = C \frac{1}{h^n}.$$

Time dobivamo ocjenu

$$\left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^q \right)^{1/q} \leq C \left(\frac{1}{h^n} \right)^{(1/q)-(1/r)} \left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^r \right)^{1/r},$$

što, primijenjeno na (4.48), konačno daje

$$\left(\sum_{K \in \mathcal{T}_h} |v|_{m,q;K}^q \right)^{1/q} \leq \frac{1}{h^{m-l}} \left(\sum_{K \in \mathcal{T}_h} |v|_{l,r;K}^r \right)^{1/r}.$$

2.3 $q < r = \infty$ Imamo

$$\begin{aligned} \left(\sum_{K \in \mathcal{T}_h} |v|_{l,\infty;K}^q \right)^{1/q} &\leq \left(\sum_{K \in \mathcal{T}_h} 1 \right)^{1/q} \max_{K \in \mathcal{T}_h} |v|_{l,\infty;K} = \text{card}(\mathcal{T}_h)^{1/q} \max_{K \in \mathcal{T}_h} |v|_{l,\infty;K} \\ &\leq C \left(\frac{1}{h^n} \right)^{(1/q)} \max_{K \in \mathcal{T}_h} |v|_{l,\infty;K}. \end{aligned}$$

Zajedno s (4.48) to daje

$$\left(\sum_{K \in \mathcal{T}_h} |v|_{m,q;K}^q \right)^{1/q} \leq \frac{1}{h^{m-l}} \max_{K \in \mathcal{T}_h} |v|_{l,\infty;K}.$$

Time je nejednakost izvedena u svim slučajevima. \square

Radi ilustracije specijalizirajmo Teorem 4.13 na dva slučaja koji se često koriste. Pretpostavimo da je $V_h \subset C(\bar{\Omega}) \cap H^1(\Omega)$ i da su ispunjene pretpostavke Teorema 4.13. Tada imamo (za $m = l = 0$ i $q = \infty, r = 2$)

$$\forall v \in V_h, \quad |v|_{\infty;\Omega} \leq \frac{C}{h^{n/2}} \|v\|_{2;\Omega},$$

te (za $m = 1, l = 0$ i $q = r = 2$)

$$\forall v \in V_h, \quad |v|_{1,2;\Omega} \leq \frac{C}{h} \|v\|_{2;\Omega}.$$

Bibliografija

- [1] S. C. Brenner and Scott L. R. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, 2002.
- [2] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 1998.

-
- [3] P. Grisvard. *Elliptic problems in nonsmooth domains*. Pitman Publishing INC, London, 1985.
- [4] Svetozar Kurepa. *Funkcionalna analiza. Elementi teorije operatora*. Školska knjiga, Zagreb, 1990.

5

Numerička integracija

Integrale koji se pojavljuju u varijacijskoj formulaciji rubne zadaće općenito moramo računati približno, pomoću formula numeričke integracije. Pri tome se uvodi dodatna greška u numerički postupak koju je potrebno procijeniti. Cilj je koristiti dovoljno točne formule kako asimptotski red točnosti metode ne bi bio umanjen.

5.1 Primjena formula numeričke integracije

Da bismo prodiskutirali kako u teoriju uključiti efekte numeričke integracije specificirat ćemo u potpunosti promatranu varijacijsku zadaću.

Neka je $\Omega \subset \mathbb{R}^n$ ograničena poligonalna domena. Promatramo homogenu Dirichletovu zadaću za diferencijalnu jednadžbu

$$-\operatorname{div}(\mathbf{A}\nabla u) = f \quad \text{u } \Omega, \quad (5.1)$$

pri čemu je $\mathbf{A}(x) = (a_{ij}(x))_{i,j=1}^n$, uniformno pozitivno definitna matrica, definirana na $\bar{\Omega}$. Preciznije, pretpostavljamo da su funkcije $f \in L^2(\Omega)$ i $a_{ij} \in L^\infty(\Omega)$, definirane na čitavom $\bar{\Omega}$, te da postoji konstanta $\beta > 0$ takva da je

$$\forall x \in \bar{\Omega}, \forall \xi \in \mathbb{R}^n, \quad \sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \geq \beta \sum_{i=1}^n \xi_i^2. \quad (5.2)$$

Prostor u kojem tražimo rješenje je $V = H_0^1(\Omega)$, a bilinearna forma i funkcional desne strane dani su formulama

$$a(u, v) = \int_{\Omega} \sum_{i,j=1}^n a_{ij} \partial_i u \partial_j v \, dx, \quad (5.3)$$

$$\langle F, v \rangle = \int_{\Omega} f v \, dx. \quad (5.4)$$

Na domeni Ω zadana je (familija) triangulacija \mathcal{T}_h koja se sastoji od konačnih elemenata (K, P_K, Σ_K) . Kao i u prethodnom poglavlju pretpostavljamo da su zadovoljeni uvjeti (H1), (H2), (H3) i (H4). Referentni element označavamo uobičajeno s $(\hat{K}, \hat{P}, \hat{\Sigma})$ i uzimamo da je minimalno $\hat{P} \subset H^1(\hat{K})$.

Pripadni prostor konačnih elemenata V_h je potprostor od $H_0^1(\Omega)$. Označimo elemente baze prostora V_h sa w_k , $k = 1, 2, \dots, M$. Tada je potrebno naći koeficijente u_k , $k = 1, 2, \dots, M$, koji zadovoljavaju

$$\sum_{k=1}^M a(w_k, w_m) u_k = \langle F, w_m \rangle, \quad m = 1, 2, \dots, M, \quad (5.5)$$

gdje je

$$a(w_k, w_m) = \sum_{K \in \mathcal{T}_h} \int_K \sum_{i,j=1}^n a_{ij} \partial_i w_k \partial_j w_m dx \quad (5.6)$$

$$\langle F, w_m \rangle = \sum_{K \in \mathcal{T}_h} \int_K f w_m dx. \quad (5.7)$$

U praksi se integrali u ovim izrazima rijetko izračunavaju egzaktno već se primijenjuju formule numeričke integracije.

Kod affine familije elemenata dovoljno je integracijsku formulu zadati na referentnom elementu. Ako je, naime,

$$F_K(\hat{x}) = B_K \hat{x} + b_K, \quad F_K(\hat{K}) = K,$$

invertibilno afino preslikavanje kojim je definiran element K , onda je

$$\int_K \phi(x) dx = |\det(B_K)| \int_{\hat{K}} \hat{\phi}(\hat{x}) d\hat{x}, \quad (5.8)$$

gdje je $\hat{\phi} = \phi \circ F_K$.

Formule numeričke integracije općenito imaju oblik

$$\int_{\hat{K}} \hat{\phi}(\hat{x}) d\hat{x} \approx \sum_{l=1}^L \hat{\omega}_l \hat{\phi}(\hat{b}_l), \quad (5.9)$$

gdje su $\hat{\omega}_l$ integracijske težine, a \hat{b}_l integracijske točke. Mi ćemo promatrati samo one integracijske formule u kojima je $\hat{\omega}_l > 0$, a integracijske točke \hat{b}_l pripadaju skupu \hat{K} .

Integracijska formula na referentnom elementu \hat{K} generira integracijsku formulu na elementu K pravilom

$$\int_K \phi(x) dx \approx \sum_{l=1}^L \omega_{l,K} \phi(b_{l,K}), \quad (5.10)$$

gdje je

$$\omega_{l,K} = |\det(B_K)| \hat{\omega}_l, \quad b_{l,K} = F_K(\hat{b}_l).$$

Uvedimo funkcionalne greške integracijskih formula

$$\hat{E}(\hat{\phi}) = \int_{\hat{K}} \hat{\phi}(\hat{x}) d\hat{x} - \sum_{l=1}^L \hat{\omega}_l \hat{\phi}(\hat{b}_l), \quad (5.11)$$

$$E_K(\phi) = \int_K \phi(x) dx - \sum_{l=1}^L \omega_{l,K} \phi(b_{l,K}). \quad (5.12)$$

Evidentno je

$$E_K(\phi) = |\det(B_K)| \hat{E}(\hat{\phi}). \quad (5.13)$$

Točnost integracijskih formula može se mjeriti skupom polinoma na kojima je formula egzaktna. Tako kažemo da je formula (5.10) egzaktna na \mathbb{P}_k ako je $\hat{E}(\hat{\phi}) = 0$ za svako $\hat{\phi} \in \mathbb{P}_k$.

Pogledajmo neke primjere integracijskih formula na simpleksima. Neka je \hat{K} simpleks s vrhovima \hat{a}_i , $i = 1, 2, \dots, n+1$. Tada je

$$\hat{a} = \frac{1}{n+1} \sum_{i=1}^{n+1} \hat{a}_i$$

baricentar simpleksa. Formula

$$\int_{\hat{K}} \hat{\phi}(\hat{x}) d\hat{x} \approx |\hat{K}| \hat{\phi}(\hat{a}), \quad (5.14)$$

je točna na polinomima stupnja manjeg ili jednakog 1 (\mathbb{P}_1). To se lako vidi iz činjenice da za baricentričke koordinate vrijedi

$$\int_{\hat{K}} \hat{\lambda}_i(\hat{x}) d\hat{x} = \frac{|\hat{K}|}{n+1}.$$

Neka je $n = 2$. Označimo s \hat{a}_{ij} za $1 \leq i < j \leq 3$ polovišta stranica, a s \hat{a}_{123} baricentar. Tada je formula

$$\int_{\hat{K}} \hat{\phi}(\hat{x}) d\hat{x} \approx \frac{|\hat{K}|}{3} \sum_{1 \leq i < j \leq 3} \hat{\phi}(\hat{a}_{ij}), \quad (5.15)$$

točna na polinomima drugog stupnja, dok je formula

$$\int_{\hat{K}} \hat{\phi}(\hat{x}) d\hat{x} \approx \frac{|\hat{K}|}{60} \left(3 \sum_{i=1}^3 \hat{\phi}(\hat{a}_i) + 8 \sum_{1 \leq i < j \leq 3} \hat{\phi}(\hat{a}_{ij}) + 27 \hat{\phi}(\hat{a}_{123}) \right), \quad (5.16)$$

točna na \mathbb{P}_3 .

Uočimo da diskretan problem koji treba riješiti sada dobiva oblik

$$\sum_{k=1}^M a_h(w_k, w_m) u_k = f_h(w_m), \quad m = 1, 2, \dots, M, \quad (5.17)$$

gdje je

$$a_h(w_k, w_m) = \sum_{K \in \mathcal{T}_h} \sum_{l=1}^L \omega_{l,K} \sum_{i,j=1}^n (a_{ij} \partial_i w_k \partial_j w_m)(b_{l,K}) \quad (5.18)$$

$$f_h(w_m) = \sum_{K \in \mathcal{T}_h} \sum_{l=1}^L \omega_{l,K} (f w_m)(b_{l,K}). \quad (5.19)$$

Stoga ga možemo formulirati na sljedeći način: Treba naći $u_h \in V_h$ koje zadovoljava

$$\forall v \in V_h, \quad a_h(u_h, v) = f_h(v), \quad (5.20)$$

gdje su bilinearna forma i funkcional desne strane dani formulama:

$$a_h(u, v) = \sum_{K \in \mathcal{T}_h} \sum_{l=1}^L \omega_{l,K} \sum_{i,j=1}^n (a_{ij} \partial_i u \partial_j v)(b_{l,K}) \quad (5.21)$$

$$f_h(v) = \sum_{K \in \mathcal{T}_h} \sum_{l=1}^L \omega_{l,K} (f v)(b_{l,K}). \quad (5.22)$$

Da bi ove formule imale smisla nužno je da su koeficijenti jednadžbe i desna strana dobro definirani u integracijskim točkama. To će biti ispunjeno u slučaju neprekidnih funkcija. Uočimo još da su kod Lagrangeovih elemenata derivacije diskontinuirane na granici elemenata, tako da će one poprimati različite vrijednosti u različitim elementima, u integracijskim točkama koje se nalaze na granici susjednih elemenata.

Promjena bilinearne forme generira dva problema: Prvo, potrebna nam je generalizacije Céa-ine leme, i drugo, otvoren je problem eliptičnosti diskretne bilinearne forme.

Uvest ćemo ovu definiciju: za bilinearnu formu $a_h: V_h \times V_h \rightarrow \mathbb{R}$ kažemo da je **uniformno** V_h -eliptična ako postoji konstanta $\tilde{\alpha}$, neovisna o potprostor V_h , takva da je

$$\forall v \in V_h, \quad \tilde{\alpha} \|v\|^2 \leq a_h(v, v). \quad (5.23)$$

Prva Strangova lema nam daje potrebnu generalizaciju Céa-ine leme. Nova apstraktna ocjena greške sastoji se od sume greške interpolacije i grešaka aproksimacije bilinearne forme i funkcionala desne strane. Ovdje ponavljamo iskaz leme koja je dokazana u Sekciji 2.9.

Teorem 5.1 (prva Strangova lema) Promatramo familiju diskretnih problema čija je pridružena bilinearna forma uniformno V_h -eliptična. Tada postoji konstanta C , neovisna o potprostor V_h , takva da je

$$\|u - u_h\| \leq C \left(\inf_{v \in V_h} \left\{ \|u - v\| + \sup_{w \in V_h} \frac{|a(v, w) - a_h(v, w)|}{\|w\|} \right\} + \sup_{w \in V_h} \frac{|\langle F, w \rangle - f_h(w)|}{\|w\|} \right). \quad (5.24)$$

5.2 Dovoljan uvjet uniformne V_h -eliptičnosti

Da bismo dokazali uniformnu V_h -eliptičnost bilinearne forme iskoristiti ćemo regularnost familije triangulacija.

Teorem 5.2 Neka je zadana formula numeričke integracije

$$\int_{\hat{K}} \hat{\phi}(\hat{x}) d\hat{x} \approx \sum_{l=1}^L \hat{\omega}_l \hat{\phi}(\hat{b}_l),$$

na referentnom elementu $(\hat{K}, \hat{P}, \hat{\Sigma})$, za koju postoji cijeli broj $k' \geq 1$, takav da je

- $\hat{P} \subset \mathbb{P}_{k'}$;
- Unija točaka $\cup_{l=1}^L \{\hat{b}_l\}$ sadrži $\mathbb{P}_{k'-1}$ unisolventan skup, ili, integracijska formula je točna na $\mathbb{P}_{2k'-2}$.

Tada postoji konstanta $\tilde{\alpha}$, neovisna o potprostor V_h , takva da je

$$\forall v \in V_h, \quad \tilde{\alpha} |v|_{1,2;\Omega}^2 \leq a_h(v, v).$$

Dokaz. (i) Pretpostavit ćemo najprije da skup $\cup_{l=1}^L \{\hat{b}_l\}$ sadrži $\mathbb{P}_{k'-1}$ unisolventan skup. Koristeći pozitivnost integracijskih težina dobivamo da za svako $\hat{p} \in \hat{P}$ vrijedi

$$\sum_{l=1}^L \hat{\omega}_l \sum_{i=1}^n (\partial_i \hat{p}(\hat{b}_l))^2 = 0 \quad \Rightarrow \quad \partial_i \hat{p}(\hat{b}_l) = 0, \quad 1 \leq i \leq n, \quad 1 \leq l \leq L.$$

Uočimo da je $\partial_i \hat{p} \in \mathbb{P}_{k'-1}$, pa zbog svojstva unisolventnosti zaključujemo da sve prve parcijalne derivacije iščezavaju na \hat{K} . Stoga je

$$\hat{p} \mapsto \left(\sum_{l=1}^L \hat{\omega}_l \sum_{i=1}^n (\partial_i \hat{p}(\hat{b}_l))^2 \right)^{1/2}$$

norma na kvocijentnom prostoru \hat{P}/\mathbb{P}_0 . S druge strane i preslikavanje $\hat{p} \mapsto |\hat{p}|_{1,2;\hat{K}}$ je norma na istom prostoru, pa zbog ekvivalencije normi na konačnodimenzionalnom prostoru zaključujemo da postoji konstanta \hat{C} , takva da je

$$\forall \hat{p} \in \hat{P}, \quad \hat{C} |\hat{p}|_{1,2;\hat{K}}^2 \leq \sum_{l=1}^L \hat{\omega}_l \sum_{i=1}^n (\partial_i \hat{p}(\hat{b}_l))^2.$$

Ako je ispunjen uvjet da je integracijska formula točna na $\mathbb{P}_{2k'-2}$, onda dobvena nejednakost prelazi u jednakost s $\hat{C} = 1$, jer je $(\partial_i \hat{p})^2 \in \mathbb{P}_{2k'-2}$.

(ii) Za proizvoljan element triangulacije K ocijenit ćemo aproksimaciju izraza

$$\int_K \sum_{i,j=1}^n a_{ij} \partial_i v \partial_j v \, dx.$$

Označimo $v|_K = p_K$; zbog pozitivne definitnosti matrice koeficijenata imamo

$$\begin{aligned} \sum_{l=1}^L \omega_{l,K} \sum_{i,j=1}^n (a_{ij} \partial_i v \partial_j v)(b_{l,K}) &= \sum_{l=1}^L \omega_{l,K} \sum_{i,j=1}^n (a_{ij} \partial_i p_K \partial_j p_K)(b_{l,K}) \\ &\geq \beta \sum_{l=1}^L \omega_{l,K} \sum_{i=1}^n (\partial_i p_K(b_{l,K}))^2. \end{aligned}$$

Neka je $F(\hat{x}) = B_K \hat{x} + b_K$ afino preslikavanje koje preslikava \hat{K} na K . Definiramo $\hat{p}_K = p_K \circ F_K$, i kako je triangulacija afina imamo $\hat{p}_K \in \hat{P}$. Nadalje, iz

$$D\hat{p}_K(\hat{b}_l) \cdot \xi = Dp_K(b_{l,K}) \cdot B_K \xi,$$

izlazi $\|D\hat{p}_K(\hat{b}_l)\| \leq \|Dp_K(b_{l,K})\| \|B_K\|$. Koristeći $\omega_{l,K} = |\det(B_K)| \hat{\omega}_l$ i tvrdnju (i) dobivamo

$$\begin{aligned} \sum_{l=1}^L \omega_{l,K} \sum_{i=1}^n (\partial_i p_K(b_{l,K}))^2 &\geq \|B_K\|^{-2} \sum_{l=1}^L \omega_{l,K} \sum_{i=1}^n (\partial_i \hat{p}_K(\hat{b}_l))^2 \\ &= |\det(B_K)| \|B_K\|^{-2} \sum_{l=1}^L \hat{\omega}_l \sum_{i=1}^n (\partial_i \hat{p}_K(\hat{b}_l))^2 \\ &\geq \hat{C} |\det(B_K)| \|B_K\|^{-2} |\hat{p}_K|_{1,2;\hat{K}}^2 \end{aligned}$$

Sada primjenom Teorema 4.1 dobivamo

$$|p_K|_{1,2;K} \leq C \|B_K^{-1}\| |\det(B_K)|^{1/2} |\hat{p}_K|_{1,2;\hat{K}}$$

pa stoga izlazi

$$\sum_{l=1}^L \omega_{l,K} \sum_{i=1}^n (\partial_i p_K(b_{l,K}))^2 \geq \hat{C} (\|B_K^{-1}\| \|B_K\|)^{-2} |p_K|_{1,2;K}^2.$$

Prema Lemi 4.1 imamo

$$\|B_K^{-1}\| \|B_K\| \leq \frac{\hat{h}}{\rho} \frac{h}{\hat{\rho}} \leq C,$$

gdje je C neovisno o $K \in \mathcal{T}_h$. Time smo dobili

$$\forall v \in V_h, \quad \sum_{l=1}^L \omega_{l,K} \sum_{i,j=1}^n (a_{ij} \partial_i v \partial_j v)(b_{l,K}) \geq \tilde{\alpha} |v|_{1,2;K}^2.$$

(iii) Sumiranjem dobivamo

$$\begin{aligned} \forall v \in V_h, \quad a_h(v, v) &= \sum_{K \in \mathcal{T}_h} \sum_{l=1}^L \omega_{l,K} \sum_{i,j=1}^n (a_{ij} \partial_i v \partial_j v)(b_{l,K}) \\ &\geq \tilde{\alpha} \sum_{K \in \mathcal{T}_h} |v|_{1,2;K}^2 = \tilde{\alpha} |v|_{1,2;\Omega}^2. \quad \square \end{aligned}$$

Primjeri. Ako je referentni element n -simpleks tipa (1), onda je dovoljno koristiti integracijsku formulu (5.14). Naime, $\{\hat{a}\}$ je \mathbb{P}_0 unisolventan skup. I drugi uvjet je zadovoljen jer je $k' = 1$, pa je $2k' - 2 = 0$, a formula je egzaktna na \mathbb{P}_1 .

Ako je referentni element trokut tipa (2) ($n = 2$), onda je dovoljno upotrijebiti formulu (5.15). Naime, $k' = 2$, a skup polovišta stranica je \mathbb{P}_1 unisolventan skup. Isto tako vrijedi da je formula (5.15) egzaktna na $\mathbb{P}_{2k'-2} = \mathbb{P}_2$.

Ako je referentni element trokut tipa (3) ($n = 2$), onda je dovoljno upotrijebiti formulu (5.16). Tada je $k' = 3$, a skup $(\cup_i \hat{a}_i) \cup (\cup_{i < j} \hat{a}_{ij})$ je \mathbb{P}_2 unisolventan. S druge strane, u ovom slučaju drugi uvjet nije zadovoljen jer formula nije egzaktna na $\mathbb{P}_{2k'-2} = \mathbb{P}_4$.

5.3 Konzistentnost. Bramble-Hilbertova lema

Nadalje ćemo promatrati samo specijalan slučaj u kome je

$$\hat{P} = \mathbb{P}_k$$

za neki $k \geq 1$. Pretpostavimo li da rješenje varijacijske zadaće leži u prostoru $H^{k+1}(\Omega)$ i da je V_h -interpolacijski operator dobro definiran, onda imamo

$$\inf_{v \in V_h} \|u - v\|_{1,2;\Omega} \leq \|u - \Pi_h u\|_{1,2;\Omega} \leq Ch^k |u|_{k+1,2;\Omega}.$$

Dakle, uz egzaktnu integraciju metoda ima red konvergencije $O(h^k)$. Zanimaju nas uvjeti uz koje numerička integracija neće smanjiti red konvergencije metode.

Uz pretpostavku da je diskretna bilinearna forma uniformno V_h -eliptička, prema prvoj Strangovoj lemi trebmo dokazati sljedeće ocjene:

$$\sup_{w \in V_h} \frac{|a(\Pi_h u, w) - a_h(\Pi_h u, w)|}{\|w\|_{1,2;\Omega}} \leq C(a_{ij}, u) h^k, \quad (5.25)$$

$$\sup_{w \in V_h} \frac{|\langle F, w \rangle - f_h(w)|}{\|w\|_{1,2;\Omega}} \leq C(f) h^k. \quad (5.26)$$

Ove ocjene predstavljaju konzistentnost metode, dok uniformna eliptičnost daje njenu stabilnost.

Budući da vrijedi

$$a(\Pi_h u, w) - a_h(\Pi_h u, w) = \sum_{K \in \mathcal{T}_h} E_K \left(\sum_{i,j=1}^n a_{ij} \partial_i(\Pi_h u) \partial_j w \right) \quad (5.27)$$

$$\langle F, w \rangle - f_h(w) = \sum_{K \in \mathcal{T}_h} E_K(fw), \quad (5.28)$$

dovoljno je dokazati sljedeće *lokalne* ocjene:

$$\forall p', p \in \mathbb{P}_k, \quad |E_K \left(\sum_{i,j=1}^n a_{ij} \partial_i p' \partial_j p \right)| \leq C(a_{ij}|_K, \partial_i p') h_K^k \|\partial_j p\|_{2;K} \quad (5.29)$$

$$\forall p \in \mathbb{P}_k, \quad |E_K(fp)| \leq C(f|_K) h_K^k \|p\|_{1,2;K}, \quad (5.30)$$

Cilj je sada sljedeći. Treba dobiti cjene (5.29) i (5.30), s konstantama na desnoj strani u odgovarajućem obliku, tako da sumiranjem možemo dobiti tražene ocjene (5.25) i (5.26).

Prvo dokazujemo pomoćni rezultat.

Teorem 5.3 (Bramble-Hilbertova lema) Neka je Ω ograničena Lipschitzova domena u \mathbb{R}^n . Za dani cijeli broj $k \geq 0$ i $p \in [1, \infty]$, neka je F linearan i neprekidan funkcional na $W^{k+1,p}(\Omega)$, sa svojstvom

$$\forall p \in \mathbb{P}_k, \quad \langle F, p \rangle = 0. \quad (5.31)$$

Tada postoji konstanta $C = C(\Omega)$, takva da je

$$\forall v \in W^{k+1,p}(\Omega), \quad |\langle F, v \rangle| \leq C \|F\|_{k+1,p;\Omega}^* |v|_{k+1,p;\Omega}, \quad (5.32)$$

gdje je $\|\cdot\|_{k+1,p;\Omega}^*$ norma u dualu prostora $W^{k+1,p}(\Omega)$.

Dokaz. Neka je $v \in W^{k+1,p}(\Omega)$ proizvoljna funkcija. Tada za bilo koje $p \in \mathbb{P}_k$ vrijedi $\langle F, v+p \rangle = \langle F, v \rangle$, pa imamo

$$\forall p \in \mathbb{P}_k, \quad |\langle F, v \rangle| = |\langle F, v+p \rangle| \leq \|F\|_{k+1,p;\Omega}^* \|v+p\|_{k+1,p;\Omega},$$

i stoga

$$|\langle F, v \rangle| \leq \|F\|_{k+1,p;\Omega}^* \inf_{p \in \mathbb{P}_k} \|v+p\|_{k+1,p;\Omega}.$$

Tvrđnja sada slijedi primjenom Teorema 4.12. \square

Sljedeća dva teorema daju nam tražene lokalne ocjene. Njihovi se dokazi baziraju na uporabi Bramble-Hilbertove leme, i kako su dosta tehnički složeni, nećemo ih ovdje iznositi. Detalji se mogu naći u Ciarletovoj knjizi [1].

Teorem 5.4 Neka postoji $k \geq 1$, takav da je

$$\hat{P} = \mathbb{P}_k, \quad (5.33)$$

$$\forall \hat{\phi} \in \mathbb{P}_{2k-2}, \quad \hat{E}(\hat{\phi}) = 0. \quad (5.34)$$

Tada postoji konstanta C , neovisna o $K \in \mathcal{T}_h$ i h , takva da vrijedi

$$\forall a \in W^{k,\infty}(K), \forall p, p' \in \mathbb{P}_k$$

$$|E_K(a \partial_i p' \partial_j p)| \leq Ch_K^k \|a\|_{k,\infty;K} \|\partial_i p'\|_{k-1,2;K} |\partial_j p|_{0,2;K} \quad (5.35)$$

$$\leq Ch_K^k \|a\|_{k,\infty;K} \|p'\|_{k,2;K} |p|_{1,2;K} \quad (5.36)$$

Dokaz. [1], str. 193.

Teorem 5.5 Neka postoji $k \geq 1$, takav da je

$$\hat{P} = \mathbb{P}_k, \quad (5.37)$$

$$\forall \hat{\phi} \in \mathbb{P}_{2k-2}, \quad \hat{E}(\hat{\phi}) = 0, \quad (5.38)$$

i neka je $q \in [1, \infty]$ broj koji zadovoljava nejednakost

$$k - \frac{n}{q} > 0. \quad (5.39)$$

Tada postoji konstanta C , neovisna o $K \in \mathcal{T}_h$ i h , takva da vrijedi

$$\forall f \in W^{k,q}(K), \forall p \in \mathbb{P}_k$$

$$|E_K(fp)| \leq Ch_K^k |K|^{(1/2)-(1/q)} \|f\|_{k,q;K} \|p\|_{1,2;K} \quad (5.40)$$

Dokaz. [1], str. 195.

Naš osnovni rezultat je sljedeći teorem:

Teorem 5.6 Neka su ispunjeni uvjeti (H1)–(H4) te neka postoji cijeli broj $k \geq 1$, takav da je

$$\hat{P} = \mathbb{P}_k, \quad (5.41)$$

$$H^{k+1}(\hat{K}) \subset C^s(\hat{K}), \quad (5.42)$$

gdje je s najviši stupanj parcijalne derivacije koja se javlja u stupnjevima slobode $\hat{\Sigma}$. Neka je k tome

$$\forall \hat{\phi} \in \mathbb{P}_{2k-2}, \quad \hat{E}(\hat{\phi}) = 0. \quad (5.43)$$

Tada ako rješenje $u \in H_0^1(\Omega)$ varijacijske zadaće koja odgovara podacima (5.3), (5.4) pripada prostoru $H^{k+1}(\Omega)$, ako je $a_{i,j} \in W^{k,\infty}(\Omega)$ za $1 \leq i, j \leq n$, i ako je $f \in W^{k,q}(\Omega)$ za neki $q \geq 2$, $k > n/q$, onda postoji konstanta C , neovisna o h , takva da vrijedi

$$\|u - u_h\|_{1,2;\Omega} \leq Ch^k (|u|_{k+1,2;\Omega} + \sum_{i,j=1}^n \|a_{i,j}\|_{k,\infty;\Omega} \|u\|_{k+1,2;\Omega} + \|f\|_{k,q;\Omega}), \quad (5.44)$$

gdje je $u_h \in V_h$ rješenje diskretnog problema.

Dokaz. U uvjetima teorema možemo primijeniti Teorem 4.3 koji daje egzistenciju konstante C , neovisne o h , takve da je

$$\|u - \Pi_h u\|_{1,2;\Omega} \leq Ch^k |u|_{k+1,2;\Omega}.$$

Polazeći od (5.27) i koristeći Teorem 5.4, dobivamo

$$\begin{aligned} |a(\Pi_h u, w) - a_h(\Pi_h u, w)| &\leq \sum_{K \in \mathcal{T}_h} |E_K \left(\sum_{i,j=1}^n a_{i,j} \partial_i (\Pi_h u|_K) \partial_j w|_K \right)| \\ &\leq C \sum_{K \in \mathcal{T}_h} h_K^k \sum_{i,j=1}^n \|a_{i,j}\|_{k,\infty;K} \|\Pi_h u\|_{k,2;K} |w|_{1,2;K} \\ &\leq Ch^k \left(\sum_{i,j=1}^n \|a_{i,j}\|_{k,\infty;\Omega} \right) \left(\sum_{K \in \mathcal{T}_h} \|\Pi_h u\|_{k,2;K}^2 \right)^{1/2} |w|_{1,2;\Omega} \end{aligned}$$

S druge strane, koristeći ponovo Teorem 4.3 dobivamo

$$\begin{aligned} \left(\sum_{K \in \mathcal{T}_h} \|\Pi_h u\|_{k,2;K}^2 \right)^{1/2} &\leq \|u\|_{k,2;\Omega} + \left(\sum_{K \in \mathcal{T}_h} \|u - \Pi_h u\|_{k,2;K}^2 \right)^{1/2} \\ &\leq \|u\|_{k,2;\Omega} + Ch |u|_{k+1,2;\Omega} \leq C \|u\|_{k+1,2;\Omega}, \end{aligned}$$

i stoga

$$\begin{aligned} &\inf_{v \in V_h} \left\{ \|u - v\|_{1,2;\Omega} + \sup_{w \in V_h} \frac{|a(v, w) - a_h(v, w)|}{\|w\|_{1,2;\Omega}} \right\} \\ &\leq \|u - \Pi_h u\|_{1,2;\Omega} + \sup_{w \in V_h} \frac{|a(\Pi_h u, w) - a_h(\Pi_h u, w)|}{\|w\|_{1,2;\Omega}} \\ &\leq Ch^k \left(|u|_{k+1,2;\Omega} + \sum_{i,j=1}^n \|a_{i,j}\|_{k,\infty;\Omega} \|u\|_{k+1,2;\Omega} \right). \end{aligned}$$

Slično, polazeći od (5.28) i koristeći Teorem 5.5, dobivamo

$$\begin{aligned} |\langle F, w \rangle - f_h(w)| &\leq \sum_{K \in \mathcal{T}_h} |E_K(fw)| \\ &\leq C \sum_{K \in \mathcal{T}_h} h_K^k |K|^{(1/2)-(1/q)} \|f\|_{k,q;K} \|w\|_{1,2;K} \\ &\leq Ch^k |\Omega|^{(1/2)-(1/q)} \|f\|_{k,q;\Omega} \|w\|_{1,2;\Omega}, \end{aligned}$$

gdje smo iskoristili nejednakost

$$\sum_K |a_K b_K c_K| \leq \left(\sum_K |a_K|^\alpha \right)^{1/\alpha} \left(\sum_K |b_K|^\beta \right)^{1/\beta} \left(\sum_K |c_K|^\gamma \right)^{1/\gamma}$$

koja vrijedi za sve $\alpha \geq 1$, $\beta \geq 1$ i $\gamma \geq 1$ koji zadovoljavaju

$$\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} = 1.$$

U našem smo slučaju uzeli $\beta = q$, $\gamma = 2$ i $1/\alpha = 1/2 - 1/q$; To je razlog zbog kojeg trebamo $q \geq 2$. Time smo dobili

$$\sup_{w \in V_h} \frac{|\langle F, w \rangle - f_h(w)|}{\|w\|_{1,2;\Omega}} \leq Ch^k |\Omega|^{(1/2)-(1/q)} \|f\|_{k,q;\Omega}.$$

Tvrđnja teorema sada slijedi iz prve Strangove leme, budući da je naša bilinearna forma uniformno V_h -eliptička. \square

Uvjet da formula numeričke integracije mora biti točna na \mathbb{P}_{2k-2} može se interpretirati na sljedeći način: Ako su koeficijenti $a_{i,j}$ konstantni onda integrali

$$\int_K a_{i,j} \partial_i u \partial_j v \, dx$$

moraju biti izračunati egzaktno.

Ako tvrdnju teorema specijaliziramo na n -simplekse, onda imamo sljedeće rezultate:

Za n -simplekse tipa (1) dobivamo $\|u - u_h\|_{1,2;\Omega} = O(h)$ ako je formula numeričke integracije točna na konstantama.

Za n -simplekse tipa (2) dobivamo $\|u - u_h\|_{1,2;\Omega} = O(h^2)$ ako je formula numeričke integracije točna na polinomima stupnja $n \leq 2$.

Za n -simplekse tipa (3) dobivamo $\|u - u_h\|_{1,2;\Omega} = O(h^3)$ ako je formula numeričke integracije točna na polinomima stupnja $n \leq 4$, itd.

Bibliografija

- [1] Philippe G. Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 4 of *studies in Mathematics and its Applications*. North-Holland, Amsterdam, 1978.

6

Princip maksimuma

6.1 Slabi i jaki princip maksimuma za eliptičke jednadžbe

Za eliptičku diferencijalnu jednadžbu u nedivergentnoj formi vrijede klasični principi maksimuma. U domeni $\Omega \subset \mathbb{R}^d$, promatramo diferencijalni operator

$$Lu = \sum_{i,j=1}^d a_{i,j}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(x) \frac{\partial u}{\partial x_i} + c(x)u, \quad (6.1)$$

pri čemu pretpostavljamo da je matrica $A(x) = (a_{i,j}(x))$ u svakoj točki domene simetrična i pozitivno definirna. To znači da postoje pozitivne funkcije $\lambda, \Lambda: \Omega \rightarrow \mathbb{R}$ takve da je za svako $x \in \Omega$

$$0 < \lambda(x) \|\vec{\xi}\|^2 \leq A(x) \vec{\xi} \cdot \vec{\xi} \leq \Lambda(x) \|\vec{\xi}\|^2, \quad \forall \vec{\xi} \in \mathbb{R}^d.$$

Pri tome ćemo još pretpostavljati da

$$\frac{|b_i(x)|}{\lambda(x)} < +\infty, \quad x \in \Omega.$$

Za diferencijalni operator L koji udovoljava tim uvjetima kažemo da je **eliptički**, a ako je $\Lambda(x)/\lambda(x)$ ograničena funkcija, onda je operator **uniformno eliptički**.

Ovdje nećemo postavljati uvjete na glatkoću koeficijenata već ćemo samo pretpostavljati da je funkcija na koju operator djeluje glatka, tako da klasične derivacije imaju smisla.

Slabi princip maksimuma ima sljedeći iskaz:

Teorem 6.1 Neka je L eliptički operator (6.1) u ograničenoj domeni Ω i neka je $u \in C^2(\Omega) \cap C(\overline{\Omega})$ funkcija koja zadovoljava $Lu \geq 0$ i neka je $c = 0$. Tada je

$$\sup_{\Omega} u = \sup_{\partial\Omega} u.$$

U slučaju $Lu \leq 0$ i $c = 0$ vrijedi

$$\inf_{\Omega} u = \inf_{\partial\Omega} u.$$

Dokaz se može naći u [1].

Pretpostavimo sada da je $c \leq 0$ u Ω i označimo s $\Omega^+ \subset \Omega$ skup na kome je $u > 0$. Zbog neprekidnosti funkcije u skup Ω^+ je otvoren i ako imamo $Lu \geq 0$ onda je

$$L_0 u = \sum_{i,j=1}^d a_{i,j}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(x) \frac{\partial u}{\partial x_i} \geq -c(x)u \geq 0,$$

na Ω^+ . Sada po Teoremu 6.1 imamo

$$\sup_{\Omega} u^+ \leq \max(0, \sup_{\partial\Omega} u) = \sup_{\partial\Omega} u^+.$$

Analogno zaključujemo kod suprotne nejednakosti.

Posljedica 6.1 Neka je L eliptički operator (6.1) u ograničenoj domeni Ω i neka je $u \in C^2(\Omega) \cap C(\bar{\Omega})$ funkcija koja zadovoljava $Lu \geq 0$ i neka je $c \leq 0$. Tada je

$$\sup_{\Omega} u \leq \sup_{\partial\Omega} u^+.$$

U slučaju $Lu \leq 0$ i $c \leq 0$ vrijedi¹

$$\inf_{\Omega} u \geq \inf_{\partial\Omega} u^-.$$

Uočimo da u slučaju $Lu = 0$ vrijedi

$$\sup_{\Omega} |u| = \sup_{\partial\Omega} |u|.$$

Uvjet $c \leq 0$ ne možemo oslabiti što se vidi na primjeru spektralne zadaće $\Delta u + \lambda u = 0$, $u = 0$ na $\partial\Omega$.

Jaki princip maksimuma tvrdi da jedino konstantna funkcija postiže maksimum u unutrašnjosti domene.

Teorem 6.2 Neka je L uniformno eliptički operator (6.1) u domeni Ω (ne nužno ograničenoj) i neka je $c = 0$. Pretpostavimo da je $u \in C^2(\Omega)$ funkcija za koju je $Lu \geq 0$ ($Lu \leq 0$). Ako u postiže svoj maksimum (minimum) u unutrašnjosti domene Ω onda je u konstanta.

Ako je $c \leq 0$ i c/λ ograničena funkcija, onda u ne može postizati svoj nenegativni maksimum (nepozitivni minimum) u unutrašnjosti domene osim ako je konstanta.

Dokaz se može naći u [1].

¹ $u^- = \max(u, 0)$ i $u^+ = \min(u, 0)$.

6.2 Slabi princip maksimuma za eliptičke jednadžbe u divergentnoj formi

Slabi princip maksimuma prenosi se i na diferencijalne operatore u divergentnoj formi i njihova slaba rješenja. Neka je zadan operator oblika

$$Lu = - \sum_{i=1}^d \frac{\partial}{\partial x_i} \left(\sum_{j=1}^d a_{i,j}(x) \frac{\partial u}{\partial x_j} + b_i(x)u \right) + \sum_{i=1}^d c_i(x) \frac{\partial u}{\partial x_i} + d(x)u, \quad (6.2)$$

gdje su $a_{i,j}$, b_i , c_i i d izmjerive funkcije na $\Omega \subset \mathbb{R}^d$. Ovom je operatoru pridružena bilinearna forma

$$a(u, v) = \int_{\Omega} \left[\sum_{i=1}^d \left(\sum_{j=1}^d a_{i,j}(x) \frac{\partial u}{\partial x_j} + b_i(x)u \right) \frac{\partial v}{\partial x_i} + \left(\sum_{i=1}^d c_i(x) \frac{\partial u}{\partial x_i} + d(x)u \right) v \right] dx, \quad (6.3)$$

Uočimo da operator (6.2) u odnosu na (6.1) ne samo što je u divergentnoj formi već ima i predznak minus ispred glavnog dijela, što će promijeniti smjerove nejednakosti.

Da bi bilinearna forma (6.3) bila dobro definirana i pretpostavit ćemo da su svi koeficijenti diferencijalnog operatora ograničeni:

$$a_{i,j}, b_i, c_i, d \in L^\infty(\Omega), \quad i, j = 1, \dots, d.$$

Nadalje, pretpostavljamo da je operator L strogo eliptičan u smislu da postoji konstanta $\lambda > 0$ takva da je za svako $x \in \Omega$,

$$0 < \lambda \sum_{i=1}^d \xi_i^2 \leq \sum_{i,j=1}^d a_{i,j}(x) \xi_i \xi_j, \quad \forall \vec{\xi} \in \mathbb{R}^d.$$

Uočimo nadalje da je u diferencijalnom operatoru (6.2) slobodni član jednak

$$\left[d(x) - \sum_{i=1}^d \frac{\partial b_i}{\partial x_i}(x) \right] u,$$

te da za princip maksimuma moramo imati zadovoljen uvjet

$$d(x) - \sum_{i=1}^d \frac{\partial b_i}{\partial x_i}(x) \geq 0,$$

(promjena nejednakosti je posljedica promjene predznaka glavnog člana). Taj ćemo uvjet pretpostavljati zadovoljen u slabom smislu:

$$\int_{\Omega} \left[d(x)v + \sum_{i=1}^d b_i \frac{\partial v}{\partial x_i} \right] dx \geq 0, \quad (6.4)$$

za sve test funkcije $v \in C_0^1(\Omega)$, $v \geq 0$. Budući da su koeficijenti d i b_i ograničeni argumentom gustoće dobivamo da nejednakost (6.4) vrijedi i za sve $v \in W_0^{1,1}(\Omega)$, $v \geq 0$.

Kažemo da funkcija $u \in H^1(\Omega)$ zadovoljava $Lu \leq 0$ (analogno $Lu \geq 0$) u Ω ako vrijedi $a(u, v) \leq 0$ (i analogno $a(u, v) \geq 0$) za svaku funkciju $v \in H_0^1(\Omega)$ takvu da je $v \geq 0$ na Ω . Evidentno je $Lu = 0$ u slabom smislu ako i samo ako je $Lu \leq 0$ i $Lu \geq 0$.

Za funkciju iz prostora Soboljeva nejednakost poput $u \leq k$ na $\partial\Omega$ (k je konstanta) definiramo s $\max(u - k, 0) \in H_0^1(\Omega)$; drugim riječima $(u - k)^+ \in H_0^1(\Omega)$. Analogno $u \geq k$ definiramo s $\min(u - k, 0) = (u - k)^- \in H_0^1(\Omega)$. Tu se radi o tome da tragovi funkcije zadovoljavaju te nejednakosti skoro svuda. Nadalje, supremum po granici domene definiramo formulom

$$\sup_{\partial\Omega} u = \inf\{k \in \mathbb{R} : u \leq k \text{ na } \partial\Omega\}.$$

(Taj se supremum često naziva esencijalni supremum i ponekad piše kao *esssup*.) Infimum se definira analogno, odnosno pomoću formule $\inf u = -\sup(-u)$. Sada možemo iskazati slabi princip maksimuma.

Teorem 6.3 Neka funkcija $u \in H^1(\Omega)$ zadovoljava $Lu \leq 0$ ($Lu \geq 0$) u Ω . Tada vrijedi

$$\sup_{\Omega} u \leq \sup_{\partial\Omega} u^+ \quad (\inf_{\Omega} u \geq \inf_{\partial\Omega} u^-).$$

Dokaz. Neka je $Lu \leq 0$ u slabom smislu, odnosno

$$\int_{\Omega} \left[\sum_{i=1}^d \left(\sum_{j=1}^d a_{i,j}(x) \frac{\partial u}{\partial x_j} + b_i(x)u \right) \frac{\partial v}{\partial x_i} + \left(\sum_{i=1}^d c_i(x) \frac{\partial u}{\partial x_i} + d(x)u \right) v \right] dx \leq 0,$$

za sve $v \in H_0^1(\Omega)$, $v \geq 0$. Izvršimo parcijalnu integraciju u članu s vektorom \vec{b} :

$$\int_{\Omega} \sum_{i=1}^d b_i(x)u \frac{\partial v}{\partial x_i} dx = \int_{\Omega} \sum_{i=1}^d \left(b_i(x) \frac{\partial(uv)}{\partial x_i} - b_i(x)v \frac{\partial u}{\partial x_i} \right) dx,$$

čime dobivamo, koristeći (6.4):

$$\begin{aligned} \int_{\Omega} \left[\sum_{i,j=1}^d a_{i,j}(x) \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + \sum_{i=1}^d (c_i(x) - b_i(x)) \frac{\partial u}{\partial x_i} v \right] dx \\ \leq \int_{\Omega} \left[\sum_{i=1}^d b_i(x) \frac{\partial(uv)}{\partial x_i} - d(x)uv \right] dx \leq 0. \end{aligned}$$

Posljednja nejednakost vrijedi za sve $v \geq 0$ za koje je $uv \geq 0$, pa ćemo stoga nadalje promatrati samo takve test funkcije. Takva funkcija neće postojati ukoliko je $u < 0$ skoro svuda, no u tom slučaju nemamo što dokazivati. Sada imamo,

$$\int_{\Omega} \left[\sum_{i,j=1}^d a_{i,j}(x) \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right] dx \leq \int_{\Omega} \left[\sum_{i=1}^d (b_i(x) - c_i(x)) \frac{\partial u}{\partial x_i} v \right] dx.$$

Koristeći ograničenost koeficijenata na desnoj strani i Cauchyjevu nejednakost dobivamo da postoji konstanta $C > 0$ takva da vrijedi

$$\int_{\Omega} \left[\sum_{i,j=1}^d a_{i,j}(x) \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} \right] dx \leq C \int_{\Omega} |\nabla u| v dx, \quad (6.5)$$

za sve $v \geq 0$ za koje je $uv \geq 0$.

Uvedimo oznaku $l = \sup_{\partial\Omega} u^+$ i pretpostavimo ja je $l < \sup_{\Omega} u$, jer smo u suprotnom gotovi. Tada možemo odabrati broj k koji zadovoljava $l \leq k < \sup_{\Omega} u$. Definirajmo test funkciju $v_k = (u - k)^+$.

Kako je $l \geq 0$ vrijedi $uv_k \geq 0$; nadalje $v_k \in H_0^1(\Omega)$ zbog $l \leq k$, pa imamo $uv_k \in W_0^{1,1}(\Omega)$. Imamo dakle dobru test funkciju koju možemo iskoristiti u nejednakosti (6.5). Uočimo još da vrijedi:

$$\nabla v_k = \begin{cases} \nabla u & \text{za } u > k \\ 0 & \text{za } u \leq k \end{cases}, \quad (6.6)$$

tako da kada v_k uvrstimo u (6.5) integracija se po desnom integralu vrši po skupu $A_k = \text{supp}(\nabla v_k) \subset \text{supp}(v_k) = \{x \in \Omega : u(x) > k\}$. To je stoga što imamo

$$\int_{\Omega} |\nabla u| v_k dx = \int_{\text{supp}(v_k)} |\nabla u| v_k dx = \int_{\text{supp}(v_k)} |\nabla v_k| v_k dx = \int_{A_k} |\nabla v_k| v_k dx$$

gdje u zadnjoj jednakosti uvažili da se može integrirati samo po dijelu nosača funkcije v_k na kome je gradijent različit od nule.

Nejednakost (6.5) sada prelazi u

$$\int_{\Omega} \left[\sum_{i,j=1}^d a_{i,j}(x) \frac{\partial u}{\partial x_j} \frac{\partial v_k}{\partial x_i} \right] dx \leq C \int_{A_k} |\nabla v_k| v_k dx,$$

odnosno, koristeći ponovo (6.6) izlazi

$$\int_{\Omega} \left[\sum_{i,j=1}^d a_{i,j}(x) \frac{\partial v_k}{\partial x_j} \frac{\partial v_k}{\partial x_i} \right] dx \leq C \int_{A_k} |\nabla v_k| v_k dx.$$

Sada možemo iskoristiti eliptičnost kako bismo dobili (s nekom drugom konstantom C):

$$\int_{\Omega} |\nabla v_k|^2 dx \leq C \int_{A_k} |\nabla v_k| v_k dx \leq C \|\nabla v_k\|_{L^2(\Omega)} \|v_k\|_{L^2(A_k)},$$

odnosno

$$\|\nabla v_k\|_{L^2(\Omega)} \leq C \|v_k\|_{L^2(A_k)}.$$

Na lijevoj strani ćemo iskoristiti Soboljevljev teorem ulaganja. Za $d \geq 3$ imamo neprekidno ulaganje $H^1(\Omega) \subset L^p(\Omega)$, gdje je

$$\frac{1}{p} = \frac{1}{2} - \frac{1}{d} \quad \Rightarrow \quad p = \frac{2d}{d-2}.$$

U slučaju $d = 2$ imamo neprekidno ulaganje $H^1(\Omega) \subset L^p(\Omega)$ za svako konačno p pa stoga možemo odabrati bilo koji $p > 2$. Nadalje, stoga, promatramo samo situaciju $d > 2$. Izlazi

$$\|v_k\|_{L^p(\Omega)} \leq C \|\nabla v_k\|_{L^2(\Omega)} \leq C \|v_k\|_{L^2(A_k)},$$

gdje koristimo da je $v_k \in H_0^1(\Omega)$ pa je $v \mapsto \|\nabla v\|_{L^2(\Omega)}$ norma ekvivalentna s $H^1(\Omega)$ -normom. Desnu pak stranu možemo ocijeniti na ovaj način:

$$\|v_k\|_{L^2(A_k)}^2 \leq \left(\int_{A_k} 1^q dx \right)^{1/q} \left(\int_{A_k} (|v_k|^2)^{p/2} dx \right)^{2/p} = |A_k|^{2/d} \|v_k\|_{L^p(\Omega)}^2,$$

jer je

$$\frac{1}{p/2} + \frac{1}{q} = 1 \quad \Rightarrow \quad q = \frac{p}{p-2} = \frac{d}{2}.$$

Time smo dobili

$$\|v_k\|_{L^p(\Omega)} \leq C |A_k|^{1/d} \|v_k\|_{L^p(\Omega)}.$$

Ukoliko je $v_k = 0$, onda je $u \leq k$ što je kontradikcija s $k < \sup_{\Omega} u$. U suprotnom dobivamo

$$1 \leq C |A_k|^{1/d} \quad \Rightarrow \quad |A_k| \geq C^{-d}. \quad (6.7)$$

Ocjena (6.7) vrijedi za sve $k \in [l, \sup_{\Omega} u)$. Konstanta na desnoj strani ne ovisi o k pa možemo prijeći na limes kada $k \rightarrow \sup_{\Omega} u$. Odaberimo u tu svrhu jedan podniz

$$k_1 < k_2 < \dots, \quad \lim_l k_l = \sup_{\Omega} u.$$

Primijetimo da $\sup_{\Omega} u$ može biti konačan ili beskonačan. Sada se lako uočava da je

$$A_{k_1} \supset A_{k_2} \supset A_{k_3} \supset \dots$$

i stoga imamo

$$A = \{x \in \Omega : u(x) = \sup_{\Omega} u\} = \bigcap_l A_{k_l}.$$

Taj skup, zbog neprekidnosti mjere, ima pozitivnu mjeru, $|A| = \lim |A_{k_l}| \geq C^{-d}$. Sada slijedi da $\sup_{\Omega} u$ mora biti konačan. Nadalje, tada je $\nabla u = 0$ s.s. na A , što je kontradikcija s $A \subset A_{k_l}$.

Kako smo u svakom slučaju došli na kontradikciju zaključujemo da je pretpostavka $l < \sup_{\Omega} u$ pogrešna i dokaz je gotov. Analogno se pokazuje druga nejednakost. \square

6.3 Diskretan princip maksimuma

6.3.1 Primjer: metoda konačnih diferencija

Kako bismi ilustrirali diskretan princip maksimuma uzet ćemo jedan jednostavan primjer. Promotrimo Dirichletovu zadaću za Laplaceovu jednadžbu

$$-\Delta u = f \quad \text{na } \Omega, \quad (6.8)$$

$$u = g \quad \text{na } \partial\Omega. \quad (6.9)$$

Pretpostavimo k tome da je $f \in C(\Omega)$ te $g \in C(\partial\Omega)$ i da imamo klasično rješenje $u \in C^2(\Omega) \cap C(\bar{\Omega})$.

Radi jednostavnosti diskretizacije uzimimo je $\Omega = (0, a) \times (0, b) \subset \mathbb{R}^2$ i pretpostavimo da možemo odabrati takav $h > 0$ da je

$$a = nh, \quad b = mh, \quad n, m \in \mathbb{N}.$$

(U općenitom slučaju bismo imali h_x i h_y , no ovdje radi jednostavnosti notacije pretpostavljamo da je $h = h_x = h_y$.)

Diskretizacijska mreža se definira na sljedeći način:

$$\begin{aligned} \Omega_h &= \{(ih, jh) : i = 1, \dots, n-1; j = 1, \dots, m-1\} \subset \Omega \\ \partial\Omega_h &= \{(ih, jh) : i \in \{0, n\}, j = 1, \dots, m-1; \text{ ili } i = 1, \dots, n-1; j \in \{0, m\}\}. \end{aligned}$$

Očito je $\partial\Omega_h \subset \partial\Omega$ te konačno definiramo $\bar{\Omega}_h = \Omega_h \cup \partial\Omega_h$.

Skup $\bar{\Omega}_h$ je diskretizacijska mreža. U unutarnjim točkama mreže $\bar{\Omega}_h \cap \Omega = \Omega_h$ derivacije aproksimiramo na uobičajeni način:

$$\frac{\partial^2 u}{\partial x^2}(x, y) = \frac{u(x+h, y) - 2u(x, y) + u(x-h, y)}{h^2} + h^2 R, \quad |R| \leq \frac{1}{12} \left\| \frac{\partial^4 u}{\partial x^4} \right\|_\infty,$$

gdje smo s $\|\cdot\|_\infty$ označili max-normu prostora $C(\bar{\Omega})$. Posve analogno postupamo s drugom derivacijom po y , što vodi na jednadžbu

$$\begin{aligned} & - \left(\frac{u(ih+h, jh) - 2u(ih, jh) + u(ih-h, jh)}{h^2} + \right. \\ & \left. + \frac{u(ih, jh+h) - 2u(ih, jh) + u(ih, jh-h)}{h^2} \right) = f(ih, jh) + h^2 R(ih, jh), \end{aligned}$$

za svako $(ih, jh) \in \Omega_h$. Odbacujući član $h^2 R(ih, jh)$ i zamijenjujući točnu vrijednost $u(ih, jh)$ s aproksimacijom $u_{i,j} \approx u(ih, jh)$, dobivamo metodu

$$\frac{1}{h^2} \left[-u_{i,j-1} - u_{i-1,j} + 4u_{i,j} - u_{i,j+1} - u_{i+1,j} \right] = f_{i,j}, \quad \forall (ih, jh) \in \Omega_h \quad (6.10)$$

$$u_{i,j} = g_{i,j}, \quad \forall (ih, jh) \in \partial\Omega_h, \quad (6.11)$$

pri čemu koristimo standardne oznake $f_{i,j} = f(ih, jh)$, $g_{i,j} = g(ih, jh)$.

Napomena 6.1 Uočimo karakterističan five point stencil u kome je vrijednost $u_{i,j}$ vezana s četiri susjedne vrijednosti: $u_{i,j-1}$, $u_{i-1,j}$, $u_{i,j+1}$ i $u_{i+1,j}$. U tom smislu kažemo da su točke (ih, jh) , $(ih + h, jh)$, $u(ih - h, jh)$, $(ih, jh + h)$ i $(ih, jh - h)$ susjedne. Nadalje, za točku kažemo da je blizu ruba ako ima barem jednu susjednu točku koja je na rubu. Nadalje, za $(ih, jh) \in \Omega_h$ korisno je uvesti oznaku

$$N_{i,j} = \{(i - 1, j), (i + 1, j), (i, j - 1), (i, j + 1)\}.$$

Svaku funkciju $u_h: \bar{\Omega}_h \rightarrow \mathbb{R}$ nazivamo mrežna funkcija. Kada u skup točaka $\bar{\Omega}_h$ uvedemo uređaj možemo mrežnu funkciju predstaviti vektorom. Standardan uređaj u skupu $\bar{\Omega}_h$ je leksikografski. Prvo numeriramo sve točke iz Ω_h idući “po recima”, a zatim numeriramo točke s $\partial\Omega_h$. Kako je $M_1 := \text{card}(\Omega_h) = (n - 1)(m - 1)$, te $M_2 := \text{card}(\partial\Omega_h) = 2(m + n)$, pri ovakvom indeksiranju prvih M_1 komponenti vektorskog prikaza mrežne funkcije pripada unutrašnjosti domene, a komponente s indeksima $i = M_1 + 1, \dots, M_1 + M_2$ pripadaju točkama s ruba domene.

Nadalje ćemo pretpostavljati da imamo leksikografski uređaj točaka u $\bar{\Omega}_h$ opisan više i za mrežnu funkciju u_h pripadni vektor ćemo označiti s U_h . Mala slova će, dakle, označavati mrežne funkcije, a velika vektore.

Mrežnu jednadžbu (6.10), (6.11) možemo sada zapisati u obliku sustava linearnih algebarskih jednadžbi:

$$A_h U_h = Q_h, \quad Q_h = -\hat{A}_h G_h + F_h. \quad (6.12)$$

Ovdje je U_h vektor koji odgovara unutaršnjim komponentama mrežne funkcije u_h , dakle ima M_1 komponenti. Analogno $G_h \in \mathbb{R}^{M_2}$ i $F_h \in \mathbb{R}^{M_1}$ su vektori koji odgovaraju mrežnih funkcija g_h i f_h . Kako g_h živi na $\partial\Omega_h$ vektor G_h ima samo M_2 komponenti, dok je f_h zadana samo na Ω_h pa njen vektor ima M_1 komponenti. Matrica A_h je kvadratna dimenzije M_1 , a matrica \hat{A}_h ima dimenziju $M_1 \times M_2$. Vektor $-\hat{A}_h G_h$ predstavlja Dirichletove vrijednosti koje su prebačene na desnu stranu.

Struktura matrica je, naravno, ovisna o načinu numeracije mrežnih točaka i u našem primjeru jednostavno ju je opisati. Matrica A_h je blok-trodijagonalna:

$$A_h = \frac{1}{h^2} \begin{bmatrix} T & -I & & & \\ -I & T & -I & & \\ & \ddots & \ddots & \ddots & \\ & & -I & T & -I \\ & & & -I & T \end{bmatrix},$$

gdje je $I \in \mathbb{R}^{n-1, n-1}$ identiteta i $T \in \mathbb{R}^{n-1, n-1}$ trodijagonalna matrica:

$$T = \begin{bmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{bmatrix}.$$

S druge strane, matrica $\hat{A}_h \in \mathbb{R}^{M_1 \times M_2}$ ima ovu strukturu:

$$(\hat{A}_h)_{i,j} = \begin{cases} -1/h^2 & \text{točka s indeksom } i \text{ je blizu ruba, a } j \text{ je na rubu} \\ 0 & \text{inače} \end{cases}.$$

Uvedimo još i proširenu matricu sustava

$$\tilde{A}_h = \begin{bmatrix} A_k | \hat{A}_h \end{bmatrix} \in \mathbb{R}^{M_1 \times M},$$

gdje je $M = M_1 + M_2$. Sada se lako provjerava da matrica sustava ima sljedeći niz svojstava:

1. $(A_h)_{i,i} > 0$ za sve $i = 1, \dots, M_1$;
2. $(\tilde{A}_h)_{i,j} \leq 0$ za sve $i = 1, \dots, M_1, j = 1, \dots, M$ i $i \neq j$;
3. $\sum_{j=1}^{M_1} (A_h)_{i,j} \geq 0$ za sve $i = 1, \dots, M_1$;
4. $\sum_{j=1}^{M_1} (A_h)_{i,j} > 0$ za sve i koji su blizu granice;
5. $\sum_{j=1}^{M_1} (\tilde{A}_h)_{i,j} = 0$ za sve $i = 1, \dots, M_1$;
6. A_h je ireducibilna.

Sva se ova svojstva neposredno provjeravaju osim ireducibilnosti koja je definirana u Sekciji 6.3.3. Tamo smo i pokazali na osnovu Varginog teorema da je matrica A_h ireducibilna. Dapače, prema svojstvima 3 i 4 ona je i ireducibilno dijagonalno dominantna (Definicija 6.3) te je po Posljedici 6.3 i regularna. Konačno, matrica A_h je simetrična i ima pozitivne dijagonalne elemente pa je stoga i pozitivno definitna (Posljedica 6.4).

Premda se regularnost matrice A_h mogla pokazati posve direktno, pojam ireducibilnosti i ireducibilne dijagonalne dominantnosti važni su za tretiranje složenijih i generalnijih primjera.

6.3.2 Ocjena greške metode konačnih diferencija

Greška je razlika između točnog i aproksimativnog rješenja. Ovdje ćemo ju mjeriti u max-normi. Preciznije, ako je $u(x, y)$ točno rješenje, onda definiramo mrežnu funkciju $\bar{u}(ih, jh) = u(ih, jh)$ i želimo ocijeniti

$$\|u_h - \bar{u}\|_\infty = \max_{(ih, jh) \in \Omega_h} |u_h(ih, jh) - u(ih, jh)|.$$

Naravno, moguće je odabrati i drzgačiju normu, na primjer diskretnu L^2 normu:

$$\|u_h - \bar{u}\|_{0,h} = h \left(\sum_{i=1}^{n-1} \sum_{j=1}^{m-1} |u_h(ih, jh) - u(ih, jh)|^2 \right)^{1/2}.$$

Nakon što smo izabrali neku normu $\|\cdot\|_h$ u kojoj ćemo mjeriti grešku, kažemo da je metoda konvergentna ako je

$$\lim_{h \rightarrow 0} \|u_h - \bar{u}\|_h = 0.$$

Red konvergencije je $p > 0$ ako postoji konstanta $C > 0$, neovisna o h , takva da je

$$\|u_h - \bar{u}\|_h \leq Ch^p.$$

Konzistentnost metode nam kaže da točno rješenje \bar{u} , odnosno pripadni vektor \bar{U} zadovoljava

$$A_h \bar{U} = Q_h + h^2 R_h, \quad \|R_h\|_\infty \leq \frac{1}{12} \left(\left\| \frac{\partial^4 u}{\partial x^4} \right\|_\infty + \left\| \frac{\partial^4 u}{\partial y^4} \right\|_\infty \right).$$

S druge strane imamo $A_h U_h = Q_h$ pa oduzimanjem dviju jednadžbi dobivamo $A_h(\bar{U} - U_h) = h^2 R_h$. Da bismo dobili ocjenu greške treba invertirati matricu A_h :

$$\|u_h - \bar{u}\|_\infty = \|U_h - \bar{U}\|_\infty = \|A_h^{-1} A_h(\bar{U} - U_h)\|_\infty \leq \|A_h^{-1}\|_\infty h^2 \|R_h\|_\infty.$$

Ponovo, umjesto max-norme mogli smo koristiti proizvoljnu normu $\|\cdot\|_h$. Pojam stabilnosti numeričke metode vezan je uz normu inverzne matrice A_h^{-1} . Općenito kažemo da je numerička metoda stabilna u $\|\cdot\|_h$ normi ako postoji konstanta $C > 0$, neovisna o h , takva da je

$$\|A_h^{-1}\|_h \leq C.$$

U našem primjeru moguće je pokazati, a to ćemo vidjeti kasnije u općenitijem slučaju u Sekciji 6.3.4, da je metoda stabilna u max-normi. Preciznije, vrijedi

$$\|A_h^{-1}\|_\infty \leq \frac{1}{16}(a^2 + b^2).$$

Time smo pokazali da je

$$\|u_h - \bar{u}\|_\infty \leq \frac{1}{192}(a^2 + b^2) \left(\left\| \frac{\partial^4 u}{\partial x^4} \right\|_\infty + \left\| \frac{\partial^4 u}{\partial y^4} \right\|_\infty \right) h^2. \quad (6.13)$$

Metoda je dakle drugog reda točnosti, no uočimo da je to dokazano uz neprirodno jake uvjete na glatkoću rješenja.

6.3.3 Ireducibilne matrice

Ovdje nas zanimaju svojstva reducibilnih i ireducibilnih matrica pa je prirodno promatrati samo kvadratne matrice koje se ne reduciraju na skalar, tj. pretpostaviti da je dimenzija matrice uvijek strogo veća od 1.

Definicija 6.1 Matrica $A = (a_{i,j}) \in \mathbb{R}^{n,n}$ je reducibilna ako postoji neprazan skup $S \subset N = \{1, 2, \dots, n\}$, $S \neq N$, takav da je

$$\forall i \in S, \forall j \in N \setminus S, \quad a_{i,j} = 0.$$

Matrica je ireducibilna ako nije reducibilna.

Jedan način karakterizacije reducibilne matrice dobivamo odgovarajućom permutacijom redaka i stupaca matrice. Prisjetimo se stoga da je permutacija $\sigma: \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$ svako bijektivno preslikavanje. Matrica permutacije σ , u oznaci P_σ , je matrica definirana formulom

$$(P_\sigma)_{i,j} = \delta_{\sigma(i),j},$$

gdje je δ Kronekerov delta-simbol. Matrica permutacije ima u svakom retku i svakom stupcu točno jednu jedinicu, a svi ostali elementi retka (stupca) jednaki su nuli. Nadalje, matrica djeluje na vektor tako da vrši permutaciju σ njegovih komponenti.

Svaka permutacijska matrica je ortogonalna jer imamo

$$(P_\sigma P_\sigma^T)_{i,j} = \sum_{k=1}^n (P_\sigma)_{i,k} (P_\sigma)_{j,k} = \sum_{k=1}^n \delta_{\sigma(i),k} \delta_{\sigma(j),k} = \delta_{i,j}.$$

Lako se provjeravaju sljedeća svojstva:

1. Množenje matrice A matricom permutacije P_σ s lijeva, $P_\sigma A$, permutira retke matrice A permutacijom σ ;
2. Množenje matrice A transponiranom matricom permutacije P_σ^T s desna, AP_σ^T , permutira stupce matrice A permutacijom σ ;
3. Za svake dvije permutacije σ i π vrijedi $P_{\sigma \circ \pi} = P_\pi P_\sigma$;
4. $P_{\sigma^{-1}} = P_\sigma^{-1} = P_\sigma^T$.

Lema 6.1 Matrica A je reducibilna ako i samo ako se može naći permutacijska matrica P takva da je

$$PAP^T = \begin{bmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{bmatrix} \quad (6.14)$$

gdje su $A_{1,1}$ i $A_{2,2}$ kvadratni blokovi.

Dokaz. Neka je matrica A ireducibilna i neka je $S \subset N$ skup iz definicije ireducibilnosti. Odaberimo permutaciju koja sve indekse iz S premješta na posljednja mjesta u N , tako da je $\sigma(S) = \{r+1, r+2, \dots, n\}$. Neka je P pripadna matrica permutacije; tada su u PAP^T svi reci matrice A koji odgovaraju indeksima $i \in S$ došli na posljednjih $n-r$ mjesta, dok su reci koji odgovaraju indeksima iz $N \setminus S$ došli na prvih r mjesta. Ista se transformacija desila sa

stupcima. Prema tome, na mjestima $r + 1 \leq i \leq n$ i $1 \leq j \leq r$, $(PAP^T)_{i,j} = a_{k,l}$ za neke $k \in S$ i $l \in N \setminus S$; stoga je $(PAP^T)_{i,j} = 0$. Nadalje, blok $A_{1,1}$ ima dimenziju $r \times r$, a blok $A_{2,2}$ ima dimenziju $n - r \times n - r$.

Obrat. Ako vrijedi (6.14), onda za S možemo uzeti $\sigma^{-1}(\{r+1, \dots, n\})$ i svojstvo reducibilnosti se lako provjerava. \square

Ako trebamo riješiti linearan sustav s reducibilnom matricom, onda se zadaća svodi na rješavanje dva sustava s dijagonalnim blokovima. Zaista, sustav

$$\begin{bmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} = \begin{bmatrix} b^{(1)} \\ b^{(2)} \end{bmatrix}$$

se svodi na

$$\begin{aligned} A_{2,2}x^{(2)} &= b^{(2)} \\ A_{1,1}x^{(1)} &= b^{(1)} - A_{1,2}x^{(2)}. \end{aligned}$$

S te strane je ireducibilnost poželjna no metoda konačnih elemenata (diferencija ili volumena) praktički uvijek vodi na ireducibilne matrice. Vidjet ćemo da ireducibilnost ima i pozitivnih posljedica no prvo nam treba jedan kriterij dokazivanja ireducibilnosti.

Svakoj matrici možemo pridružiti usmjereni graf. Prisjetimo se da je usmjereni graf uređeni par $G = (V, E)$ koji se sastoji od nepraznog skupa točaka V (vrhovi grafa) i skupa E uređenih parova vrhova (bridovi grafa).

Kvadratnoj matrici reda n pridružujemo usmjereni graf na sljedeći način: svakom indeksu $i \in \{1, 2, \dots, n\}$ pridružimo jednu ročku P_i , tako da $\{P_1, P_2, \dots, P_n\}$ predstavlja skup od n različitih točaka. Zatim za svaki element matrice $a_{i,j} \neq 0$ grafu pridružimo usmjereni brid (P_i, P_j) . Tako konstruirani graf ovisi samo o profilu matrice, a ne i iznosz njenih elemenata.

Lako se provjerava sljedeća tvrdnja: Ako je P matrica permutacije, onda matrice A i PAP^T imaju usmjerene grafove koji se razlikuju samo u labeliranju vrhova grafa.

Unutar usmjerenog grafa postoji usmjerena staza koja povezuje točku $P_i \in V$ s točkom $P_j \in V$ ukoliko postoje točke $P_{i_1}, \dots, P_{i_r} \in V$ takve da je

$$(P_i, P_{i_1}), (P_{i_1}, P_{i_2}), (P_{i_2}, P_{i_3}), \dots, (P_{i_r}, P_j) \in E.$$

Definicija 6.2 Usmjereni graf (V, E) je povezan ako za svaki uređen par međusobno različitih vrhova $(P_i, P_j) \in V$ postoji usmjerena staza unutar grafa koja povezuje točku P_i s P_j .

Teorem 6.4 (Varga, 1962) Matrica je ireducibilna ako i samo ako je njen usmjereni graf povezan.

Dokaz. Neka je matrica A ireducibilna i neka je i_0 proizvoljan indeks. Želimo pokazati da je on povezan usmjerenom stazom sa svakim drugim indeksom j . Stoga uvedimo skup indeksa

$$T = \{j : \text{indeks } i_0 \text{ je povezan s indeksom } j \text{ usmjerenom stazom}\}.$$

Zbog ireducibilnosti matrice skup T je neprazan. Naime, u suprotnom bi za sve $j \neq i_0$ bilo $a_{i_0,j} = 0$, pa bi uzimajući $S = \{i_0\}$ mogli pokazati da je matrica A reducibilna.

Ostaje nam pokazati da je $T = N = \{1, 2, \dots, n\}$. Pretpostavimo suprotno, da $S = N \setminus T$ neprazan skup. Uzmemo li $j \in T$ i $k \in S$ dovoljno je pokazati da je nužno $a_{j,k} = 0$ jer smo time došli do zaključka da je matrica A reducibilna, što je u suprotnosti s pretpostavkom i slijedi $T = N$.

Ostaje nam samo primijetiti da kada bi bilo $a_{j,k} \neq 0$ onda bi mogli konstruirati usmjerenu stazu od i_0 do k jer zbog $j \in T$ možemo konstruirati stazu do k , a $a_{j,k} \neq 0$ daje da je $(P_j, P_k) \in E$. No time smo došli u proturječje s pretpostavkom $j \in S$ koja povlači da takva staza ne postoji. Stoga je nužno $a_{j,k} = 0$ i tvrdnja je dokazana.

Obrat. Neka je pripadni usmjereni graf povezan. Ako bi matrica bila ireducibilna mogli bismo naći neprazan skup $S \subset N$, takav da je $T = N \setminus S$ neprazan i $a_{j,k} = 0$ za sve $j \in S$, $k \in T$. Povežimo indekse j i k usmjerenom stazom:

$$(P_j, P_{i_1}), (P_{i_1}, P_{i_2}), (P_{i_2}, P_{i_3}), \dots, (P_{i_r}, P_k) \in E.$$

Neka je i_l najveći indeks u ovom nizu za koji je $i_l \in S$, pri čemu uzimamo da je $i_0 = j$ i $i_{r+1} = k$. Očito $i_l \neq k$ te stoga postoji $i_{l+1} \in T$ takav da je $(P_{i_l}, P_{i_{l+1}}) \in E$, odnosno $a_{i_l, i_{l+1}} \neq 0$. No to je kontradikcija s definicijom skupa S i T . Time je dokazano da je matrica ireducibilna. \square

Primjer. Pokažimo da je matrica A_h iz Sekcije 6.3.1 ireducibilna. Za skup vrhova možemo uzeti točke iz Ω_h . Svaki vrh je tada evidentno povezan sa svojim susjednim točkama (vidi Napomenu 6.1) pa je evidentno da svake dvije točke možemo povezati usmjerenom stazom. \square

Prisjetimo se da je matrica A reda n strogo dijagonalno dominantna ako je

$$|a_{i,i}| > \sigma_i, \quad \sigma_i = \sum_{j \neq i} |a_{i,j}|, \quad \forall i = 1, \dots, n. \quad (6.15)$$

Ponekad se preciznije govori o strogoj dijagonalnoj dominantnosti po recima jer je isti pojam moguće formirati i po stupcima.

Stroga dijagonalna dominantnost povlači regularnost matrice. To se lako pokazuje na osnovu Gershgorinovog teorema. Uvedimo oznake

$$D_i = \{z \in \mathbb{C} : |z - a_{i,i}| \leq \sum_{j \neq i} |a_{i,j}|\},$$

$$D'_i = \{z \in \mathbb{C} : |z - a_{i,i}| \leq \sum_{j \neq i} |a_{j,i}|\}.$$

Teorem 6.5 Neka je A kvadratna matrica reda n i $\sigma(A)$ njen spektar. Tada vrijedi:

$$\sigma(A) \subset (\cup_{i=1}^n D_i) \cap (\cup_{i=1}^n D'_i).$$

Dokaz. Neka je λ jedna svojstvena vrijednost i x pripadni svojstveni vektor, normaliziran tako da je $\|x\|_\infty = 1$. Tada iz $Ax = \lambda x$ dobivamo

$$\sum_{j=1}^n a_{i,j} x_j = \lambda x_i, \quad 1 \leq i \leq n.$$

Uzmimo indeks i na kojem se postiže maksimum u vektoru x ; $|x_i| = 1$. Eventualnim množenjem vektora s -1 možemo postići $x_i = 1$. Sad imamo

$$\sum_{j \neq i}^n a_{i,j} x_j = \lambda - a_{i,i},$$

odnosno

$$|\lambda - a_{i,i}| \leq \sum_{j \neq i}^n |a_{i,j}| |x_j| \leq \sum_{j \neq i}^n |a_{i,j}| = \sigma_i,$$

gdje smo iskoristili normiranje vektora. Time smo pokazali da je $\sigma(A) \subset \cup_{i=1}^n D_i$. Budući da transponirana matrica ima isti spektar dobivamo analognu ocjenu s transponiranom matricom. \square

Sada se lako dokazuje

Posljedica 6.2 Strogo dijagonalno dominantna matrica je regularna.

Dokaz. Dovoljno je vidjeti da nula nije svojstvena vrijednost, a to slijedi iz Gershgorinovog teorema prema kome bi $0 \in \sigma(A)$ narušilo strogu dijagonalnu dominantnost. \square

Stroga dijagonalna dominantnost je u mnogim slučajevima suviše jako svojstvo.

Definicija 6.3 Kvadratna matrica A reda n je ireducibilno dijagonalno dominantna ako je ireducibilna i ako vrijedi:

$$|a_{i,i}| \geq \sum_{j \neq i} |a_{i,j}|, \quad \forall i = 1, \dots, n, \quad (6.16)$$

$$|a_{i,i}| > \sum_{j \neq i} |a_{i,j}|, \quad \text{za bar jedan } 1 \leq i \leq n. \quad (6.17)$$

Htjeli bismo pokazati da je ireducibilna dijagonalno dominantna matrica regularna, no sada definicija dijagonalne dominantnosti ne isključuje da je nula svojstvena vrijednost matrice A . Ipak, ako jeste svojstvena vrijednost ona mora biti na rubu ∂K , gdje je $K = \cup_{i=1}^n D_i$, inače bi bio narušen uvjet (6.16).

Sljedeća verzija Gershgorinovog teorema nam daje više informacija o svojstvenim vrijednostima na rubu Gershgorinovih krugova, ako je matrica ireducibilna.

Teorem 6.6 Neka je A ireducibilna kvadratna matrica reda n . Tada je svojstvena vrijednost $\lambda \in \partial K$, $K = \cup_{i=1}^n D_i$, ako i samo ako je $|\lambda - a_{i,i}| = \sigma_i$ za sve $i = 1, \dots, n$.

Dokaz. Uzmimo svojstvenu vrijednost $\lambda \in \partial K$ i neka je x pripadni svojstveni vektor koji ćemo normirati tako da bude $\|x\|_\infty = 1$. Tada možemo naći indeks r takav da je $1 = x_r \geq |x_j|$ za sve j . Kao u dokazu Teorema 6.5

$$|\lambda - a_{r,r}| \leq \sum_{j \neq r}^n |a_{r,j}| |x_j| \leq \sum_{j \neq r}^n |a_{r,j}| = \sigma_r,$$

no možemo uočiti i više: Kako je sada $\lambda \in \partial K$, to je $|\lambda - a_{i,i}| = \sigma_i$ pa je

$$\sum_{j \neq i}^n |a_{i,j}| |x_j| = \sum_{j \neq i}^n |a_{i,j}|.$$

To znači da je za svako j takvo da je $a_{i,j} \neq 0$, $|x_j| = 1$. Zbog ireducibilnosti matrice A postoji barem jedan $j \neq i$ takav da je $a_{i,j} \neq 0$; označimo ga s r_1 . Tada kao gore pokazujemo da je $|\lambda - a_{r_1,r_1}| = \sigma_{r_1}$. Postupak dalje nastavljamo nalazeći indeks $r_2 \neq r_1$ takav da je $a_{r_1,r_2} \neq 0$.

Neka je sada j proizvoljan indeks. Prema ireducibilnosti i Varginom teoremu možemo naći usmjerenu stazu u grafu matrice od indeksa i do indeksa j ,

$$(P_i, P_{r_1}), (P_{r_1}, P_{r_2}), \dots, (P_{r_l}, P_j) \in E,$$

odnosno takvu da je

$$a_{i,r_1} \neq 0, a_{r_1,r_2} \neq 0, \dots, a_{r_l,j} \neq 0,$$

i stoga dobivamo $|\lambda - a_{j,j}| = \sigma_j$. Time je tvrdnja dokazana. \square

Posljedica 6.3 Ireducibilno dijagonalno dominantna matrica je regularna.

Dokaz. Ako je nula svojstvena vrijednost matrice, onda je $0 \in \partial K$. Prema prethodnom teoremu $|a_{i,i}| = \sigma_i$ za sve $1 \leq i \leq n$, što je u suprotnosti s (6.17). \square

Posljedica 6.4 Ako je matrica A simetrična ireducibilno (ili strogo) dijagonalno dominantna s pozitivnom dijagonalom, tada je ona pozitivno definitna.

Dokaz. Zbog simetrije sve svojstvene vrijednosti su realne, a zbog pozitivnosti dijagonalnih elemenata one su pozitivne. Stroga pozitivnost slijedi iz toga što nula nije svojstvena vrijednost. \square

6.3.4 Diskretan princip maksimuma

Vratimo se našem primjeru Laplaceove jednadžbe (6.10), (6.11). Uvedimo oznaku

$$\partial\Omega_h^* = \partial\Omega_h \setminus \{(0, 0), (a, 0), (0, b), (a, b)\}.$$

Sada možemo iskazati jako princip maksimuma.

Teorem 6.7 Neka je u_h mrežna funkcija definirana na $\bar{\Omega}_h$ s (6.10) i (6.11) te neka je f_h mrežna funkcija koja zadovoljava $f_h \leq 0$ na Ω_h . Ako funkcija u_h postiže svoj maksimum na $\Omega_h \cup \partial\Omega_h^*$ u nekoj točki iz Ω_h , tada je u_h konstanta na $\Omega_h \cup \partial\Omega_h^*$. Posebno vrijedi

$$\max_{\Omega_h} u_h \leq \max_{\partial\Omega_h^*} u_h.$$

Dokaz. Pretpostavimo da se maksimum funkcija u_h dostiže u nekoj točki $(i_0h, j_0h) \in \Omega_h$ i da je jednak $\bar{u} = u_h(i_0h, j_0h)$. Zbog $f_{i,j} \leq 0$ imamo

$$4\bar{u} \leq \sum_{(i,j) \in N_{i_0,j_0}} u_{i,j} \leq 4\bar{u}.$$

gdje smo iskoristili oznaku N_{i_0,j_0} iz Napomene 6.1. Iz dobivenog slijedi da su sve četiri vrijednosti jednake, onda se ponavljanjem argumenta lako dobiva da je mrežna funkcija konstantna svugdje osim u vrhovima pravokutnika $\Omega = (0, a) \times (0, b)$. rhovi pravokutnika nisu uključeni stoga što nisu vezani s drugim vrijednostima, odnosno ne ulaze u jednadžbe. \square

Zadatak 6.1 *Dokažite odgovarajuću tvrdnju za $f_h \geq 0$.*

Sada prelazimo na generalizaciju tog rezultata na općenitiju situaciju. Neka je zadana diskretna jednadžba

$$A_h U_h = Q_h, \quad Q_h = -\hat{A}_h G_h + F_h, \quad (6.18)$$

gdje je $A_h \in \mathbb{R}^{M_1 \times M_1}$, $\hat{A}_h \in \mathbb{R}^{M_1 \times M_2}$.

Ovdje je U_h vektor koji odgovara unutarnjim komponentama mrežne funkcije u_h , $G_h \in \mathbb{R}^{M_2}$ i $F_h \in \mathbb{R}^{M_1}$. Vektor G_h odgovara Dirichletovim rubnim vrijednosti za koje pretpostavljamo da postoja ($M_1 \geq 1$), dok F_h odgovara funkciji desne strane.

Uvedimo još i proširenu matricu sustava i prošireni vektor rješenja

$$\tilde{A}_h = \begin{bmatrix} A_h & \hat{A}_h \end{bmatrix} \in \mathbb{R}^{M_1 \times M}, \quad \tilde{U}_h = \begin{bmatrix} U_h \\ G_h \end{bmatrix},$$

gdje je $M = M_1 + M_2$ i pretpostavimo da su zadovoljena sljedeća svojstva:

- (M1) $(A_h)_{i,i} > 0$ za sve $1 \leq i \leq M_1$;
- (M2) $(A_h)_{i,j} \leq 0$ za sve $1 \leq i, j \leq M_1$ i $i \neq j$;
- (M3) $\sum_{j=1}^{M_1} (A_h)_{i,j} \geq 0$ za sve $1 \leq i \leq M_1$;
- (M4) $\sum_{j=1}^{M_1} (A_h)_{i,j} > 0$ za barem jedan $1 \leq i \leq M_1$;
- (M5) A_h je ireducibilna;
- (M6) $(\hat{A}_h)_{i,j} \leq 0$ za sve $1 \leq i \leq M_1$ i $M_1 + 1 \leq j \leq M$;
- (M7) $\sum_{j=1}^M (\tilde{A}_h)_{i,j} \geq 0$ za sve $1 \leq i \leq M_1$;
- (M8) Za svako $M_1 + 1 \leq j \leq M$ postoji bar jedno $1 \leq i \leq M_1$ takvo da je $(\hat{A}_h)_{i,j} \neq 0$.

Napomena 6.2 *U (M8) pretpostavljamo da smo eliminirali one rubne vrijednosti koje nisu povezane s drugim funkcijskim vrijednostima, kao što su to bili vrhovi pravokutnika u primjeru s početka.*

Napomena 6.3 *Generalizirat ćemo relaciju biti susjed na sljedeći način: Indeksi $1 \leq i \leq M_1$ i $1 \leq j \leq M$ su susjedni ako je $(\tilde{A}_h)_{i,j} \neq 0$. Nadalje ćemo za svaki $1 \leq i \leq M_1$ s N_i označiti skup svih susjednih indeksa.*

Uočimo da (M2) i (M6) možemo zapisati u obliku

$$(M2)' \quad (\tilde{A}_h)_{i,j} \leq 0 \text{ za sve } 1 \leq i \leq M_1, 1 \leq j \leq M \text{ i } i \neq j;$$

i da tada (M7) prima oblik

$$(M7) \quad (A_h)_{i,i} \geq \sum_{j \in N_i} |(\tilde{A}_h)_{i,j}| \text{ za sve } 1 \leq i \leq M_1.$$

Napomena 6.4 *Ireducibilnost matrice A_h možemo formulirati na ovaj način: Svaka dva indeksa $1 \leq i \leq M_1$ i $1 \leq j \leq M$ možemo povezati lancem susjednih indeksa.*

Pokazuje se korisnim uvesti parcijalni uređaj u prostore matrica i vektora po komponentama. Tako ćemo pisati $F_h \leq 0$ umjesto $(F_h)_i \leq 0$ za sve $1 \leq i \leq M_1$ $A_h \geq 0$ umjesto $(A_h)_{i,j} \geq 0$ za sve $1 \leq i, j \leq M_1$ i posve analogno za sve ostale relacije.

Sada možemo dokazati, u poopćenoj situaciji, jaki princip maksimuma:

Teorem 6.8 *Neka su ispunjeni uvjeti (M1)–(M8) i neka je $F_h \leq 0$ i neka je U_h rješenje sustava (6.18). Ako prošireni vektor rješenja \tilde{U}_h prima nenegativnu maksimalnu vrijednost za neki indeks $1 \leq i \leq M_1$, tada su sve komponente vektora \tilde{U}_h jednake. Posebno vrijedi*

$$\max_{1 \leq i \leq M} (\tilde{U}_h)_i \leq \max \left\{ 0, \max_{M_1+1 \leq j \leq M} (\hat{U}_h)_j \right\}.$$

Dokaz. Pretpostavimo da se maksimum proširenog vektora postiže na indeksu $1 \leq i \leq M_1$ i neka je $\bar{u} = (U_h)_i \geq 0$. Koristeći oznaku N_i iz Napomene 6.3 možemo pisati:

$$\begin{aligned} (A_h)_{i,i} \bar{u} &= - \sum_{N_i} (\tilde{A}_h)_{i,j} (\tilde{U}_h)_j + (F_h)_i \\ &\leq - \sum_{N_i} (\tilde{A}_h)_{i,j} (\tilde{U}_h)_j = \sum_{N_i} |(\tilde{A}_h)_{i,j}| (\tilde{U}_h)_j \\ &\leq \sum_{N_i} |(\tilde{A}_h)_{i,j}| \bar{u} \leq (A_h)_{i,i} \bar{u}, \end{aligned}$$

gdje smo koristili (M2), (M6) i (M7), a u zadnjem koraku i pretpostavku $\bar{u} \geq 0$. Zbog (M1) je $(A_h)_{i,i} \neq 0$ pa svugdje moraju vrijediti jednakosti. Nadalje, kako je po definiciji skupa N_i (Napomena 6.3) $(\tilde{A}_h)_{i,j} \neq 0$ to imamo da je $(\tilde{U}_h)_j = \bar{u}$ za sve $j \in N_i$. Zbog ireducibilnosti

matrice A_h (svojstvo (M5)) svaka dva indeksa $1 \leq i, j \leq M_1$ možemo povezati lancem međusobno susjednih indeksa i stoga je $(\tilde{U}_h)_j = \bar{u}$ za sve $1 \leq j \leq M_1$. Onda su i sve točke na granici, koje su povezane s točkama u unutrašnjosti međusobno jednake, tj. prema (M8) slijedi da je $(\tilde{U}_h)_i = \bar{u}$ za sve $1 \leq i \leq M$. \square

Zadatak 6.2 *Dokažite odgovarajuću tvrdnju za $F_h \geq 0$.*

Svojstvo ireducibilnosti matrice A_h možemo oslabiti ako je suma elemenata svako retka proširene matrice jednaka nuli. Imamo ovaj skup uvjeta:

(N1) $(A_h)_{i,i} > 0$ za sve $1 \leq i \leq M_1$;

(N2) $(A_h)_{i,j} \leq 0$ za sve $1 \leq i, j \leq M_1$ i $i \neq j$;

(N3) $\sum_{j=1}^{M_1} (A_h)_{i,j} \geq 0$ za sve $1 \leq i \leq M_1$;

(N4) Za svaki indeks $1 \leq i_1 \leq M_1$ za koji je suma u (N3) jednaka nuli postoji niz indeksa i_2, \dots, i_{l+1} takvih da je $(A_h)_{i_k, i_{k+1}} \neq 0$ za $k = 1, 2, \dots, l$ te da vrijedi

$$\sum_{j=1}^{M_1} (A_h)_{i_{l+1}, j} > 0;$$

(N5) $(\hat{A}_h)_{i,j} \leq 0$ za sve $1 \leq i \leq M_1$ i $M_1 + 1 \leq j \leq M$;

(N6) $\sum_{j=1}^M (\tilde{A}_h)_{i,j} = 0$ za sve $1 \leq i \leq M_1$;

Uočimo da je uvjet (M4), koji se više ne pojavljuje eksplicitno u (N1)-(N6), sadržan u (N4).

Napomena 6.5 *Motivirani primjerom Laplaceove jednadžbe uvest ćemo ovakvu terminologiju: Za indeks $1 \leq i \leq M_1$ kažemo da je daleko od granice ako je $\sum_{j=1}^{M_1} (A_h)_{i,j} = 0$. Indeks je blizu granice ako je $e \sum_{j=1}^{M_1} (A_h)_{i,j} > 0$, dok su indeksi $i > M_1$ na granici. Svojstvo (N5) sada kaže da se svaki indeks koji je daleko od granice može povezati lancem susjednih indeksa s točkom koja je blizu granice. Svaka točka blizu granice susjedna je bar jednom točkom na granici (zbog (N7)).*

Sada možemo dokazati ovaj slabi princip maksimuma:

Teorem 6.9 Neka su ispunjeni uvjeti (N1)–(N6) i neka je $F_h \leq 0$ i neka je U_h rješenje sustava (6.18). Tada je

$$\max_{1 \leq i \leq M} (\tilde{U}_h)_i \leq \max_{M_1+1 \leq j \leq M} (\hat{U}_h)_j.$$

Dokaz. Pretpostavimo da se maksimum proširenog vektora postiže na indeksu $1 \leq i \leq M_1$ i neka je $\bar{u} = (U_h)_i$. Na isti način kao i u Teoremu 6.8 dobivamo

$$\begin{aligned} (A_h)_{i,i}\bar{u} &\leq - \sum_{N_i} (\tilde{A}_h)_{i,j}(\tilde{U}_h)_j = \sum_{N_i} |(\tilde{A}_h)_{i,j}|(\tilde{U}_h)_j \\ &\leq \sum_{N_i} |(\tilde{A}_h)_{i,j}|\bar{u} = (A_h)_{i,i}\bar{u}, \end{aligned}$$

gdje smo u zadnjem koraku iskoristili (N7) te imamo jednakost umjesto nejednakosti i pretpostavka pozitivnosti maksimuma \bar{u} nam nije potrebna. Kao u Teoremu 6.8 zaključujemo da je vrijednost rješenja u svim susjelim indeksima također jednaka \bar{u} . Kako se svaka takva točka može lancem susjeda povezati s točkom koja je blizu granice, ona se može povezati i s točkama na grabici, što dokazuje tvrdnju. \square

Zadatak 6.3 Dokažite odgovarajuću tvrdnju za $F_h \geq 0$.

Sada lako slijedi princip uspoređivanja:

Lema 6.2 Pretpostavimo da vrijede uvjeti (M1)–(M8) ili (N1)–(N6). Neka su $U_h^1, U_h^2 \in \mathbb{R}^{M_1}$ rješenja sustava

$$A_h U_h^i = -\hat{A}_h G_h^i + F_h^i, \quad i = 1, 2,$$

gdje $G_h^1, G_h^2 \in \mathbb{R}^{M_2}$ te $F_h^1, F_h^2 \in \mathbb{R}^{M_1}$ zadovoljavaju $G_h^1 \leq G_h^2$ i $F_h^1 \leq F_h^2$. Tada je

$$U_h^1 \leq U_h^2.$$

Dokaz. Razlika rješenja $V_h = U_h^1 - U_h^2$ zadovoljava

$$A_h V_h = -\hat{A}_h (G_h^1 - G_h^2) + F_h^1 - F_h^2.$$

Prema Teoremu 6.8 ili Teoremu 6.9, te zbog $F_h^1 - F_h^2 \leq 0$ i $G_h^1 - G_h^2 \leq 0$ dobivamo

$$\max_{1 \leq i \leq M} (V_h)_i \leq 0,$$

odakle slijedi $U_h^1 \leq U_h^2$. \square

Lema 6.3 Neka je $A_h \in \mathbb{R}^{M_1 \times M_1}$ matrica sa svojstvima (N1), (N2), (N3) i (N4). Tada za svako $U_h \in \mathbb{R}^{M_1}$ vrijedi

$$A_h U_h \geq 0 \quad \text{povlači} \quad U_h \geq 0.$$

Dokaz. Želimo primijeniti Lemu 6.2 s $U_h^1 = 0$, $U_h^2 = U_h$, $G_h^1 = G_h^2 = 0$, $F_h^1 = 0$ i $F_h^2 = A_h U_h$. Pri tome trebamo definirati matricu \hat{A}_h zadovoljimo (N5) i (N6). To je evidentno moguće, i štoviše, odabrana vrijednost nema nikakvog utjecaja na primjenu Leme 6.2. Kako su sve pretpostavke leme zadovoljene zaključujemo da je $U_h^2 \geq U_h^1$, odnosno $U_h \geq 0$. \square

U Lemi 6.3 smo pokazali da je matrica koja zadovoljava (N1)–(N4) monotona.

Definicija 6.4 Matrica A je monotona ako $Ax \geq 0$ implicira $x \geq 0$.

Lema 6.4 Kvadratna matrica A je monotona ako i samo ako je regularna i $A^{-1} \geq 0$.

Dokaz. Neka je matrica A monotona. Ako je $Ax = 0$ onda je $x \geq 0$. Uzimajući $-x$ zaključujemo da je i $x \leq 0$, pa je $x = 0$. Time je regularnost dokazana. Nadalje, ako je e_i i -ti vektor kanonske baze, onda iz $e_i = A(A^{-1}e_i) \geq 0$ slijedi $A^{-1}e_i \geq 0$, odakle se vidi da su svi elementi inverzne matrice pozitivni.

Obrat, ako je A regularna i $A^{-1} \geq 0$, onda iz $Ax \geq 0$ neposredno izlazi $x = A^{-1}(Ax) \geq 0$. \square

Definicija 6.5 M-matrica A je monotona sa svojstvima $a_{i,i} > 0$ i $a_{i,j} \leq 0$ za $i \neq j$.

Dakle, svaka monotona matrica A_h koja zadovoljava uvjete (N1) i (N2) je M-matrica. Sada možemo dokazati drugu verziju principa uspoređivanja.

Posljedica 6.5 Neka je $A_h \in \mathbb{R}^{M_1 \times M_1}$ monotona matrica i neka vrijedi (N5). Neka su $U_h^1, U_h^2 \in \mathbb{R}^{M_1}$ rješenja sustava

$$A_h U_h^i = -\hat{A}_h G_h^i + F_h^i, \quad i = 1, 2,$$

gdje $G_h^1, G_h^2 \in \mathbb{R}^{M_2}$ te $F_h^1, F_h^2 \in \mathbb{R}^{M_1}$ zadovoljavaju $G_h^1 \leq G_h^2$ i $F_h^1 \leq F_h^2$. Tada je

$$U_h^1 \leq U_h^2.$$

Dokaz. Jednadžbu

$$A_h(U_h^1 - U_h^2) = -\hat{A}_h(G_h^1 - G_h^2) + F_h^1 - F_h^2$$

pomnožimo s inverznom matricom:

$$(U_h^1 - U_h^2) = A_h^{-1}(-\hat{A}_h)(G_h^1 - G_h^2) + A_h^{-1}(F_h^1 - F_h^2).$$

Tvrdnja sada slijedi iz monotonosti. \square

Izvedimo sada apriorne ocjene na bazi principa maksimuma. U tu svrhu ćemo koristiti oznaku $\mathbf{1}$ za vektor odgovarajuće dimenzije čije su sve komponente jednake 1. Nadalje, koristimo oznaku $|\cdot|_\infty$ za max-normu funkcije, mrežne finkcije ili vektora.

Teorem 6.10 Neka su ispunjene pretpostavke (N1)-(N5) i neka postoje vektori $W_h^1, W_h^2 \in \mathbb{R}^{M_1}$ za koje vrijedi

$$A_h W_h^1 \geq \mathbf{1}, \quad A_h W_h^2 \geq -\hat{A}_h \mathbf{1}. \quad (6.19)$$

Tada rješenje jednadžbe $A_h U_h = -\hat{A}_h G_h + F_h$ zadovoljava

$$-(|F_h|_\infty W_h^1 + |G_h|_\infty W_h^2) \leq U_h \leq |F_h|_\infty W_h^1 + |G_h|_\infty W_h^2 \quad (6.20)$$

$$|U_h|_\infty \leq |F_h|_\infty |W_h^1|_\infty + |G_h|_\infty |W_h^2|_\infty. \quad (6.21)$$

Ako vrijedi (N1)-(N4) i (6.19), onda je ($\|\cdot\|_\infty$ je matična norma koju generira $|\cdot|_\infty$ norma)

$$\|A_h^{-1}\|_\infty \leq |W_h^1|_\infty. \quad (6.22)$$

Dokaz. Uočimo prvo da je

$$\begin{aligned} -|F_h|_\infty \mathbf{1} &\leq F_h \leq |F_h|_\infty \mathbf{1}, \\ -|G_h|_\infty \mathbf{1} &\leq G_h \leq |G_h|_\infty \mathbf{1}. \end{aligned}$$

Uvedimo vektor $V_h = |F_h|_\infty W_h^1 + |G_h|_\infty W_h^2 - U_h$ i uočimo da je zbog (6.19)

$$\begin{aligned} A_h V_h &= |F_h|_\infty A_h W_h^1 + |G_h|_\infty A_h W_h^2 - A_h U_h \\ &\geq |F_h|_\infty \mathbf{1} - |G_h|_\infty \hat{A}_h \mathbf{1} + \hat{A}_h G_h - F_h \\ &= (|F_h|_\infty \mathbf{1} - F_h) - \hat{A}_h (|G_h|_\infty \mathbf{1} - G_h) \geq 0. \end{aligned}$$

Ovdje smo iskoristili $-\hat{A}_h \geq 0$, što izlazi iz (N5). Prema Lemi 6.3 matrica A_h je monotona, pa dobivamo $V_h \geq 0$. Time je dokazana jedna nejednakost u (6.20). Druga se dokazuje analogno. Uzimajući maksimum dobivamo (6.21). Za dokaz nejednakosti (6.22) nam ne treba pretpostavka (N5) koja se odnosu na matricu \hat{A}_h . U tom slučaju uzimamo naprosto da je $G_h = 0$, $F_h = A_h U_h$ i iz dokazanog dobivamo

$$|U_h|_\infty \leq |A_h U_h|_\infty |W_h^1|_\infty.$$

Kako to sada vrijedi za proizvoljan vektor U_h izlazi

$$|A_h^{-1} U_h|_\infty \leq |U_h|_\infty |W_h^1|_\infty,$$

i odavde izlazi (6.22). \square

Zadatak 6.4 *Dovršite dokaz Teorema 6.10.*

Iz Teorema 6.10 vidimo da ako želimo dokazati stabilnost metode moramo naći vektor W_h takav da je

$$A_h W_h \geq \mathbf{1}, \quad |W_h|_\infty \leq C,$$

gdje C ne ovisi o h .

Primijetimo da vektor W_h zbog monotonosti matrice mora imati pozitivne koordinate.

Vratimo se sada na naš primjer diskretizacije konačnim diferencijama Laplaceove jednadžbe na pravokutniku. Da bismo konstruirali traženi vektor prvo konstruiramo funkciju

$$w(x, y) = \frac{1}{4}x(a-x) + \frac{1}{4}y(b-y),$$

koja zadovoljava

$$\begin{aligned} -\Delta w &= 1, & \text{na } \Omega \\ w &\geq 0, & \text{na } \partial\Omega \end{aligned}$$

Budući da je funkcija w polinom drugog stupnja (u separiranoj formi) to vidimo da vektor W_h koji se dobije tako da se uzmu vrijednosti funkcije w u mrežnim točkama zadovoljava $\tilde{A}_h W_h = \mathbf{1}$, odnosno

$$A_h W_h = -\hat{A}_h \tilde{W}_h + \mathbf{1} \geq \mathbf{1},$$

gdje smo s \hat{W}_h označili vektor koji se dobiva iz vrijednosti funkcije w u točkama iz $\partial\Omega_h^*$. Nadalje,

$$|w|_\infty = \frac{1}{16}(a^2 + b^2).$$

Time je stabilnost dokazana.

Pogledajmo sada kako bismo tu tehniku primijenili na proizvoljnu rubnu zadaću oblika:

$$Lu = f, \quad \text{na } \Omega \quad (6.23)$$

$$u = g, \quad \text{na } \partial\Omega. \quad (6.24)$$

Trebali bismo naći rješenje w te zadaće za odabrane $f \geq 1$ i $g \geq 0$. Zatim bismo za vektor W_h uzeli vrijednosti funkcije w u točkama mreže. Evidentno je $|W_h|_\infty = \|w\|_\infty$.

Pretpostavimo da je dana diskretizacija zadaće (6.23), (6.24) oblika $A_h U_h = -\hat{A}_h \hat{W}_h + F_h$ koja je konzistentna s jednadžbom u smislu

$$T_h = A_h W_h + \hat{A}_h \hat{W}_h - F_h, \quad |T_h|_\infty \rightarrow 0 \quad \text{kada } h \rightarrow 0.$$

Uočimo da je u našem prethodnom primjeru diskretna jednadžba bila egzaktno zadovoljena ($T_h = 0$) no općenito to nije moguće postići. Sada imamo

$$A_h W_h = T_h - \hat{A}_h \hat{W}_h + F_h \quad (6.25)$$

Sada iz $f \geq 1$ i $G \geq 0$ dobivamo da postoji konstanta $\alpha > 0$ takva da je

$$-\hat{A}_h \hat{W}_h + F_h \geq \alpha \mathbf{1},$$

no dovoljno nam je da to vrijedi samo za dovoljno male $h > 0$. Sada iz (6.25) dobivamo

$$A_h W_h \geq -|T_h|_\infty \mathbf{1} + \alpha \mathbf{1},$$

što daje $A_h W_h \geq (\alpha/2)\mathbf{1}$ za sve h dovoljno male. Time smo dobili da je traženi vektor $2W_h/\alpha$, a konstanta u ocjeni stabilnosti $C = 2\|w\|_\infty/\alpha$.

6.4 Primjena na metodu konačnih elemenata

Diskretan princip maksimuma dobro je proučen u slučaju eliptičke jednadžbe u dvije prostorne varijable ($d = 2$) i konformnih konačnih elemenata prvog reda (P_1 elementi). Uvjeti koji se dobivaju za elemente višeg reda i u trodimenzionalnom prostoru puno su kompleksniji od onih koje ćemo izvesti za $d = 2$ za linearne elemente, ili u nekim situacijama, diskretan princip maksimuma uopće ne vrijedi.

Vratimo se ponovo na primjer Dirichletove zadaće za Laplaceovu jednadžbu

$$-\Delta u = f, \quad \text{na } \Omega \quad (6.26)$$

$$u = g, \quad \text{na } \partial\Omega, \quad (6.27)$$

gdje je $\Omega \subset \mathbb{R}^2$ poligonalna ograničena domena. Zadaću (6.26), (6.27) diskretiziramo linearnim konformnim elementima na triangulaciji sastavljenoj od trokuta. Diskretan sustav će ponovo biti oblika

$$A_h U_h = -\hat{A}_h G_h + F_h,$$

gdje su unutarnji vrhovi triangulacije numerirani od 1 do M_1 , a Dirichletove su točke (vrhovi na $\partial\Omega$) numerirane od $M_1 + 1$ do $M = M_1 + M_2$. Diskretizirane vrijednosti rubnog uvjeta i funkcije desne strane označene su s $G_h \in \mathbb{R}^{M_2}$ i $F_h \in \mathbb{R}^{M_1}$, dok je $A_h \in \mathbb{R}^{M_1 \times M_1}$ i $\hat{A}_h \in \mathbb{R}^{M_1 \times M_2}$.

Elementi proširene matrice sustava

$$\tilde{A}_h = \begin{bmatrix} A_h & \hat{A}_h \end{bmatrix} \in \mathbb{R}^{M_1 \times M},$$

dani su formulom

$$(\tilde{A}_h)_{i,j} = \int_{\Omega} \nabla \phi_j \cdot \nabla \phi_i \, dx = \int_{\text{supp}\phi_i \cap \text{supp}\phi_j} \nabla \phi_j \cdot \nabla \phi_i \, dx$$

gdje je ϕ_i bazna funkcija pridružena nodalnoj točki (vrhu trokuta) s indeksom i . Ako je $i \neq j$ i ako je bar jedna nodalna točka unutarnja onda se presjek nosača dviju baznih funkcija sastoji od dva trokuta, dok se u slučaju da su obje nodalne točke na rubu domene sastoji od samo jednog trokuta.

Lema 6.5 Neka je \mathcal{T}_h konformna triangulacija poligonalne domene $\Omega \subset \mathbb{R}^2$ sastavljena od trokuta. Neka je $K \in \mathcal{T}_h$ trokut i neka su $a_i, a_j \in K$ dva različita vrha trokuta. Označimo s $\alpha_{i,j}^K$ unutarnji vrh trokuta koji je nasuprotan bridu $[a_i, a_j]$. Tada je

$$\int_K \nabla \phi_j \cdot \nabla \phi_i \, dx = -\frac{1}{2} \cot \alpha_{i,j}^K.$$

Dokaz. Treći vrh u trokutu K označimo s a_k , tako da su sva tri vrha trokuta a_i, a_j, a_k . Bazna se funkcija ϕ_j poništava na stranici trokuta koja je nasuprotna vrhu a_j (u kojem je $\phi_j(a_j) = 1$). To znači da je derivacija u smjeru nasuprotne stranice trokuta jednaka nuli i prema tome čitav gradijent funkcije ima smjer normale na nasuprotnu stranicu. To možemo zapisati u obliku

$$\nabla \phi_j = -|\nabla \phi_j| \mathbf{n}^j \quad (6.28)$$

gdje je \mathbf{n}^j jedinična vanjska normala (gleda izvan trokuta) na stranicu nasuprotnu vrhu a_j (stranicu razapetu s a_k i a_i). Predznak minus dolazi odatle što funkcija ϕ_j raste u suprotnom smjeru od \mathbf{n}^j . Iz

$$-|\nabla \phi_j| = \nabla \phi_j \cdot \mathbf{n}^j = \frac{0 - 1}{h_j} = -\frac{1}{h_j} \quad (6.29)$$

gdje smo s h_j označili visinu trokuta povučenu iz vrha a_j . Stoga je

$$\nabla \phi_j = -\frac{1}{h_j} \mathbf{n}^j,$$

pa imamo

$$\nabla \phi_j \cdot \nabla \phi_i = \frac{1}{h_j h_i} \mathbf{n}^j \cdot \mathbf{n}^i = -\frac{1}{h_j h_i} \cos \alpha_{i,j}^K.$$

Konačno, trebamo izbaciti visine iz izraza, pomoću izraza za površinu trokuta:

$$2|K| = h_j |a_k - a_i| = h_i |a_k - a_j| = |a_k - a_i| |a_k - a_j| \sin \alpha_{i,j}^K,$$

gdje smo u zadnjem koraku iskoristili sinusov poučak. Sada je

$$\frac{1}{h_j h_i} = \frac{|a_k - a_i| |a_k - a_j|}{4|K|^2} = \frac{1}{2|K| \sin \alpha_{i,j}^K}.$$

Time je konačno

$$\nabla \phi_j \cdot \nabla \phi_i = -\frac{\cot \alpha_{i,j}^K}{2|K|}$$

i tvrdnja slijedi integracijom. \square

Posljedica 6.6 Neka su K_1 i K_2 dva susjedna trokuta triangulacije \mathcal{T}_h koji imaju zajedničku stranicu s vrhovima a_i i a_j . Tada je

$$\int_{K_1 \cup K_2} \nabla \phi_j \cdot \nabla \phi_i dx = -\frac{1}{2} \frac{\sin(\alpha_{i,j}^{K_1} + \alpha_{i,j}^{K_2})}{\sin(\alpha_{i,j}^{K_1}) \sin(\alpha_{i,j}^{K_2})}.$$

Dokaz. Primjenom adicione formule za kotangens. \square

Pogledajmo sada pod kojim uvjetima su ispunjeni uvjeti (N1)–(N6). Prvo, lako uočavamo da je (N1) zadovoljen jer je

$$(A_h)_{i,i} = \int_{\Omega} |\nabla \phi_i|^2 dx > 0.$$

Nadalje, budući da je $\sum \phi_j = 1$ na $\bar{\Omega}$ slijedi

$$\sum_{j=1}^M (\tilde{A}_h)_{i,j} = \sum_{j=1}^M \int_{\Omega} \nabla \phi_j \cdot \nabla \phi_i dx = \int_{\Omega} \nabla 1 \cdot \nabla \phi_i dx = 0,$$

i time je dokazano (N6). Da bismo zadovoljili (N2) i (N5) moramo na triangulaciju postaviti sljedeći uvjet:

- (Uvjet na kuteve) Za svaki par trokuta u \mathcal{T}_h koji imaju zajedničku stranicu suma unutarnjih kuteva nasuprotnih toj stranici mora biti manja ili jednaka π . Ako trokut ima stranicu na granici koja nije Dirichletova, onda nasuprotni kut mora biti manji ili jednak $\pi/2$.

Ako je ispunjen uvjet na kuteve onda prema Lemi 6.5 i Posljedici 6.6 slijedi (N2) i (N5). Naime, svaki par indeksa (i, j) , $i \neq j$, odgovara matičnom elementu $\tilde{A}_{i,j} = 0$ ili odgovara paru vrhova (a_i, a_j) koji čine stranicu jednog trokuta. Tu stranicu mogu dijeliti dva trokuta ili samo jedan. Ako ju dijele dva trokuta, recimo K_1 i K_2 , onda za negativnost elementa, $\tilde{A}_{i,j} \leq 0$, prema Posljedici 6.6 mora vrijediti $\alpha_{i,j}^{K_1} + \alpha_{i,j}^{K_2} \leq \pi$. Ako je stranica pala na granicu domene, onda se može nalaziti u jednom jedinom elementu, recimo K , i tada prema Lemi 6.5 mora vrijediti $\alpha_{i,j}^K \leq \pi/2$.

Sada, naravno, jednostavno slijedi i (N3) jer je zbog (N6)

$$\sum_{j=1}^{M_1} (A_h)_{i,j} = - \sum_{j=M_1+1}^M (\hat{A}_h)_{i,j} \geq 0, \quad \text{za } 1 \leq i \leq M_1.$$

Težak dio koji ostaje za dokazati je (N4). Za njega nam treba sljedeća pretpostavka

- (Uvjet povezanosti) Svaka dva vrha u triangulaciji koji ne leže na Dirichletovoj granici mogu se povezati poligonalnom linijom koja je sastavljena od stranica trokuta koje ne leže na Dirichletovoj granici.

Mi nećemo ići u dokaz (N4) na osnovu uvjeta na kuteve i uvjeta povezanosti već ćemo samo primijetiti da ako uvjet na kuteve osnažimo tako da se u njemu traži stroga nejednakost, onda je u (N3) stroga nejednakost za svaki indeks i blizu granice. Na osnovu toga i uvjeta povezanosti lako se dokazuje (N4). Kao rezultat Teorema 6.9 dobivamo: Za $f \in L^2(\Omega)$, $f \leq 0$

$$\max_{x \in \bar{\Omega}} u_h(x) \leq \max_{x \in \partial\Omega} u_h(x).$$

Pogledajmo kako bismo proširili ovaj rezultat na složenije diferencijalne operatore. Najjednostavnije proširenje je ono u kome Laplaceovu jednadžbu zamijenim s

$$-\operatorname{div}(k \nabla u) = f,$$

gdje je k po trokutima konstantna skalarna funkcija. U tom slučaju dobivamo

$$(\tilde{A}_h)_{i,j} = -\frac{1}{2} [k_{K_1} \cot(\alpha_{i,j}^{K_1}) + k_{K_2} \cot(\alpha_{i,j}^{K_2})].$$

Negativnost tog člana ne predstavlja više čisto geometrijski uvjet i stoga je težak za analizu. Slučaj pune difuzijske matrice još daleko složeniji.

Primijetimo još da prisutnost članova nižeg reda ne čini probleme ukoliko pripadna diskretizacijska matrica ima pozitivnu dijagonalu i negativne vandijagonalne elemente. U slučaju jakog konvektivnog člana potrebno je koristiti speijalnu diskretizaciju kako bi se postigao taj efekt.

Bibliografija

- [1] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 1998.

7

Adaptivne metode i aposteriorna ocjena greške

7.1 Adaptivni konačni elementi

Brojne zadaće za parcijalne diferencijalne jednadžbe modeliraju fenomene koji su prostorno jako lokalizirani tako što u jednom dijelu domene, obično relativno malom, imaju singularitete ili snažne oscilacije. Tada obično u manjem dijelu domene moramo koristiti vrlo finu triangulaciju za reprezentaciju rješenja, dok u preostalom dijelu dostaje i grublja diskretizacija. U jednostavnijim situacijama takvi se lokalizirani fenomeni nalaze u dijelovima domene koji su na unaprijed poznati, na primjer rubni singulariteti kao u primjerima iz Sekcije 4.5. Općenito, potrebno je iz izračunate aproksimacije detektirati takva područja pomoću *procjenitelja greške*.

Nakon što smo izračunali aproksimativno rješenje varijacijske zadaće u_h na nekoj mreži \mathcal{T}_h i detektirali područje u kojem trebamo povećati preciznost aproksimacije preostaju nam dva glavna tipa adaptivne strategije: *h-shema* i *p-shema*.

U *p-shemi* na svakom elementu K na kojem je detektirana nedostatna preciznost aproksimacije obogaćujemo prostor funkcija P_K s dodatno baznom funkcijom; triangulacija pri tome ostaje nepromijenjena. Da bi taj pristup bio efikasan potrebno je prijeći s nodalne baze u prostoru konačnih elemenata na tzv. hijerarhijsku bazu. Taj je pristup novijeg datuma i s njim se ovdje više nećemo baviti.

U *h-shemi* svaki se element K , na kojem je detektirana nedostatna preciznost aproksimacije, dijeli u više manjih elemenata. Pri tome su moguće dvije tehnike: Ako se vrši subdivizija jednog elementa, onda se dijeli i susjedni elementi kako bi se dobila triangulacija u kojoj susjedni elementi dijele čitavu stranicu. Ta se tehnika koristi na trokutima (i tetraedrima). Važno je pri tome paziti da se ne izlazi iz okvira regularne familije triangulacija. Druga mogućnost je da se susjedni elementi ne profinjuju. U tom slučaju dobivena triangulacija ne zadovoljava uvjet zajedniče stranice, pa stoga postoje nodalne točke na granici dvaju elemenata koje pripadaju samo jednom od dva elementa. Takve su točke *neregularne* (*hanging nodes*) i da bismo na takvoj mreži postigli neprekidnost funkcije iz prostora konačnih elemenata vrijednost funkcije u neregularnoj točki

određujemo interpolacijom, na osnovu vrijednosti u susjednim regularnim točkama. Ta se tehnika obično koristi kod pravokutnih elemenata.

U složenijim strategijama moguće je pored profinjenja mreže u pojedinim područjima vršiti i okrupnjavanje mreže tamo gdje je greška isuviše mala. Takva mogućnost postoji kada procjenitelj daje uz ocjenu greške i procjenu lokalno optimalnog parametra h . Spomenimo konačno da za implementaciju tih algoritama treba koristiti posebno dizajniranu strukturu podataka za pamćenje triangulacije.

7.2 Aposteriorni procjenitelj greške

Postoje različiti tipovi aposteriornih procjenitelja greške, a najjednostavniji među njima su *rezidualni procjenitelji*. Od svakog procjenitelja se traži da je izračunljiv jedino iz aproksimativnog rješenja na zadanoj triangulaciji i poznatih podataka, te da je definiran *po elementima* (tj. da ima lokalni karakter). Uzmimo radi određenosti da imamo zadaću: naći $u \in V = H_0^1(\Omega)$, takvo da je

$$\forall v \in V, \quad a(u, v) = (f, v)_{L^2(\Omega)},$$

i neka je $V_h \subset V$ prostor konačnih elemenata konstruiran nad triangulacijom \mathcal{T}_h . Diskretan problem tada glasi: naći $u_h \in V_h$, takvo da je

$$\forall v \in V_h, \quad a(u_h, v) = (f, v)_{L^2(\Omega)},$$

Sada možemo definirati:

Definicija 7.1 Funkcija $e(h, u_h, f)$ se naziva aposteriorni procjenitelj greške ako je

$$\|u - u_h\| \leq e(h, u_h, f).$$

Ako procjenitelj $e(h, u_h, f)$ ima oblik

$$e(h, u_h, f) = \left(\sum_{K \in \mathcal{T}_h} e_K(u_h, f)^2 \right)^{1/2},$$

onda se $e_K(u_h, f)$ naziva indikator greške.

Pomoću indikatora greške možemo oformiti sljedeći adaptivni algoritam koji se bazira na *ekvidistribuciji* greške.

1. Na triangulaciji \mathcal{T}_h naći rješenje u_h ;
2. Za svako $K \in \mathcal{T}_h$ računati $e_K(u_h, f)$ i formirati $e(h, u_h, f)$.
3. Ako je $e(h, u_h, f) < \epsilon$ (ϵ je tolerancija) onda smo gotovi.

4. U suprotnom za svaki element $K \in \mathcal{T}_h$ odrediti optimalni h_K iz uvjeta da je $e_K(u_h, f) \approx \epsilon/\sqrt{N}$, $N = \text{card}(\mathcal{T}_h)$.
5. Izvršiti profinjenje/okrupnjenje triangulacije u skladu s novoizračunatim dimenzijama triangulacije.
6. GOTO 1.

U gornjem je algoritmu najvažniji peti korak u kojem treba osmisliti algoritam profinjenja i okrupnjenja triangulacije. Jedan dobar način se sastoji u tome da izračunamo željenu veličinu adaptirane mreže i konstruiramo novu mrežu u skladu s njom. Na primjer, indikator greške koji ćemo izvesti imat će oblik

$$e_K(u_h, f) = h_K r_K(u_h, f)$$

gdje r_K ne ovisi o h_K , odnosno ovisi vrlo slabo tako da se može ocijeniti s izrazom koji ne ovisi o h_K , i koji mu je vrlo blizak. Označimo novu mrežu koju želimo konstruirati s $\mathcal{T}_{\tilde{h}}$, i njoj pripadne veličine označimo tildom. Tada želimo postići

$$\tilde{e}_K(\tilde{u}_h, f) = \tilde{h}_K \tilde{r}_K(\tilde{u}_h, f) \approx \frac{\epsilon}{\sqrt{\tilde{N}}}, \quad \tilde{N} = \text{card}(\mathcal{T}_{\tilde{h}}).$$

Ako uvedmo pretpostavku da je

$$\tilde{r}_K(\tilde{u}_h, f) \approx r_K(u_h, f), \quad \tilde{N} \approx C^2 N,$$

gdje je C neka konstanta, onda imamo za željenu dimenziju nove mreže

$$\tilde{h}_K \approx h_K \frac{\epsilon}{e_K C \sqrt{N}}$$

Budući da smo ovdje nepoznate veličine na novoj mreži aproksimirali njihovim vrijednostima na staroj mreži, bit će potrebno napraviti nekoliko iteracija algoritma da se dođe do rješenja koje ima zadanu toleranciju.

Algoritam ekvidistribucije greške može biti vrlo neefikasan u smislu da ima tendenciju približno uniformnog profinjavanja mreže. Stoga su razvijene brojne druge strategije profinjavanja.

Jedna se jednostavna strategija sastoji u tome da se nova veličina mreže računa prema formuli

$$\tilde{h}_K \approx h_K f_K, \quad f_K = \frac{e_K}{C e^*}, \quad e^* = \frac{1}{N_h} \sum_{K \in \mathcal{T}_h} e_K,$$

gdje je $C \approx 1$ konstanta koju biramo, a e^* aritmetička sredina indikatora po cijeloj mreži ($N_h = \text{card}(\mathcal{T}_h)$). Pri tom se često ograničava f_K na neki interval $[a, b]$ i uzima se

$$f_K = \max(\min(f_K, b), a),$$

(npr. $a = 1/3$, $b = 3$). U toj se strategiji profinjenje vrši tamo gdje je indikator greške veći od prosječnog. Rezultat će jako ovisiti o početnoj mreži i bit će potrebno nekoliko iteracija metode.

7.3 Indikator za Laplaceovu jednadžbu

Uzmimo primjer homogene Dirichletove rubne zadaće za Laplaceovu jednadžbu:

$$-\Delta u = f \quad \text{u } \Omega, \quad (7.1)$$

$$u = 0 \quad \text{na } \partial\Omega \quad (7.2)$$

gdje je $\Omega \subset \mathbb{R}^d$ ograničena domena i $f \in L^2(\Omega)$.

Varijacijska formulacija zadaće ima oblik: naći $u \in H_0^1(\Omega)$ za koji vrijedi:

$$a(u, v) = F(v) \quad \forall v \in H_0^1(\Omega),$$

gdje je

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x}, \quad F(v) = \int_{\Omega} f v \, d\mathbf{x}.$$

Prikladna aproksimacija konačnim elementima ima oblik: naći $u_h \in V_h$ takvo da je

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

gdje je $V_h \subset H_0^1(\Omega)$, prostor konačnih elemenata.

Koristeći koercitivnost bilinearne forme dobivamo

$$\begin{aligned} \|u - u_h\|_{H^1(\Omega)} &\leq \frac{1}{\alpha} \frac{a(u - u_h, u - u_h)}{\|u - u_h\|_{H^1(\Omega)}} \leq \frac{1}{\alpha} \sup_{v \in H_0^1(\Omega)} \frac{a(u - u_h, v)}{\|v\|_{H^1(\Omega)}} \\ &= \frac{1}{\alpha} \sup_{v \in H_0^1(\Omega)} \frac{a(u - u_h, v - v_h)}{\|v\|_{H^1(\Omega)}}, \end{aligned}$$

za svako $v_h \in V_h$, jer vrijedi relacija ortogonalnosti $a(u - u_h, v_h) = 0$ za svako $v_h \in V_h$. Parcijalnom integracijom u bilinearnoj formi dobivamo:

$$\begin{aligned} a(u - u_h, v - v_h) &= \int_{\Omega} \nabla(u - u_h) \cdot \nabla(v - v_h) \, d\mathbf{x} \\ &= \int_{\Omega} \nabla u \cdot \nabla(v - v_h) \, d\mathbf{x} - \int_{\Omega} \nabla u_h \cdot \nabla(v - v_h) \, d\mathbf{x} \\ &= \int_{\Omega} [\operatorname{div}(\nabla u(v - v_h)) - \Delta u(v - v_h)] \, d\mathbf{x} \\ &\quad - \int_{\Omega} \nabla u_h \cdot \nabla(v - v_h) \, d\mathbf{x} \\ &= \int_{\partial\Omega} (\nabla u \cdot \mathbf{n})(v - v_h) \, d\sigma - \int_{\Omega} \Delta u(v - v_h) \, d\mathbf{x} \\ &\quad - \sum_{K \in \mathcal{T}_h} \int_K \nabla u_h \cdot \nabla(v - v_h) \, d\mathbf{x}. \end{aligned}$$

Iskoristimo sada činjenicu da je $v - v_h = 0$ na $\partial\Omega$ te da je $-\Delta u = f$, a u zadnjem integralu napravimo parcijalnu integraciju po svakom elementu posebno. Izlazi:

$$\begin{aligned} a(u - u_h, v - v_h) &= \int_{\Omega} f(v - v_h) \, d\mathbf{x} - \sum_{K \in \mathcal{T}_h} \int_K [\operatorname{div}(\nabla u_h(v - v_h)) - \Delta u_h(v - v_h)] \, d\mathbf{x} \\ &= \sum_{K \in \mathcal{T}_h} \int_K (f + \Delta u_h)(v - v_h) \, d\mathbf{x} - \sum_{K \in \mathcal{T}_h} \int_{\partial K} (\nabla u_h \cdot \mathbf{n})(v - v_h) \, d\sigma. \end{aligned}$$

Uočimo da suma u posljednjem članu ide samo po unutarnjim stranicama (bridovima u 2D) triangulacije jer je $v - v_h = 0$ na $\partial\Omega$. S druge strane, svaka unutarnja strana se u ukupnoj sumi obide dva puta i pri tome vanjska jedinična normala \mathbf{n} ima suprotan predznak. Stoga se u ukupnoj sumi javljaju skokovi normalne derivacije na granicama elemenata. Da bismo to precizirali uvedimo orijentaciju stranica triangulacije, tj. svakoj stranici γ dvaju susjednih elemenata K_1 i K_2 pridružimo jednu (od dvije moguće) normalu na stranicu \mathbf{n} . Za skok normalne derivacije uvedimo oznaku

$$[\nabla u_h \cdot \mathbf{n}] = (\nabla u_h \cdot \mathbf{n})|_{K_1} - (\nabla u_h \cdot \mathbf{n})|_{K_2},$$

pri čemu je K_1 element iz kojeg gleda odabrana normala \mathbf{n} . Sada možemo pisati

$$a(u - u_h, v - v_h) = \sum_{K \in \mathcal{T}_h} \int_K (f + \Delta u_h)(v - v_h) \, d\mathbf{x} - \sum_{K \in \mathcal{T}_h} \frac{1}{2} \int_{\partial K} [\nabla u_h \cdot \mathbf{n}](v - v_h) \, d\sigma.$$

Jednu polovicu smo stavili stoga što se svaka strana obilazi dva puta.

Koristeći Schwartz-Cauchyjevu nejednakost dolazimo do sljedećih ocjena:

$$\begin{aligned} a(u - u_h, v - v_h) &\leq \sum_{K \in \mathcal{T}_h} \|f + \Delta u_h\|_{L^2(K)} \|v - v_h\|_{L^2(K)} \\ &\quad + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \sum_{\gamma \subset \partial K} \|[\nabla u_h \cdot \mathbf{n}]\|_{L^2(\gamma)} \|v - v_h\|_{L^2(\gamma)} \\ &= \sum_{K \in \mathcal{T}_h} (h_K \|f + \Delta u_h\|_{L^2(K)}) \left(\frac{1}{h_K} \|v - v_h\|_{L^2(K)}\right) \\ &\quad + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \sum_{\gamma \subset \partial K} (h_\gamma^{1/2} \|[\nabla u_h \cdot \mathbf{n}]\|_{L^2(\gamma)}) \left(\frac{1}{h_\gamma^{1/2}} \|v - v_h\|_{L^2(\gamma)}\right) \end{aligned}$$

gdje je h_K dijаметar elementa K , a h_γ dijаметar stranice $\gamma \subset \partial K$. Tu ocjenu možemo pisati u obliku

$$\begin{aligned} a(u - u_h, v - v_h) &\leq \sum_{K \in \mathcal{T}_h} (h_K \|f + \Delta u_h\|_{L^2(K)} + \sum_{\gamma \subset \partial K} \frac{1}{2} h_\gamma^{1/2} \|[\nabla u_h \cdot \mathbf{n}]\|_{L^2(\gamma)}) \\ &\quad \times \max\left(\frac{1}{h_K} \|v - v_h\|_{L^2(K)}, \max_{\gamma \subset \partial K} \frac{1}{h_\gamma^{1/2}} \|v - v_h\|_{L^2(\gamma)}\right) \end{aligned}$$

Uvedimo oznaku

$$e_K(u_h, f) = h_K \|f + \Delta u_h\|_{L^2(K)} + \sum_{\gamma \subset \partial K} \frac{1}{2} h_\gamma^{1/2} \|[\nabla u_h \cdot \mathbf{n}]\|_{L^2(\gamma)}. \quad (7.3)$$

Kako je funkcija v_h proizvoljna odabrat ćemo $v_h = \Pi_h v$, gdje je Π_h interpolacijski operator. Na nesreću, ne možemo upotrijebiti standardan interpolacijski operator jer funkcija v ne leži, općenito, u $H^2(\Omega)$. U tu svrhu nam treba sofisticiraniji interpolacijski operator (vidi Bernardi, Girault [1]), kojeg ćemo označavati na isti način i koji zadovoljava:

$$\begin{aligned} \|v - \Pi_h v\|_{L^2(K)} &\leq Ch_K \|v\|_{H^1(V(K))}, \\ \|v - \Pi_h v\|_{L^2(\gamma)} &\leq Ch_\gamma^{1/2} \|v\|_{H^1(K_1 \cup K_2)}, \end{aligned}$$

gdje su K_1 i K_2 susjedni element sa zajedničkom stranicom γ , a $V(K)$ je unija svih elemenata koji imaju neprazan presjek s K (uključujući, naravno, i one koji dijele samo zajednički vrh ili brid).

Na osnovu gornjih formula dobivamo

$$\max\left(\frac{1}{h_K} \|v - v_h\|_{L^2(K)}, \max_{\gamma \subset \partial K} \frac{1}{h_\gamma^{1/2}} \|v - v_h\|_{L^2(\gamma)}\right) \leq C \|v\|_{H^1(V(K))}$$

gdje je $V(K)$ unija svih elemenata koji imaju neprazan presjek s K . Ako je triangulacija regularna broj takvih elemenata je ograničen neovisno o finoći triangulacije. Time smo dobili

$$\begin{aligned} a(u - u_h, v - v_h) &\leq C \sum_{K \in \mathcal{T}_h} e_K(u_h, f) \|v\|_{H^1(V(K))} \\ &\leq C \left(\sum_{K \in \mathcal{T}_h} e_K(u_h, f)^2 \right)^{1/2} \left(\sum_{K \in \mathcal{T}_h} \|v\|_{H^1(V(K))}^2 \right)^{1/2} \\ &\leq CN \left(\sum_{K \in \mathcal{T}_h} e_K(u_h, f)^2 \right)^{1/2} \|v\|_{H^1(\Omega)}, \end{aligned}$$

gdje je N maksimalan broj susjednih elemenata u triangulaciji. Time dobivamo ocjenu

$$\begin{aligned} \|u - u_h\|_{H^1(\Omega)} &\leq \frac{1}{\alpha} \sup_{v \in H^1(\Omega)} \frac{CN \left(\sum_{K \in \mathcal{T}_h} e_K(u_h, f)^2 \right)^{1/2} \|v\|_{H^1(\Omega)}}{\|v\|_{H^1(\Omega)}} \\ &= \frac{CN}{\alpha} \left(\sum_{K \in \mathcal{T}_h} e_K(u_h, f)^2 \right)^{1/2}. \end{aligned}$$

Time smo dokazali da je

$$e(h, u_h, f) = \left(\sum_{K \in \mathcal{T}_h} e_K(u_h, f)^2 \right)^{1/2}$$

greška a posteriori, a $e_K(u_h, f)$ indikator greške.

Bibliografija

- [1] V. Girault C. Bernardi. A local regularisation operator for triangular and quadrilateral finite elements. *SIAM J. Numer. Anal.*, 35(5):1893–1916, 1998.

8

Paraboličke rubne zadaće

U ovom poglavlju promatramo evolucijske zadaće prvog reda paraboličkog tipa. Cilj nam je postaviti varijacijsku formulaciju zadaće, opisati aproksimaciju konačnim elementima i napraviti procjenu greške.

8.1 Primjeri inicijalno-rubnih zadaća

Uzmimo primjer rubne zadaće za jednadžbu provođenja

$$\frac{\partial u}{\partial t} - \mu \Delta u = f \quad \text{u } \Omega \times (0, T), \quad (8.1)$$

gdje je $\Omega \subset \mathbb{R}^2$ ograničena domena i $\mu > 0$ konstanta. Diferencijalnoj jednadžbi treba pridružiti početni i rubni uvjet. Najjednostavniji rubni uvjet je Dirichletovog tipa:

$$u = 0 \quad \text{na } \partial\Omega \times (0, T), \quad (8.2)$$

$$u|_{t=0} = u_0 \quad \text{u } \Omega. \quad (8.3)$$

Funkcija $u_0(\mathbf{x})$ je zadana i da bi bila kompatibilna s rubnim uvjetom mora se poništavati na rubu domene Ω .

Općenito, ako je L eliptički diferencijalni operator drugog reda,

$$Lu = - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} (a_{ij} \frac{\partial u}{\partial x_j}) + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} + a_0 u$$

možemo promatrati evolucijsku zadaću prvog reda

$$\frac{\partial u}{\partial t} + Lu = f \quad \text{u } \Omega \times (0, T), \quad (8.4)$$

s pridruženim rubnim i početnim uvjetima. Ako je operator L eliptičan, onda kažemo da je evolucijska jednadžba (8.4) paraboličkog tipa. Uz tu zadaću možemo postaviti sve tipove rubnih uvjeta koji se mogu postaviti za pripadnu eliptičku jednadžbu $Lu = f$.

8.2 Varijacijska formulacija

Varijacijska formulacija evolucijske rubne zadaće formira se posve analogno kao za eliptičku zadaću. Pri tome se kompliciraju funkcijski prostori u kojima se traži rješenje. Pogledajmo prvo jednadžbu provođenja (8.1) u inicijalno-rubne uvjete (8.2)–(8.3).

Kao i kod eliptičke zadaće tražit ćemo da rješenje $u(\cdot, t)$ pripada prostoru $H_0^1(\Omega)$, za svako $t \in [0, T]$. Stoga se prirodno funkcija $u(\mathbf{x}, t)$ ne promatra kao funkcija varijabli \mathbf{x} i t , već kao preslikavanje $t \mapsto u(\mathbf{x}, t)$, koje svakom $t \in (0, T)$ pridružuje jednu funkciju iz prostora $H_0^1(\Omega)$. Od te se funkcije minimalno traži da joj je norma kvadratno integrabilna, tj.

$$\int_0^T \|u(t)\|_{H_0^1(\Omega)}^2 dt = \int_0^T \int_{\Omega} \left(u(\mathbf{x}, t)^2 + \sum_{i=1}^d \left(\frac{\partial u(\mathbf{x}, t)}{\partial x_i} \right)^2 \right) d\mathbf{x} dt < \infty.$$

Skup svih takvih funkcija je linearan prostor (štoviše i Hilbertov) i označava se s

$$L^2(0, T; H_0^1(\Omega)).$$

Za funkciju desne strane $f(x, t)$ pretpostavit ćemo da je kvadratno integrabilna na $Q_T = \Omega \times (0, T)$, tj.

$$\int_0^T \int_{\Omega} f(\mathbf{x}, t)^2 d\mathbf{x} dt = \int_0^T \|f(t)\|_{L^2(\Omega)}^2 dt < \infty.$$

Ponovo je prirodno f promatrati kao funkciju sa $(0, T)$ u $L^2(\Omega)$ i imamo

$$L^2(Q_T) = L^2(0, T; L^2(\Omega)).$$

Varijacijsku formulaciju dobivamo množenjem diferencijalne jednadžbe s test funkcijom iz $H_0^1(\Omega)$ i parcijalnom integracijom. Kao i kod eliptičke jednadžbe Dirichletov rubni uvjet se ugrađuje u prostor:

$$\int_{\Omega} \frac{\partial u(\mathbf{x}, t)}{\partial t} v(\mathbf{x}) d\mathbf{x} - \mu \int_{\Omega} \Delta u(\mathbf{x}, t) v(\mathbf{x}) d\mathbf{x} = \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}) d\mathbf{x}$$

Parcijalna integracija:

$$\begin{aligned} \int_{\Omega} \Delta u(\mathbf{x}, t) v(\mathbf{x}) d\mathbf{x} &= \int_{\Omega} \operatorname{div}(\nabla u(\mathbf{x}, t) v(\mathbf{x})) - \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}) d\mathbf{x} \\ &= \int_{\partial\Omega} \nabla u(\mathbf{x}, t) \cdot \mathbf{n} v(\mathbf{x}) ds - \int_{\Omega} \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}) d\mathbf{x} \\ &= - \int_{\Omega} \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}) d\mathbf{x}, \end{aligned}$$

gdje smo iskoristili da je $v(\mathbf{x}) = 0$ za sve $\mathbf{x} \in \partial\Omega$. Time smo dobili

$$\int_{\Omega} \frac{\partial u(\mathbf{x}, t)}{\partial t} v(\mathbf{x}) d\mathbf{x} + \mu \int_{\Omega} \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}) d\mathbf{x} = \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}) d\mathbf{x}, \quad (8.5)$$

za sve $v \in H_0^1(\Omega)$ i sve $t \in (0, T)$. Rješenje $u(\mathbf{x}, t)$ traži se u prostoru funkcija iz $L^2(0, T; H_0^1(\Omega)) \cap C([0, T]; L^2(\Omega))$ koje imaju vremenske derivacije

$$\frac{\partial u}{\partial t} \in L^2(0, T; L^2(\Omega)).$$

Uočimo još da smo birali test funkciju koja ne ovisi o vremenu, tako da možemo vremensku derivaciju izvući ispred integrala:

$$\frac{d}{dt} \int_{\Omega} u(\mathbf{x}, t) v(\mathbf{x}) \, d\mathbf{x} + \mu \int_{\Omega} \nabla u(\mathbf{x}, t) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} f(\mathbf{x}, t) v(\mathbf{x}) \, d\mathbf{x}, \quad (8.6)$$

za sve $v \in H_0^1(\Omega)$ i sve $t \in (0, T)$.

Napomena 8.1 Na vremensku derivaciju se obično postavlja slabiji uvjet, ali je za to potrebno uvesti dodatne funkcijske prostore. \square

Pogledajmo sada općeniti primjer diferencijalne jednadžbe (8.4) s inicijalno-rubnim uvjetima (8.2)–(8.3). Pretpostavimo da rubni problem

$$\begin{cases} Lu = f, & \text{u } \Omega \\ u = 0, & \text{na } \partial\Omega \end{cases}$$

ima varijacijsku formulaciju

$$\begin{cases} \text{Naći } u \in V \\ a(u, v) = F(v) \quad \forall v \in V. \end{cases} \quad (8.7)$$

Zbog homogenog Dirichletovog rubnog uvjeta ćemo imati $V = H_0^1(\Omega)$, dok će za općeniti rubni uvjet biti $H_0^1(\Omega) \subset V \subset H^1(\Omega)$. Tada je pripadna varijacijska formulacija evolucijske zadaće:

$$\begin{cases} \text{Naći } u \in L^2(0, T; V) \cap C([0, T]; L^2(\Omega)) \\ \frac{d}{dt}(u, v)_{L^2(\Omega)} + a(u, v) = F(v), & \forall v \in V \\ u(\cdot, 0) = u_0 \end{cases} \quad (8.8)$$

Ta se varijacijska jednadžba, kao i u slučaju jednadžbe provođenja, dobiva množenjem test funkcijom $v \in V$ (koja ne ovisi o vremenu) i parcijalom integracijom u prostornom dijelu. Vremensku derivaciju zatim možemo izvući ispred integrala. Dirichletov rubni uvjet je kao i u eliptičkoj zadaći ugrađen u prostor, dok su Neumannovi i Robienovi uvjeti zadovoljeni implicitno samom varijacijskom zadaćom. Inicijalni rubni uvjet je zadan eksplicitno i stoga rješenje mora biti neprekidno po varijabli t , tj. $u \in C([0, T]; L^2(\Omega))$. Za rješenje varijacijske zadaće (8.8) kažemo da je slabo rješenje diferencijalne zadaće (8.4).

Napomena 8.2 Ako koeficijenti diferencijalne jednadžbe ovise o vremenu t , onda će i pripadana bilinearna forma ovisiti eksplicitno o vremenu. U tom slučaju ju treba zapisivati u obliku $a(t; u, v)$ i treba pretpostaviti da je koercitivna i neprekidna za sve $t \in [0, T]$, s konstantama koercitivnosti i neprekidnosti neovisnim o vremenu. Isto je s funkcionalom F na desnoj strani. Ako desna strana diferencijalne jednadžbe (8.4) ovisi o vremenu, onda imamo linearan funkcional koji ovisi o vremenu: $F_t(v)$, na primjer

$$F_t(v) = \int_{\Omega} f(\mathbf{x}, t)v(\mathbf{x}) \, d\mathbf{x}.$$

I ovdje treba pretpostaviti da je funkcional neprekidan “uniformno po t ”, tj. da je

$$\sup_{0 \leq t \leq T} \|F_t\|_{V'} < +\infty.$$

8.3 Korektnost zadaće (8.8)

Pretpostavimo da je eliptička zadaća (8.7) korektno postavljena, tj. da su zadovoljeni uvjeti Lax-Milgramove leme. Tada *apriornu ocjenu* za rješenje parabolike zadaće (8.4) možemo izvesti na sljedeći način. U varijacijski jednadžbu (8.8) ubacimo $v = u(t)$:

$$\left(\frac{\partial u}{\partial t}, v \right)_{L^2(\Omega)} + a(t; u(t), u(t)) = F_t(u(t)).$$

U prvom člani možemo vremensku derivaciju izvući ispred integrala:

$$\left(\frac{\partial u}{\partial t}, v \right)_{L^2(\Omega)} = \frac{1}{2} \frac{d}{dt} \|u(t)\|_{L^2(\Omega)}^2.$$

U druga dva člana iskoristimo koercitivnos bilinearne forme i neprekidnost linearnog funkcionala:

$$a(t; u(t), u(t)) \geq \alpha \|u(t)\|_V^2, \quad |F_t(u(t))| \leq \|f(t)\|_{L^2(\Omega)} \|u(t)\|_{L^2(\Omega)} \leq \|f(t)\|_{L^2(\Omega)} \|u(t)\|_V.$$

Time dobivamo

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|u(t)\|_{L^2(\Omega)}^2 + \alpha \|u(t)\|_V^2 &\leq \|F_t\|_{V'} \|u(t)\|_V \\ &\leq \frac{2}{\alpha} \|f(t)\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \|u(t)\|_V^2 \end{aligned}$$

Sada će se zadnji član pokratiti s drugim na lijevoj strani. Integriranjem od 0 to t i uvažavanjem početnog uvjeta dobivamo

$$\frac{1}{2} \|u(t)\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \int_0^t \|u(t)\|_V^2 \, dt \leq \frac{2}{\alpha} \int_0^t \|f(t)\|_{L^2(\Omega)}^2 \, dt + \frac{1}{2} \|u_0\|_{L^2(\Omega)}^2.$$

Iz ove ocjene lako slijedi

$$\sup_{0 \leq t \leq T} \|u(t)\|_{L^2(\Omega)}^2 + \alpha \int_0^T \|u(t)\|_V^2 dt \leq \frac{4}{\alpha} \int_0^T \|f(t)\|_{L^2(\Omega)}^2 dt + \|u_0\|_{L^2(\Omega)}^2. \quad (8.9)$$

Ocjena (8.9) naziva se *energetska ocjena*. Iz nje zaključujemo da je slabo rješenje inicijalno-rubne zadaće (8.4) jedinstveno. Naime, zbog linearnosti zadaće, razlika dva rješenja $u = u_1 - u_2$ zadovoljava istu diferencijalnu jednadžbu s homogenim inicijalnim i rubnim uvjetima. Tada je $F_t \equiv 0$ i $u_0 \equiv 0$, pa je desna strana u (8.9) jednaka nuli, a onda je i $u \equiv 0$.

Energetska ocjena nam daje i neprekidnu zavisnost slabog rješenja zadaće (8.4) o zadanim podacima (početni i rubni uvjeti i desna strana). Za korektnost ostaje samo dokazati egzistenciju rješenja, u što nećemo ulaziti. Navodimo samo rezultat.

Teorem 8.1 *Neka je $V \subset H^1(\Omega)$ zatvoreni potprostor, $u_0 \in L^2(\Omega)$, $f \in L^2(\Omega \times (0, T))$ i neka je bilinearna forma $a: V \times V \rightarrow \mathbb{R}$ bilinearna, neprekidna i koercitivna. Tada zadaća (8.8) ima jedinstveno rješenje $u \in L^2(0, T; V) \cap C([0, T]; L^2(\Omega))$ koje zadovoljava (8.9).*

Za dokaz vidi [1], str. 366.

Uvjet koercitivnosti može se zamijeniti slabijim uvjetom *slabe koercitivnosti*. Kažemo je bilinearna forma $a: V \times V \rightarrow \mathbb{R}$, gdje je $V \subset H^1(\Omega)$ zatvoreni potprostor, slabo koercitivna ako postoje konstante $\alpha > 0$ i $\lambda \geq 0$, takve da je

$$\forall v \in V, \quad a(v, v) + \lambda \|v\|_{L^2(\Omega)}^2 \geq \alpha \|v\|_{H^1(\Omega)}^2. \quad (8.10)$$

Nejednakost (8.10) naziva se Gardingova nejednakost. Da bismo se u to uvjerali pretpostavimo da je u slabo rješenje diferencijalne jednadžbe (8.4) i uvedimo funkciju

$$u_\lambda(\mathbf{x}, t) = e^{-\lambda t} u(\mathbf{x}, t).$$

Lako se vidi da je ova funkcija rješenje diferencijalne jednadžbe

$$\frac{\partial u_\lambda}{\partial t} + Lu_\lambda + \lambda u_\lambda = e^{-\lambda t} f.$$

Novi diferencijalni operator je $L_\lambda = L + \lambda$ i njemu pripadna bilinearna forma je

$$a_\lambda(u, v) = a(u, v) + \lambda(u, v)_{L^2(\Omega)}.$$

Prema Gardingovoj nejednakosti ta je forma koercitivna, pa prema Teoremu 8.1 postoji jedinstveno rješenje pripadne varijacijske zadaće. No tada je $u(\mathbf{x}, t) = e^{\lambda t} u_\lambda(\mathbf{x}, t)$ slabo rješenje polazne zadaće. Time se vidi da je Gardingova nejednakost dovoljna za korektnost zadaće.

8.4 Diskretizacija

Izvodi se u dva koraka. Prvo diskretiziramo prostorni dio i dobivamo sustav običnih diferencijalnih jednažbi. Zatim rješavamo taj sustav nekom numeričkom metodom za obične diferencijalne jednažbe. Pri tome se, radi stabilnosti, biraju implicitne metode.

Prostorna diskretizacija se radi kao i u eliptičkom slučaju zamjenom osnovnog prostora V . Prostor V zamijenimo (konačnodimenzionalnim) potprostorom konačnih elemenata $V_h \subset V$ i dobivamo zadaću:

$$\begin{cases} \text{Naći } u_h \in L^2(0, T; V_h) \\ \frac{d}{dt}(u_h, v)_{L^2(\Omega)} + a(u_h, v) = (f, v)_{L^2(\Omega)}, \quad \forall v \in V_h \\ u_h(\cdot, 0) = u_{0h} \end{cases} \quad (8.11)$$

gdje je $u_{0h} \in V_h$ aproksimacija početnog podatka (npr. $u_{0h} = \Pi_h u_0$, gdje je Π_h interpolacijski operator za prostor V_h).

Ovime dobivamo sustav običnih diferencijalnih jednažbi. Odaberimo nodalnu bazu $\{\phi_1, \phi_2, \dots, \phi_N\}$ u prostoru V_h i zapišimo rješenje varijacijske jednažbe (8.11) u obliku

$$u_h(\mathbf{x}, t) = \sum_{j=1}^N U_j(t) \phi_j(\mathbf{x}).$$

Aproksimativni početni podatak u_{0h} mpžemo analogno zapisati,

$$u_{0h}(\mathbf{x}) = \sum_{j=1}^N U_j^0 \phi_j(\mathbf{x}),$$

gdje su koeficijenti U_j^0 poznati. Uzimajući za test funkciju v redom elemente baze ϕ_i , $i = 1, 2, \dots, N$, dobivamo linearan sustav običnih diferencijalnih jednažbi

$$\mathbf{M} \frac{dU}{dt}(t) + \mathbf{A}U(t) = F(t), \quad U(0) = U^0. \quad (8.12)$$

Ovdje je

$$U = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_N \end{bmatrix}, \quad F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_N \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} m_{1,1} & m_{1,2} & \cdots & m_{1,N} \\ m_{2,1} & m_{2,2} & \cdots & m_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ m_{N,1} & m_{N,2} & \cdots & m_{N,N} \end{bmatrix},$$

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{bmatrix},$$

gdje je

$$m_{i,j} = (\phi_j, \phi_i)_{L^2(\Omega)}, \quad a_{i,j} = a(\phi_j, \phi_i), \quad F_i = F(\phi_i) = (f, \phi_i)_{L^2(\Omega)}.$$

Uočimo da je matrica \mathbf{M} simetrična, pozitivno definitna, a da je \mathbf{A} simetrična ako je forma $a(\cdot, \cdot)$ isimetrična i pozitivno definitna ako je forma koercitivna.

Vremenska diskretizacija. Koriste se uglavnom implicitne metode. Implicitna Eulerova metoda s konstantnim vremenskim korakom Δt daje

$$\mathbf{M} \frac{U^n - U^{n-1}}{\Delta t} + \mathbf{A}U^n = F^n, \quad U^0 \text{ je zadano.}$$

Ovdje je $U^n \approx U(t_n)$, $t_n = n\Delta t$. Time u svakom koraku trebamo riješiti sustav

$$(\mathbf{M} + \Delta t \mathbf{A})U^n = \Delta t F^n + \mathbf{M}U^{n-1}.$$

Ovdje na diskretnoj razini vidimo da nije nužno da matrica \mathbf{A} bude regularna, već $\mathbf{M} + \Delta t \mathbf{A}$ treba biti regularna.

Druga vrsta diskretizacije koja se često koristi je Crank-Nicolsonova. U njoj vremensku derivaciju aproksimiramo centralnom diferencijom u točki $t_{n-1/2} = (t_n + t_{n-1})/2$:

$$\mathbf{M} \frac{U^n - U^{n-1}}{\Delta t} + \frac{1}{2} \mathbf{A}(U^n + U^{n-1}) = \frac{1}{2}(F^n + F^{n-1}), \quad U^0 \text{ je zadano.}$$

Ona vodi na rješavanje sustava

$$(\mathbf{M} + \frac{1}{2} \Delta t \mathbf{A})U^n = \frac{\Delta t}{2}(F^n + F^{n-1}) + (\mathbf{M} - \frac{1}{2} \Delta t \mathbf{A})U^{n-1}.$$

8.5 Korektnost poludiskretizirane zadaće

Budući da je zadaća (8.11) posve analogna kontinuiranoj zadaći (8.8), na posve isti način se za nju dokazuje energetsna nejednakost:

$$\sup_{0 \leq t \leq T} \|u_h(t)\|_{L^2(\Omega)}^2 + \alpha \int_0^T \|u_h(t)\|_V^2 dt \leq \frac{4}{\alpha} \int_0^T \|f(t)\|_{L^2(\Omega)}^2 dt + \|u_{0h}\|_{L^2(\Omega)}^2.$$

Za ocjenu greške rješenja trebat će nam dodatna ocjena na vremensku derivaciju od u_h . Nju ćemo lako dobiti uz dodatnu pretpostavku da je bilinearna forma $a(\cdot, \cdot)$ simetrična i neovisna o t . U jednakost

$$\left(\frac{d}{dt}u_h, v\right)_{L^2(\Omega)} + a(u_h, v) = (f, v)_{L^2(\Omega)}$$

uvrstimo $v = \frac{d}{dt}u_h$. Izlazi

$$\left\| \frac{du_h}{dt} \right\|_{L^2(\Omega)}^2 + a(u_h, \frac{du_h}{dt}) = (f, \frac{du_h}{dt})_{L^2(\Omega)}$$

Zbog pretpostavki o bilinearnoj formi imamo

$$a(u_h, \frac{du_h}{dt}) = \frac{1}{2} \frac{d}{dt} a(u_h, u_h),$$

što daje ocjenu

$$\begin{aligned} \left\| \frac{du_h}{dt} \right\|_{L^2(\Omega)}^2 + \frac{1}{2} \frac{d}{dt} a(u_h, u_h) &= (f, \frac{du_h}{dt})_{L^2(\Omega)} \leq \|f(t)\|_{L^2(\Omega)} \left\| \frac{du_h}{dt} \right\|_{L^2(\Omega)} \\ &\leq \frac{1}{2} \|f(t)\|_{L^2(\Omega)}^2 + \frac{1}{2} \left\| \frac{du_h}{dt} \right\|_{L^2(\Omega)}^2 \end{aligned}$$

Skraćivanjem i integriranjem po t dobivamo

$$\int_0^t \left\| \frac{du_h}{dt} \right\|_{L^2(\Omega)}^2 dt + a(u_h(t), u_h(t)) \leq \int_0^t \|f(t)\|_{L^2(\Omega)}^2 dt + a(u_{0h}, u_{0h}).$$

Koercitivnost i neprekidnost forme $a(\cdot, \cdot)$ daju

$$\int_0^t \left\| \frac{du_h}{dt} \right\|_{L^2(\Omega)}^2 dt + \alpha \|u_h(t)\|_V^2 \leq \int_0^t \|f(t)\|_{L^2(\Omega)}^2 dt + M \|u_{0h}\|_V^2. \quad (8.13)$$

za svako $t \in (0, T)$. Uzimanjem supremuma op t dobivamo drugu apriornu ocjenu:

$$\int_0^T \left\| \frac{du_h}{dt} \right\|_{L^2(\Omega)}^2 dt + \alpha \sup_{0 \leq t \leq T} \|u_h(t)\|_{H^1(\Omega)}^2 \leq \int_0^T \|f(t)\|_{L^2(\Omega)}^2 dt + M \|u_{0h}\|_{H^1(\Omega)}^2.$$

Ona postavlja dodatni uvjet na glatkoću početnog uvjeta, jer moramo imati $u_{0h} \in H^1(\Omega)$.

8.6 Ocjena greške za poludiskretiziranu zadaću

Sada želimo ocijeniti razliku $e_n(t) = u(t) - u_h(t)$, gdje je u rješenje kontinuirane zadaće (8.8), a u_h rješenje poludiskretizirane zadaće (8.11).

U tu svrhu ćemo pretpostaviti da je bilinearna forma $a(\cdot, \cdot)$ neprekidna i koercitivna s konstantama M i α . Nadalje, pretpostavit ćemo da je rješenje kontinuirane zadaće (8.8) dovoljno regularno (iz prostor $L^2(0, T; H^2(\Omega))$ te vrijedi (8.14)), a prostor V_h ćemo konstruirati pomoću P_1 elemenata.

Kada je aproksimacija prostora V konformna, tj. $V_h \subset V$ možemo oduzeti varijacijske jednadžbe kontinuiranog i diskretiziranog problema:

$$\begin{aligned}\frac{d}{dt}(u(t), v)_{L^2(\Omega)} + a(u(t), v) &= F(v) \\ \frac{d}{dt}(u_h(t), v)_{L^2(\Omega)} + a(u_h(t), v) &= F(v),\end{aligned}$$

uzimajući $v \in V_h$. Time dobivamo (zbog linearnosti zadaće)

$$\frac{d}{dt}(e_h(t), v)_{L^2(\Omega)} + a(e_h(t), v) = 0, \quad \forall v \in V_h.$$

Uzmimo proizvoljnu funkciju $w \in V_h$ i stavimo $v = w - u_h(t)$:

$$\left(\frac{de_h(t)}{dt}, w - u_h(t)\right)_{L^2(\Omega)} + a(e_h(t), w - u_h(t)) = 0.$$

Jednostavnom manipulacijom dobivamo

$$\begin{aligned}\left(\frac{de_h(t)}{dt}, e_h(t)\right)_{L^2(\Omega)} - \left(\frac{de_h(t)}{dt}, u(t) - w\right)_{L^2(\Omega)} + a(e_h(t), e_h(t)) \\ - a(e_h(t), u(t) - w) = 0.\end{aligned}$$

$$\begin{aligned}\frac{1}{2} \frac{d}{dt} \|e_h(t)\|_{L^2(\Omega)}^2 + a(e_h(t), e_h(t)) &= \left(\frac{de_h(t)}{dt}, u(t) - w\right)_{L^2(\Omega)} + a(e_h(t), u(t) - w) \\ &\leq \left\| \frac{de_h(t)}{dt} \right\|_{L^2(\Omega)} \|u(t) - w\|_{L^2(\Omega)} + M \|e_h(t)\|_{H^1(\Omega)} \|u(t) - w\|_{H^1(\Omega)}\end{aligned}$$

Na lijevoj strani ćemo iskoristiti koercitivnos bilinearne forme, pa dobivamo

$$\begin{aligned}\frac{1}{2} \frac{d}{dt} \|e_h(t)\|_{L^2(\Omega)}^2 + \alpha \|e_h(t)\|_{H^1(\Omega)}^2 &\leq \left\| \frac{de_h(t)}{dt} \right\|_{L^2(\Omega)} \|u(t) - w\|_{L^2(\Omega)} \\ &\quad + \frac{\alpha}{2} \|e_h(t)\|_{H^1(\Omega)}^2 + \frac{M^2}{2\alpha} \|u(t) - w\|_{H^1(\Omega)}^2,\end{aligned}$$

gdje smo na desnoj strana zadnji produkt normi ocijenili na uobičajeni način. Nakon skraćivanja dobivamo:

$$\begin{aligned}\frac{d}{dt} \|e_h(t)\|_{L^2(\Omega)}^2 + \alpha \|e_h(t)\|_{H^1(\Omega)}^2 &\leq 2 \left\| \frac{de_h(t)}{dt} \right\|_{L^2(\Omega)} \|u(t) - w\|_{L^2(\Omega)} \\ &\quad + \frac{M^2}{\alpha} \|u(t) - w\|_{H^1(\Omega)}^2.\end{aligned}$$

Uz pretpostavku da je $u(t) \in H^2(\Omega)$ za svako t , možemo uzeti $w = \Pi_h(u(t))$ i iskoristiti ocjene za grešku interpolacije:

$$\begin{aligned} \|u(t) - \Pi_h(u(t))\|_{L^2(\Omega)} &\leq Ch^2|u(t)|_{H^2(\Omega)} \\ \|u(t) - \Pi_h(u(t))\|_{H^1(\Omega)} &\leq Ch|u(t)|_{H^2(\Omega)}. \end{aligned}$$

Time dobivamo nejednakost

$$\frac{d}{dt} \|e_h(t)\|_{L^2(\Omega)}^2 + \alpha \|e_h(t)\|_{H^1(\Omega)}^2 \leq Ch^2 \left\| \frac{de_h(t)}{dt} \right\|_{L^2(\Omega)} |u(t)|_{H^2(\Omega)} + Ch^2 |u(t)|_{H^2(\Omega)}^2,$$

koju ćemo proitegirirati od 0 to t :

$$\begin{aligned} \|e_h(t)\|_{L^2(\Omega)}^2 + \alpha \int_0^t \|e_h(t)\|_{H^1(\Omega)}^2 dt &\leq Ch^2 \int_0^t \left\| \frac{de_h(t)}{dt} \right\|_{L^2(\Omega)} |u(t)|_{H^2(\Omega)} dt \\ &\quad + Ch^2 \int_0^t |u(t)|_{H^2(\Omega)}^2 dt. \end{aligned}$$

Ostaje nam još ocijeniti član

$$\left\| \frac{de_h(t)}{dt} \right\|_{L^2(\Omega)} = \left\| \frac{du(t)}{dt} - \frac{du_h(t)}{dt} \right\|_{L^2(\Omega)} \leq \left\| \frac{du(t)}{dt} \right\|_{L^2(\Omega)} + \left\| \frac{du_h(t)}{dt} \right\|_{L^2(\Omega)}$$

Za rješenje kontinuiranog problema ćemo pretpostaviti da je

$$\int_0^T \left\| \frac{du(t)}{dt} \right\|_{L^2(\Omega)}^2 dt < +\infty, \quad (8.14)$$

odnosno da je $\frac{du(t)}{dt} \in L^2(\Omega \times (0, T))$. Time smo došli do ocjene

$$\begin{aligned} \|e_h(t)\|_{L^2(\Omega)}^2 + \alpha \int_0^t \|e_h(t)\|_{H^1(\Omega)}^2 dt &\leq Ch^2 \int_0^t \left\| \frac{du(t)}{dt} \right\|_{L^2(\Omega)} |u(t)|_{H^2(\Omega)} dt \\ &\quad + Ch^2 \int_0^t \left\| \frac{du_h(t)}{dt} \right\|_{L^2(\Omega)} |u(t)|_{H^2(\Omega)} dt + Ch^2 \int_0^t |u(t)|_{H^2(\Omega)}^2 dt \\ &\leq Ch^2 \left[\int_0^t \left\| \frac{du(t)}{dt} \right\|_{L^2(\Omega)}^2 dt + \int_0^t \left\| \frac{du_h(t)}{dt} \right\|_{L^2(\Omega)}^2 dt + \int_0^t |u(t)|_{H^2(\Omega)}^2 dt \right]. \end{aligned}$$

Za diskretno rješenje imamo ocjenu (8.13), pa izlazi

$$\int_0^t \left\| \frac{du_h(t)}{dt} \right\|_{L^2(\Omega)}^2 dt \leq \int_0^t \|f(t)\|_{L^2(\Omega)}^2 dt + M \|u_{0h}\|_{H^1(\Omega)}^2$$

odnosno

$$\begin{aligned} \|e_h(t)\|_{L^2(\Omega)}^2 + \alpha \int_0^t \|e_h(t)\|_{H^1(\Omega)}^2 dt \\ \leq Ch^2 \left[\int_0^t \left\| \frac{du(t)}{dt} \right\|_{L^2(\Omega)}^2 dt + \int_0^t \|f(t)\|_{L^2(\Omega)}^2 dt + M \|u_{0h}\|_{H^1(\Omega)}^2 + \int_0^t |u(t)|_{H^2(\Omega)}^2 dt \right]. \end{aligned}$$

Uzimajući supremum po t dobivamo ocjenu

$$\begin{aligned} & \sup_{0 \leq t \leq T} \|e_h(t)\|_{L^2(\Omega)}^2 + \alpha \int_0^T \|e_h(t)\|_{H^1(\Omega)}^2 dt \\ & \leq Ch^2 \left[\int_0^T \left\| \frac{du(t)}{dt} \right\|_{L^2(\Omega)}^2 dt + \int_0^T \|f(t)\|_{L^2(\Omega)}^2 dt + M \|u_{0h}\|_{H^1(\Omega)}^2 + \int_0^T \|u(t)\|_{H^2(\Omega)}^2 dt \right]. \end{aligned}$$

Desna strana ne ovisi o h , osim kroz član $\|u_{0h}\|_{H^1(\Omega)}^2$, no budući da $u_{0h} \rightarrow u_0$ u $H^1(\Omega)$, to mora i

$$\|u_{0h}\|_{H^1(\Omega)}^2 \rightarrow \|u_0\|_{H^1(\Omega)}^2$$

kada $h \rightarrow 0$. Stoga je $\|u_{0h}\|_{H^1(\Omega)}^2$ uniformno ograničena, neovisno o h . Iz toga zaključujemo da postoji konstanta C , neovisna o h , koja ovisi o raznim normama točnog rješenja, takva da je

$$\sup_{0 \leq t \leq T} \|e_h(t)\|_{L^2(\Omega)}^2 + \alpha \int_0^T \|e_h(t)\|_{H^1(\Omega)}^2 dt \leq Ch^2.$$

To dokazuje da je metoda linearno konvergentna u normama na lijevoj strani.

Bibliografija

- [1] Alberto Valli Alfio Quarteroni. *Numerical Approximation of Partial Differential Equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer, Berlin, 1997.

Dodatak A

Funkcijski prostori

A.1 Diferencijalni operatori

A.1.1 Multiindeksi

Domena u \mathbb{R}^d je svaki otvoren i povezan skup. Kanonsku bazu u \mathbb{R}^d označavamo s $(\mathbf{e}_1, \dots, \mathbf{e}_d)$. Skalarni produkt označavamo točkom (\cdot) , a normu s $|\cdot|$.

Parcijalne derivacije funkcije $u: \Omega \rightarrow \mathbb{R}$ označavamo s

$$\frac{\partial u}{\partial x_i} = \partial_i u = u_{x_i}, \quad \frac{\partial^2 u}{\partial x_i \partial x_j} = \partial_{ij}^2 u = u_{x_i x_j}, \dots$$

Za označavanje viših derivacija uvodimo pojam multiindeksa. Multiindeks je svaka d -torka $\alpha = (\alpha_1, \dots, \alpha_d)$, $\alpha_i \in \mathbb{N} \cup \{0\}$, a duljina multiindeksa je broj

$$|\alpha| = \alpha_1 + \dots + \alpha_d.$$

Govorimo još da je α multiindeks reda $|\alpha|$. Zbrajanje i oduzimanje multiindeksa naslijeđuje se iz \mathbb{R}^d , $\alpha \pm \beta = (\alpha_1 \pm \beta_1, \dots, \alpha_d \pm \beta_d)$, dok se parcijalni uređaj uvodi prema pravilu

$$\alpha \leq \beta \Leftrightarrow \alpha_i \leq \beta_i \quad \forall i.$$

Evidentno nas oduzimanje $\alpha - \beta$ ostavlja u domeni multiindeksa samo ako je $\beta \leq \alpha$.

Multiindeks koristimo za kratko označavanje parcijalnih derivacija:

$$\partial^\alpha u(\mathbf{x}) = \left(\frac{\partial}{\partial x_1} \right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_d} \right)^{\alpha_d} u(\mathbf{x}) = \frac{\partial^{|\alpha|} u(\mathbf{x})}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}; \quad (\text{A.1})$$

te za označavanje monoma:

$$\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}. \quad (\text{A.2})$$

Uočimo da je \mathbf{x}^α monom stupnja $|\alpha|$ i da je za svako $t \in \mathbb{R}$, $(t\mathbf{x})^\alpha = t^{|\alpha|}\mathbf{x}^\alpha$. Notacija multiindeksa uzima u obzir da poredak primjene parcijalnih derivacija nije bitan, odnosno u multiindeks notaciji nije moguće razlikovati različite poretke primjena parcijalnih derivacija.

Pogledajmo ovaj primjer u \mathbb{R}^2 (sumira se po svim multiindeksima α čija je duljina 3):

$$\sum_{|\alpha|=3} \partial^\alpha u = \frac{\partial^3 u}{\partial x_1^3} + \frac{\partial^3 u}{\partial x_1^2 \partial x_2} + \frac{\partial^3 u}{\partial x_1 \partial x_2^2} + \frac{\partial^3 u}{\partial x_2^3}.$$

Ako isti izraz raspišemo u \mathbb{R}^3 dobivamo 10 članova:

$$\begin{aligned} \sum_{|\alpha|=3} \partial^\alpha u &= \frac{\partial^3 u}{\partial x_1^3} + \frac{\partial^3 u}{\partial x_1^2 \partial x_2} + \frac{\partial^3 u}{\partial x_1 \partial x_2^2} + \frac{\partial^3 u}{\partial x_1^2 \partial x_3} + \frac{\partial^3 u}{\partial x_1 \partial x_2^2} + \frac{\partial^3 u}{\partial x_2^2 \partial x_3} + \frac{\partial^3 u}{\partial x_2 \partial x_3^2} \\ &+ \frac{\partial^3 u}{\partial x_1 \partial x_2 \partial x_3} + \frac{\partial^3 u}{\partial x_2^3} + \frac{\partial^3 u}{\partial x_3^3}. \end{aligned}$$

Tu vidimo kompaktnost i praktičnost zapisa pomoću multiindeksa.

Broj multiindeksa reda k u \mathbb{R}^d je jednak $\binom{d+k-1}{k}$ (vidi [5], Primjer 38, str 98). To je ujedno broj svih parcijalnih derivacija reda k i broj svih monoma stupnja k .

Generalizacija binomnog teorema (multinomni teorem, vidi [5], str 123) jednostavno se izražava pomoću multiindeksa ako uvedemo $\alpha! = \alpha_1! \alpha_2! \cdots \alpha_d!$. Za svako $k \geq 0$ i $\mathbf{x} \in \mathbb{R}^d$ imamo

$$(x_1 + x_2 + \cdots + x_d)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \mathbf{x}^\alpha. \quad (\text{A.3})$$

A.1.2 Taylorov polinom

Taylorov teorem možemo zapisati u sljedećem obliku: za svako $f \in C^k([0, 1])$, $k \geq 1$, vrijedi

$$f(1) = \sum_{l=0}^{k-1} \frac{1}{l!} f^{(l)}(0) + \frac{1}{(k-1)!} \int_0^1 t^{k-1} f^{(k)}(1-t) dt. \quad (\text{A.4})$$

Primjenom multinomne formule (A.3) na $f(t) = u(\mathbf{x} + t\mathbf{h})$ dobivamo

$$\frac{d^k}{dt^k} f(t) \Big|_{t=0} = \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_d \frac{\partial}{\partial x_d} \right)^k u(\mathbf{x}) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha u(\mathbf{x}) \mathbf{h}^\alpha,$$

što daje formulu

$$u(\mathbf{x} + \mathbf{h}) = \sum_{|\alpha| < k} \frac{1}{\alpha!} \partial^\alpha u(\mathbf{x}) \mathbf{h}^\alpha + \sum_{|\alpha|=k} \mathbf{h}^\alpha \frac{k!}{\alpha!} \int_0^1 t^{k-1} \partial^\alpha u(\mathbf{x} + (1-t)\mathbf{h}) dt. \quad (\text{A.5})$$

Na primjer, u \mathbb{R}^2 , uzimajući $k = 3$ dobivamo

$$u(\mathbf{x} + \mathbf{h}) = u(\mathbf{x}) + [h_1 \partial_1 u(\mathbf{x}) + h_2 \partial_2 u(\mathbf{x})] + \frac{1}{2} [h_1^2 \partial_1^2 u(\mathbf{x}) + 2h_1 h_2 \partial_{12}^2 u(\mathbf{x}) + h_2^2 \partial_2^2 u(\mathbf{x})] \\ + \int_0^1 t^2 \left[\frac{h_1^3}{2} \partial_1^3 u(\mathbf{z}(t)) + \frac{3h_1^2 h_2}{2} \partial_1^2 \partial_2 u(\mathbf{z}(t)) + \frac{3h_1 h_2^2}{2} \partial_1 \partial_2^2 u(\mathbf{z}(t)) + \frac{h_2^3}{2} \partial_2^3 u(\mathbf{z}(t)) \right] dt.$$

gdje je $\mathbf{z}(t) = \mathbf{x} + (1-t)\mathbf{h}$. Taylorova formula (A.5) vrijedi uz uvjet da je funkcija u k -puta neprekidno derivabilna i da je $\mathbf{z}(t) \in \Omega$ za $t \in [0, 1]$.

Iz Taylorovog razvoja imamo sljedeći zaključak: Ako je $\partial^\alpha u \equiv 0$ za sve $|\alpha| = k+1$, onda je $u \in \mathbb{P}_k$ (polinom stupnja manjeg ili jednakog k). Sličan zaključak vrijedi i za funkcije iz \mathbb{Q}_k .

Funkcije iz \mathbb{P}_k možemo pisati u obliku

$$p(\mathbf{x}) = \sum_{|\alpha| < k+1} a_\alpha \mathbf{x}^\alpha,$$

gdje su a_α neki koeficijenti. S druge strane, funkcije iz \mathbb{Q}_k su polinomi stupnja strogo manjeg od $k+1$ u svakoj varijabli posebno, pa ih možemo zapisivati u obliku

$$q(\mathbf{x}) = \sum_{\alpha \in A_{k+1}} a_\alpha \mathbf{x}^\alpha, \quad A_{k+1} = \left\{ \alpha : \frac{\partial^{k+1}}{\partial x_i^{k+1}} \mathbf{x}^\alpha = 0, \forall i = 1, \dots, d \right\}. \quad (\text{A.6})$$

Ako je sada funkcija $u \in C^{kd}(\Omega)$ takva da je

$$\frac{\partial^{k+1} u}{\partial x_i^{k+1}} \equiv 0, \quad \forall i = 1, \dots, d, \quad (\text{A.7})$$

onda zaključujemo da je $u \in \mathbb{Q}_k$. Zaista, iz uvjeta (A.7) slijedi da je $\partial^\alpha u \equiv 0$ za sve $|\alpha| > kd$ (barem jedan α_i mora biti veći od k), pa je stoga $u \in \mathbb{P}_{kd}$. Iz posve istog razloga je $\partial^\alpha u \equiv 0$ za svako $\alpha \notin A_{k+1}$ i stoga je $u(\mathbf{x})$ oblika (A.6) i zaključak slijedi.

A.1.3 Diferencijalni operatori

$D^k u(x) = \{\partial^\alpha u(x) : |\alpha| = k\}$ je skup svih k -tih parcijalnih derivacija. Posebno imamo gradijent

$$Du(x) = (\partial_1 u(x), \dots, \partial_n u(x)) = \nabla u(x);$$

i hesijan

$$D^2 u(x) = \begin{bmatrix} \partial_{11}^2 u(x) & \cdots & \partial_{1n}^2 u(x) \\ \vdots & \ddots & \vdots \\ \partial_{n1}^2 u(x) & \cdots & \partial_{nn}^2 u(x) \end{bmatrix}.$$

Laplaceov operator:

$$\Delta u(x) = \text{tr}(D^2 u(x)) = \sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}.$$

Gradijent vektorske funkcije $\mathbf{u}: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ je matrica

$$D\mathbf{u}(x) = \nabla\mathbf{u}(x) = \begin{bmatrix} \partial_1 u_1(x) & \cdots & \partial_n u_1(x) \\ \vdots & \ddots & \vdots \\ \partial_1 u_n(x) & \cdots & \partial_n u_n(x) \end{bmatrix}.$$

U i -tom retku matrice nalazi se gradijent i -te komponente funkcije. Laplaceov operator vektorske funkcije definira se po komponentama:

$$\Delta\mathbf{u}(x) = (\Delta u_i(x)).$$

Divergencija vektorske funkcije:

$$\operatorname{div} \mathbf{u}(x) = \sum_{i=1}^n \frac{\partial u_i(x)}{\partial x_i}.$$

Posebno je

$$\Delta u(x) = \operatorname{div}(\nabla u(x)).$$

Rotacija vektorske funkcije u \mathbb{R}^3 :

$$\operatorname{rot} \mathbf{u}(x) = \begin{bmatrix} \partial_2 u_3(x) - \partial_3 u_2(x) \\ \partial_3 u_1(x) - \partial_1 u_3(x) \\ \partial_1 u_2(x) - \partial_2 u_1(x) \end{bmatrix} = (\text{formalno}) = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ u_1(x) & u_2(x) & u_3(x) \end{vmatrix}.$$

Zadatak A.1 Neka je $\phi(x)$ skalarna glatka funkcija, a $\mathbf{u}(x)$ vektorska. Dokažite formulu

$$\operatorname{div}(\phi(x)\mathbf{u}(x)) = \phi(x) \operatorname{div} \mathbf{u}(x) + \nabla\phi(x) \cdot \mathbf{u}(x). \quad \square$$

Teorem o divergenciji. Za dovoljno regularno područje $\Omega \subset \mathbb{R}^n$ i dovoljno glatku vektorsku funkciju $\mathbf{F}: \Omega \rightarrow \mathbb{R}^n$ vrijedi

$$\int_{\Omega} \operatorname{div} \mathbf{F}(x) dx = \int_{\partial\Omega} \mathbf{F}(x) \cdot \mathbf{n}(x) dS$$

gdje je \mathbf{n} polje jedinične vanjske normale na $\partial\Omega$. Posebno je

$$\int_{\Omega} \Delta u(x) dx = \int_{\partial\Omega} \nabla u(x) \cdot \mathbf{n}(x) dS.$$

Pišemo

$$\frac{\partial u(x)}{\partial \mathbf{n}} = \nabla u(x) \cdot \mathbf{n}(x).$$

Za linearnu PDJ-u drugog reda kažemo da je u divergentnoj formi ako se može zapisati u obliku

$$-\sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\sum_{j=1}^n a_{ij}(x) \frac{\partial u}{\partial x_j} \right) + \sum_{i=1}^n b_i(x) \frac{\partial u}{\partial x_i} + c(x)u = f(x).$$

Takvu jednadžbu češće pišemo u vektorskoj notaciji

$$-\operatorname{div}(\mathbf{A}(x)\nabla u) + \mathbf{b}(x) \cdot \nabla u + c(x)u = f(x).$$

Ako su koeficijenti a_{ij} dovoljno glatki, onda se jednadžba može iz nedivergentnog oblika prebaciti u divergentni i obratno. Pri tome se matrica \mathbf{A} ne mijenja.

Rubne zadaće za PDJ. Za različite tipove PDJ-bi zadaju se različiti uvjeti na granici domene. Evo nekoliko primjera.

$$\begin{aligned} -\Delta u &= f & x \in \Omega & \quad \text{Dirichletova zadaća} \\ u &= u_0 & x \in \partial\Omega \end{aligned}$$

$$\begin{aligned} -\Delta u &= f & x \in \Omega & \quad \text{Neumannova zadaća} \\ \frac{\partial u}{\partial \mathbf{n}} &= g & x \in \partial\Omega \end{aligned}$$

$$\begin{aligned} u_t - \Delta u &= f & x \in \Omega, t > 0 \\ u &= u_0 & x \in \partial\Omega, t > 0 & \quad \text{rubni uvjet} \\ u|_{t=0} &= v_0 & x \in \Omega & \quad \text{početni uvjet} \end{aligned}$$

$$\begin{aligned} u_{tt} - \Delta u &= f & x \in \Omega, t > 0 \\ u &= u_0 & x \in \partial\Omega, t > 0 & \quad \text{rubni uvjet} \\ u|_{t=0} &= v_0 & x \in \Omega & \quad 1. \text{ početni uvjet} \\ u_t|_{t=0} &= v_1 & x \in \Omega & \quad 2. \text{ početni uvjet} \end{aligned}$$

U ovim primjerima f , g , u_0 , v_0 i v_1 su zadani podaci, a u je nepoznanica.

Rubni i početni uvjeti za diferencijalnu jednadžbu određeni su fizikalnim smislom promatranog problema. S matematičke strane oni moraju biti odabrani tako da zadaća bude korektno postavljena u smislu ove definicije:

Definicija A.1 (Hadamard) Rubna zadaća za PDJ je dobro postavljena (korektna) ako ima jedinstveno rješenje koje neprekidno ovisi o zadanim podacima.

A.2 Prostori glatkih funkcija

Neka je $\Omega \subset \mathbb{R}^d$ otvoren i ograničen skup. Tada skup funkcija

$$C(\Omega) = \{u: \Omega \rightarrow \mathbb{R}: u \text{ je neprekidna na } \Omega\},$$

predstavlja linearan prostor u odnosu na uobičajene operacije zbrajanja funkcija i množenja skalarom. Za svako $k \in \mathbb{N}$ možemo definirati potprostor glatkih funkcija

$$C^k(\Omega) = \{u \in C(\Omega): D^\alpha u \in C(\Omega) \text{ za sve } |\alpha| \leq k\},$$

koje imaju neprekidne sve parcijalne derivacije do uključivo k -tog reda. U skladu s time se ponekad koristi oznaka $C^0(\Omega) = C(\Omega)$.

Funkcije iz $C^k(\Omega)$ prostora su općenito neograničene i nisu dobro definirane na granici domene $\partial\Omega$. Stoga se uvode potprostori tih prostora koji se sastoje od funkcija neprekidnih “do ruba”, tj. funkcija koje se mogu po neprekidnosti proširiti na otvoreni skup koji sadrži $\bar{\Omega}$; za takve funkcije kažemo da su neprekidne na $\bar{\Omega}$.

$$\begin{aligned} C(\bar{\Omega}) &= \{u: \Omega \rightarrow \mathbb{R}: u \text{ je neprekidna na } \bar{\Omega}\}, \\ C^k(\bar{\Omega}) &= \{u \in C(\bar{\Omega}): D^\alpha u \in C(\bar{\Omega}) \text{ za sve } |\alpha| \leq k\}, \quad k \in \mathbb{N}. \end{aligned}$$

Kao i prije pišemo $C^0(\bar{\Omega}) = C(\bar{\Omega})$ i koristimo sljedeće dvije oznake:

$$C^\infty(\Omega) = \bigcap_{k=1}^{\infty} C^k(\Omega), \quad C^\infty(\bar{\Omega}) = \bigcap_{k=1}^{\infty} C^k(\bar{\Omega}).$$

U prostor $C^k(\bar{\Omega})$ se uvodi norma na sljedeći način:

$$\|u\|_{C^k(\bar{\Omega})} = \max_{|\alpha| \leq k} \sup_{\mathbf{x} \in \bar{\Omega}} |\partial^\alpha u(\mathbf{x})|. \quad (\text{A.8})$$

Posebno, u slučaju $k = 0$ imamo

$$\|u\|_{C(\bar{\Omega})} = \sup_{\mathbf{x} \in \bar{\Omega}} |u(\mathbf{x})|.$$

$C^k(\bar{\Omega})$ Nije teško provjeriti da je sa (A.8) dana jedna norma na $C^k(\bar{\Omega})$. Štoviše, u toj normi je prostor $C^k(\bar{\Omega})$ potpun.

Teorem A.1 Neka je $\Omega \subset \mathbb{R}^d$ otvoren i ograničen skup. Za sve $k \geq 0$ prostor $C^k(\bar{\Omega})$ s normom (A.8) je Banachov.

U vezi s rubnim zadaćama promatrat ćemo potprostor od $C^k(\bar{\Omega})$ koji se sastoji od funkcija koje se poništavaju na $\partial\Omega$, zajedno sa svim parcijalnim derivacijama do uključivo reda k , i označavat ćemo ga s $C_0^k(\Omega)$. U teoriji distribucija se promatraju funkcije koje se poništavaju već na nekoj pozitivnoj udaljenosti od ruba domene pa se stoga uvodi pojam nosača funkcije, $\text{supp}(u)$:

$$\text{supp}(u) = \overline{\{x \in \Omega: u(x) \neq 0\}}.$$

Sada se neovisno o tome je li skup Ω ograničen ili neograničen uvodi prostor

$$C_c^k(\Omega) = \{u \in C^k(\Omega) : u \text{ ima kompaktan nosač}\}.$$

Konačno, prostor test-funkcija iz teorije distribucija definira se kao

$$\mathcal{D}(\Omega) = C_c^\infty(\Omega) = \bigcap_{k=1}^{\infty} C_c^k(\Omega).$$

A.3 L^p prostori

Pretpostavljamo da je $\Omega \subset \mathbb{R}^n$ otvoren skup opskrbljen Lebesgueovom mjerom naslijeđenom s \mathbb{R}^n . Za dvije funkcije $f, g: \Omega \rightarrow \mathbb{R}$ kažemo da su jednake skoro svuda i pišemo

$$f = g \quad \text{s.s.}$$

ako postoji skup $A \subset \Omega$ Lebesgueove mjere jednake nuli, takav da je

$$\forall x \in \Omega \setminus A, \quad f(x) = g(x).$$

Svake takve dvije funkcije smatrat ćemo jednakim.

Definicija A.2 Za $p \in \mathbb{R}$, $1 \leq p < \infty$ definiramo

$$L^p(\Omega) = \{f: \Omega \rightarrow \mathbb{R} : f \text{ je izmjeriva i } \int_{\Omega} |f(x)|^p dx < \infty\}.$$

Uvodimo oznaku

$$\|f\|_{L^p} = \left[\int_{\Omega} |f(x)|^p dx \right]^{1/p}.$$

Definicija A.3

$$L^\infty(\Omega) = \{f: \Omega \rightarrow \mathbb{R} : f \text{ je izmjeriva i } \exists \text{ konstanta } C \text{ tdj. } |f(x)| \leq C \text{ s.s.}\}.$$

Uvodimo oznaku

$$\|f\|_{L^\infty} = \inf\{C : |f(x)| \leq C \text{ s.s.}\}.$$

Uvijek imamo

$$|f(x)| \leq \|f\|_{L^\infty} \text{ s.s.}$$

Ako je f neprekidna funkcija, onda je

$$\|f\|_{L^\infty(\Omega)} = \sup_{x \in \Omega} |f(x)|.$$

Definicija A.4 Za funkciju kažemo da je lokalno integrabilna na Ω ako je integrabilna na svakom kompaktnom podskupu od Ω . Prostor svih lokalno integrabilnih funkcija označavamo $L^1_{\text{loc}}(\Omega)$.

Nejednakost

$$ab \leq \frac{1}{2}(a^2 + b^2),$$

koja je korisna u mnogim situacijama, može se generalizirati na sljedeći način. Prvo uvedemo konjugirane indekse (eksponente):

Definicija A.5 Neka je $1 \leq p \leq \infty$. Konjugirani indeks p' definiran je relacijom

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Za $p = \infty$ imamo $p' = 1$ i obratno.

Zatim imamo:

Lema A.1 Youngova nejednakost: Neka su $a, b \geq 0$ i $1 < p < \infty$. Tada je

$$ab \leq \frac{1}{p}a^p + \frac{1}{p'}b^{p'}.$$

Dokaz. Koristimo činjenicu da je funkcija $E(x) = e^x$ konveksna. Imamo,

$$\begin{aligned} ab &= e^{\ln a + \ln b} = E\left(\frac{1}{p} \ln a^p + \frac{1}{p'} \ln b^{p'}\right) \\ &\leq \frac{1}{p}E(\ln a^p) + \frac{1}{p'}E(\ln b^{p'}) = \frac{1}{p}a^p + \frac{1}{p'}b^{p'}. \quad \square \end{aligned}$$

Zadatak A.2 Pokažite da za svako $a, b \geq 0$, $\varepsilon > 0$ i $1 < p < \infty$ vrijedi

$$ab \leq \varepsilon a^p + \frac{1}{p'(p\varepsilon)^{p'/p}} b^{p'}.$$

Lema A.2 Hölderova nejednakost. Neka je $1 \leq p \leq +\infty$ te $f \in L^p(\Omega)$ i $g \in L^{p'}(\Omega)$. Tada je $fg \in L^1(\Omega)$ i vrijedi

$$\int_{\Omega} |f(x)g(x)| dx \leq \|f\|_p \|g\|_{p'}.$$

Dokaz. a) Slučajevi $p = 1$ i $p = \infty$ su trivijalni.

b) $1 < p < \infty$. Tvrdnja je trivijalna ako je jedna od funkcija jednaka nuli. Pretpostavimo stoga da su obje različite od nule. Primjenom Youngove nejednakosti dobivamo

$$\int_{\Omega} |f(x)g(x)| dx \leq \frac{1}{p} \int_{\Omega} |f(x)|^p dx + \frac{1}{p'} \int_{\Omega} |g(x)|^{p'} dx.$$

Ako tu nejednakost primijenim na funkcije $f(x)/\|f\|_p$ i $g(x)/\|g\|_{p'}$ dobivamo

$$\frac{1}{\|f\|_p \|g\|_{p'}} \int_{\Omega} |f(x)g(x)| dx \leq \frac{1}{p} + \frac{1}{p'} = 1,$$

odakle slijedi tvrdnja. \square

Lema A.3 *Nejednakost Minkowskog. Za svako $1 \leq p \leq \infty$ vrijedi*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Dokaz. Primjenom Hölderove nejednakosti slijedi

$$\begin{aligned} \|f + g\|_p^p &= \int_{\Omega} |f(x) + g(x)|^p dx \leq \int_{\Omega} |f(x) + g(x)|^{p-1} (|f(x)| + |g(x)|) dx \\ &\leq \left(\int_{\Omega} |f(x) + g(x)|^p dx \right)^{(p-1)/p} \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p} \\ &\quad + \left(\int_{\Omega} |f(x) + g(x)|^p dx \right)^{(p-1)/p} \left(\int_{\Omega} |g(x)|^p dx \right)^{1/p} \\ &= \|f + g\|_p^{p-1} (\|f\|_p + \|g\|_p). \quad \square \end{aligned}$$

Iz nejednakosti Minkowskog lako se pokazuje da $L^p(\Omega)$ ima strukturu linearnog prostora i da je $\|f\|_p$ norma na $L^p(\Omega)$ za $1 \leq p \leq \infty$. Ovi prostori imaju sljedeća svojstva:

Teorem A.2 • *Prostor $L^p(\Omega)$ je Banachov za $1 \leq p \leq \infty$;*

• *Prostor $L^2(\Omega)$ je Hilbertov sa skalarnim produktom*

$$(f, g)_{L^2} = \int_{\Omega} f(x)g(x) dx.$$

Napomena A.1 *U slučaju kada želimo naglasiti domenu, koristit ćemo oznaku za normu $\|f\|_{p;\Omega}$ ili $\|f\|_{L^p(\Omega)}$. \square*

Sljedeće dvije leme govore o funkcijama iz L^∞ prostora.

Lema A.4 *Neka je domena Ω ograničena (ili samo konačne mjere). Tada za svako $u \in L^\infty(\Omega)$ vrijedi*

$$\|u\|_\infty = \lim_{p \rightarrow \infty} \|u\|_p.$$

Lema A.5 *Neka je domena Ω ograničena (ili samo konačne mjere). Ako je*

$$u \in \bigcap_{p=1}^{\infty} L^p(\Omega) \quad \text{i} \quad \sup_{p \in \mathbb{N}} \|u\|_p < \infty,$$

onda je $u \in L^\infty(\Omega)$.

A.4 $W^{1,p}$ prostori

Pojam parcijalne derivacije generalizira se na sljedeći način:

Definicija A.6 Za funkciju $f \in L^p(\Omega)$, $1 \leq p \leq \infty$ kažemo da ima slabu parcijalnu derivaciju po varijabli x_i , ako postoji funkcija $g \in L^1_{\text{loc}}(\Omega)$, takva da za svako $\phi \in \mathcal{D}(\Omega)$ vrijedi

$$\int_{\Omega} f(x) \frac{\partial \phi}{\partial x_i}(x) dx = - \int_{\Omega} g(x) \phi(x) dx.$$

Funkcija g je tada slaba parcijalna derivacija funkcije f po x_i i koristimo standardnu oznaku

$$g = \frac{\partial f}{\partial x_i}.$$

Iz ove definicije je jasno da je klasična parcijalna derivacija, ako postoji, ujedno i slaba derivacija. Nadalje, lako se pokazuje da je slaba derivacija jedinstvena (dokažite!). Stoga je ovaj pojam dobro proširenje klasičnog pojma derivacije. U tekstu koji slijedi sve parcijalne derivacije se uzimaju u slabom smislu.

Za $1 \leq p \leq \infty$ definiramo

$$W^{1,p}(\Omega) = \{u \in L^p(\Omega) : \frac{\partial u}{\partial x_i} \in L^p(\Omega), 1 \leq i \leq n\}.$$

Prostor $W^{1,p}(\Omega)$ je očito linearan. Za $1 \leq p < \infty$ normu u njemu definiramo formulom

$$\|u\|_{1,p} = \left(\int_{\Omega} |u(x)|^p dx + \sum_{i=1}^n \int_{\Omega} \left| \frac{\partial u}{\partial x_i}(x) \right|^p dx \right)^{1/p}.$$

Za $p = \infty$ imamo

$$\|u\|_{1,\infty} = \max(\|u\|_{\infty}, \max_{1 \leq i \leq n} \left\| \frac{\partial u}{\partial x_i} \right\|_{\infty}).$$

Istaknuto mjesto ima slučaj $p = 2$, pa stoga taj prostor imamo posebnu oznaku

$$W^{1,2}(\Omega) = H^1(\Omega).$$

U taj prostor može se uvesti skalarni produkt formulom

$$(u, v)_1 = \int_{\Omega} u(x)v(x) dx + \sum_{i=1}^n \int_{\Omega} \frac{\partial u}{\partial x_i}(x) \frac{\partial v}{\partial x_i}(x) dx.$$

Prostore $W^{1,p}(\Omega)$ nazivamo prostorima Soboljeva prvog reda. Posve analogno se definiraju prostori Soboljeva višeg reda ($m = 1, 2, 3, \dots$).

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) : \partial^{\alpha} u \in L^p(\Omega), \text{ za sve } |\alpha| \leq m\}.$$

Normu definiramo na sljedeći način: Za $1 \leq p < \infty$,

$$\|u\|_{m,p} = \left(\sum_{|\alpha| \leq m} \int_{\Omega} |\partial^{\alpha} u(x)|^p dx \right)^{1/p}.$$

te za $p = \infty$,

$$\|u\|_{m,\infty} = \max_{|\alpha| \leq m} \|\partial^{\alpha} u\|_{\infty}.$$

Ponovo, u slučaju $p = 2$ dobivamo prostor sa skalarnim produktom koji označavamo

$$W^{m,2}(\Omega) = H^m(\Omega),$$

a skalarni produkt je dan formulom

$$(u, v)_m = \sum_{|\alpha| \leq m} \int_{\Omega} \partial^{\alpha} u(x) \partial^{\alpha} v(x) dx.$$

Napomena A.2 U slučaju kada želimo naglasiti domenu, koristit ćemo oznaku za normu $\|u\|_{m,p;\Omega}$ ili $\|u\|_{W^{m,p}(\Omega)}$.

Svi prostori Soboljeva su potpuni.

Teorem A.3 Za $m \in \{1, 2, \dots\}$ i $p \in [1, \infty]$ prostor $W^{m,p}(\Omega)$ je Banachov. Prostor $H^m(\Omega)$ je Hilbertov.

Neka je $\partial\Omega$ oznaka za rub domene (granicu) Ω . Mi ćemo uvijek pretpostavljati da je to ploha određene glatkoće (vidi sekciju A.4.2). Definiramo prostor:

$$W_0^{1,p}(\Omega) = \{\phi \in W^{1,p}(\Omega) : \phi(x) = 0 \text{ za } x \in \partial\Omega\}.$$

Koristimo oznaku

$$H_0^1(\Omega) = W_0^{1,2}(\Omega).$$

U prostorima $W^{m,p}$ se definiraju polunorme

$$|u|_{m,p} = \left(\sum_{|\alpha|=m} \int_{\Omega} |\partial^{\alpha} u(x)|^p dx \right)^{1/p},$$

za $1 \leq p < \infty$, te analogno za $p = \infty$. U tim se polunormama pojavljuju samo parcijalne derivacije najvišeg reda.

Teorem A.4 (Poincaréova nejednakost) Ako je domena Ω ograničena barem u jednom smjeru, onda postoji konstanta $c = c(\Omega, p)$ takva da za svako $u \in W_0^{1,p}(\Omega)$ vrijedi

$$\|u\|_{L^p} \leq c|u|_{1,p}, \quad (1 \leq p < \infty). \quad (\text{A.9})$$

Dokaz. Dokazat ćemo teorem u slučaju $p = 2$. Bez smanjenja općenitosti možemo uzeti da se domena nalazi unutar pruge $a < x_n < b$. Uzmimo proizvoljnu funkciju $v \in H_0^1(\Omega)$ i proširimo ju nulom izvan skupa Ω . Imamo

$$v(x', x_n) = \int_a^{x_n} \frac{\partial v}{\partial x_n}(x', t) dt.$$

Cauchyjeva nejednakost daje

$$|v(x', x_n)|^2 \leq (x_n - a) \int_a^{x_n} \left| \frac{\partial v}{\partial x_n}(x', t) \right|^2 dt,$$

pa integriranjem po prvih $n - 1$ varijabli slijedi

$$\begin{aligned} \int_{\mathbb{R}^{n-1}} |v(x', x_n)|^2 dx' &\leq (x_n - a) \int_a^{x_n} \int_{\mathbb{R}^{n-1}} \left| \frac{\partial v}{\partial x_n}(x', t) \right|^2 dx' dt. \\ \int_a^b \int_{\mathbb{R}^{n-1}} |v(x', x_n)|^2 dx' dx_n &\leq \frac{1}{2}(b - a)^2 \int_a^b \int_{\mathbb{R}^{n-1}} \left| \frac{\partial v}{\partial x_n}(x', t) \right|^2 dx' dt. \end{aligned}$$

Budući da je funkcija jednaka nuli izvan Ω slijedi

$$\int_{\Omega} |v(x)|^2 dx \leq \frac{1}{2}(b - a)^2 \int_{\Omega} \left| \frac{\partial v}{\partial x_n}(x) \right|^2 dx. \quad \square$$

Promjenom Poincaréove nejednakosti dobivamo sljedeću ekvivalenciju normi na $H_0^1(\Omega)$:

$$\forall u \in H_0^1(\Omega), \quad \frac{1}{\sqrt{1 + c^2}} \|u\|_{1,2} \leq |u|_{1,2} \leq \|u\|_{1,2}.$$

Zadatak A.3 Dokažite teorem u slučaju $p \neq 2$. \square

Napomena A.3 Poincaréova nejednakost (A.9) vrijedi i za funkcije koje se poništavaju samo na dijelu granice domene Ω . Dokaz te tvrdnje više nije konstruktivan kao dokaz Teorema A.4, već, na primjer, treba koristiti metodu od protivnog i pojam slabe konvergencije.

A.4.1 Singulariteti $W^{1,p}$ funkcija

Funkcije iz prostora L^p mogu biti neograničene (za $p < \infty$). Moraju li funkcije is $W^{1,p}$ biti ograničene ovisi o odnosu dimenzije prostora i indeksa p . Za $n = 1$ imamo ovaj rezultat

Teorem A.5 $W^{1,p}(a, b) \subset C([a, b])$ i vrijedi

$$u(x) - u(y) = \int_x^y u'(z) dz.$$

Sve funkcije iz $W^{1,p}(a, b)$ su neprekidne na segmentu $[a, b]$, pa su stoga i ograničene. Već u dvije dimenzije ovo svojstvo se gubi.

Zadatak A.4 Pokažite da funkcija

$$u(x, y) = \ln \ln \frac{2}{r}, \quad r^2 = x^2 + y^2,$$

pripada prostoru $H^1(\Omega)$, $\Omega = B(0, 1) \subset \mathbb{R}^2$.

Zadatak A.5 Pokažite da funkcija

$$u(x) = r^{-\alpha}, \quad r = |x|,$$

pripada prostoru $H^1(\Omega)$, $\Omega = B(0, 1) \subset \mathbb{R}^n$, $n > 2$, za $\alpha < (n - 2)/2$.

Iz ovih primjera vidimo da funkcije sa slabom derivacijom ne moraju biti niti neprekidne niti ograničene.

Teorem A.6 Za $1 \leq p \leq \infty$ i bilo koju domenu $\Omega \subset \mathbb{R}^n$ vrijedi

- Za $1 \leq p < n$ vrijedi $W_0^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ za $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$;
- Za $p = n$ vrijedi $W_0^{1,p}(\Omega) \subset L^q(\Omega)$ za sve $q \in [p, \infty)$;
- Za $p > n$ vrijedi $W_0^{1,p}(\Omega) \subset C(\overline{\Omega})$ i štoviše imamo

$$|u(x) - u(y)| \leq C \|u\|_{1,p} |x - y|^\alpha$$

gdje je $\alpha = 1 - n/p$ i $C = C(n, p, \Omega)$.

Isti teorem vrijedi i za funkcije iz $W^{1,p}(\Omega)$, ako se na Ω stave dodatne pretpostavke o glatkoći granice.

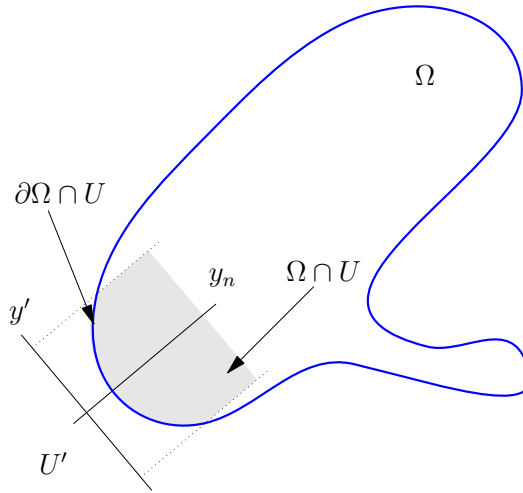
A.4.2 Regularnost granice

Uzet ćemo radi jednostavnosti da je domena Ω ograničena, što je za naše potrebe sasvim dovoljno.

Definicija A.7 Ograničena domena $\Omega \subset \mathbb{R}^n$ je Lipschitzova (klase C^k) ako za svaku točku $x \in \partial\Omega$ postoji okolina U i sustav ortogonalnih koordinata $y = (y', y_n)$, gdje je $y' = (y_1, \dots, y_{n-1})$, takvi da vrijedi:

1. U novim koordinatam U je hiperkocka

$$U = \{y \in \mathbb{R}^n : -a_j < y_j < a_j, j = 1, 2, \dots, n\}.$$



Slika A.1: Lokalna parametrizacija granice

2. Postoji Lipschitzova (klase C^k) funkcija ϕ definirana na

$$U' = \{y' \in \mathbb{R}^n : -a_j < y_j < a_j, j = 1, 2, \dots, n-1\},$$

koja zadovoljava

$$\forall y' \in U', \quad |\phi(y')| \leq a_n/2$$

$$\Omega \cap U = \{y : y_n > \phi(y')\}, \quad \partial\Omega \cap U = \{y : y_n = \phi(y')\}.$$

Ova definicija znači da se lokalno granica skupa može prikazati kao graf Lipschitzove funkcije, te da se skup lokalno nalazi samo s jedne strane granice. Uočimo još da zbog ograničenosti domene, njena granica je kompaktan skup te stoga možemo uvijek odabrati konačan skup ovakvih karata koje pokrivaju cijelu granicu.

U Lipschitzovoj domeni vrjede teoremi ulaganja za $W^{1,p}$ prostor.

Teorem A.7 Za $1 \leq p \leq \infty$ i bilo ograničenu Lipschitzovu domenu $\Omega \subset \mathbb{R}^n$ vrijedi

- Za $1 \leq p < n$ vrijedi $W^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ za $\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}$;
- Za $p = n$ vrijedi $W^{1,p}(\Omega) \subset L^q(\Omega)$ za sve $q \in [p, \infty)$;
- Za $p > n$ vrijedi $W^{1,p}(\Omega) \subset C(\bar{\Omega})$ i štoviše imamo

$$|u(x) - u(y)| \leq C \|u\|_{1,p} |x - y|^\alpha$$

gdje je $\alpha = 1 - n/p$ i $C = C(n, p, \Omega)$.

Kada je rub domene dovoljno gladak (Lipschitzov) onda se mogu definirati L^p prostori na rubu $\partial\Omega$. Pri tome vrijedi:

Teorem A.8 (Teorem o tragu) Neka je $\Omega \subset \mathbb{R}^n$ ograničena Lipschitzova domena. Tada postoji konstanta $C > 0$ takva da je

$$\forall v \in W^{1,p}(\Omega), \quad \|v\|_{L^p(\partial\Omega)} \leq C\|v\|_{1,p}.$$

Napomena o literaturi. Prostori diferencijabilnih funkcija i prostori Soboljeva obrađeni su na jednostavan način u [3]. Više detalja se može naći u [4] i [2]. Osnovne definicije i rezultati se mogu naći u svakoj knjizi iz konačnih elemenata (npr, [1]).

Bibliografija

- [1] Alberto Valli Alfio Quarteroni. *Numerical Approximation of Partial Differential Equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer, Berlin, 1997.
- [2] H. Brezis. *Analyse Fonctionnelle*. Masson, Paris, 1983.
- [3] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 1998.
- [4] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 1998.
- [5] Darko Veljan. *Kombinatorika s teorijom grafova*. Školska knjiga, Zagreb, 1988.

A.5 Zadaci

1. Dokažite Taylorovu formulu (A.4).

Bibliografija

- [1] Alberto Valli Alfio Quarteroni. *Numerical Approximation of Partial Differential Equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer, Berlin, 1997.
- [2] H. Brezis. *Analyse Fonctionnelle*. Masson, Paris, 1983.
- [3] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 1998.
- [4] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 1998.

- [5] Darko Veljan. *Kombinatorika s teorijom grafova*. Školska knjiga, Zagreb, 1988.