

# Numerikus Módszerek III (második, javított kiadás)

Stoyan Gisbert (szöveg), Takó Galina (programok)

Magyarország, 2008 november

ELTE Informatikai Kar 2008

Az 1997-es kiadás a Typotex kiadóban készült,  
annak lektora volt Gáspár Csaba, Szent István Egyetem

### Előszó az 1. kiadáshoz

Ezzel a könyvvel folytatjuk az előző két kötet vonalát az alkalmazások szempontjából különösen fontos parciális differenciálegyenletek terén. A másodrendű egyenletekkel kapcsolatos differencia, véges térfogat és végeelem módszereket elég részletesen ismertetjük. Sok feladat és két program (amely az elliptikus, parabolikus és hiperbolikus egyenletekkel kapcsolatos), valamint kiadásos jegyzékek egészítik ki a könyvet.

A szerzők hálások az ELTEnek a kiadvány anyagi támogatásáért és köszönik a Typotex kiadó munkáját. Különös köszönetet mondunk Tikk Domonkos műszaki szerkesztőnknek, aki mindhárom kötetet gondozta, valamint Gáspár Csaba lektorunknak, aki konstruktív kritikájával (és a szöveg magyartalan fogalmazásainak csökkentésével) segítségünkre volt.

A hiányzó érdekes témák és módszerek (pl. integrálegyenletek, peremintegrál- és Lagrange-féle módszerek stb.) ismertetésére a stafettabotot neki adjuk át.

Reméljük, hogy a jelenlegi könyv is hozzájárul a magyar egyetemek és főiskolák oktatásához, a doktori képzéshez, valamint az ország gazdasági és műszaki fejlődéséhez.

Stoyan Gisbert, Takó Galina

Budapest, 1996. szeptember 13.

### Előszó a 2. kiadáshoz

Az ELTE Informatikai Kar jegyzet-támogatási pályázat jóvoltából készülhetett ez a 2. (internetes) kiadás – az első kiadás után 12 évvel. Ezalatt az idő alatt nyilván folytatódott a kutatási és oktatási munkám a parciális differenciálegyenletek numerikus megoldásának terén. Ennek megfelelően most már egy sok helyen javított szöveget tudok nyújtani : a teljes könyvből az itt inkludált szöveg (15., 16. fejezet, 17. fejezet eleje) nyelvtani és matematikai hibákra van korrigálva (kevésbé a feladatok és a \*\*\* helyek, ahol pl. hivatkozások vannak a 17-18. fejezetre). Az irodalmat is felújítottam. Különösen a végeelem konvergencia-elméletén (15.7.5. pont) és a parabolikus esetben a másod- és harmadrendű peremfeltételekre vonatkozó becsléseken (16.4.5. pont) dolgoztam.

Sajnálom, hogy a sorozat 2. kötete (közönséges differenciálegyenletekről : Numerikus módszerek II, Typotex 1995) felújítására nem volt lehetőség, míg az első kötet 3. kiadása már ingyenesen olvasható az interneten.

Stoyan Gisbert

Budapest, 2008. november 14.

Takó Galina programjai a 3. kötethez :

8. Ellipsz (kétdimenziós elliptikus differenciálegyenlet megoldása többrácsos módszerrel)
9. Hipy (parabolikus és hiperbolikus egyenlet megoldása súlyozott differenciasémával)

# Tartalomjegyzék

<b>15</b>	<b>Elliptikus egyenletek</b>	<b>9</b>
15.1	Bevezetés: Elliptikus egyenletek alkalmazott feladatokban . . .	9
15.2	Elméleti háttér . . . . .	11
15.3	Véges differencia eljárások II . . . . .	16
15.3.1	A Poisson–egyenlet approximációja differenciasémával .	17
15.3.2	A diszkrét maximumelv . . . . .	21
15.3.3	A diszkrét Poisson–egyenlet megoldása . . . . .	32
15.3.4	Harmadfajú peremfeltételek . . . . .	37
15.3.5	A Poisson–egyenlet általános tartományban . . . . .	43
15.4	A többrácsos módszer . . . . .	49
15.4.1	Az alapötlet . . . . .	49
15.4.2	A simító iterációk . . . . .	51
15.4.3	Alapvető többrácsos algoritmusok . . . . .	54
15.4.4	A többrácsos iterációk művelet- és tárigénye . . . . .	58
15.4.5	A kétrácsos módszer konvergenciája . . . . .	61
15.4.6	Az egydimenziós eset . . . . .	63
15.4.7	A simítási és az approximációs tulajdonság . . . . .	70
15.4.8	A többrácsos módszer konvergenciája . . . . .	79
15.4.9	Nemlineáris egyenletek megoldása . . . . .	83
15.4.10	Befejező megjegyzések . . . . .	86
15.5	Differencia-approximációk; kiegészítések . . . . .	87
15.5.1	A „diszkrét” Green-féle függvény becslése . . . . .	87
15.5.2	Változó együtthatójú differenciáloperátorok . . . . .	92
15.5.3	„Diszkrét” beágyazási tételek . . . . .	97
15.5.4	Differenciasémák vizsgálata $L_2$ -beli jobboldal esetén . .	100
15.5.5	Dirac-féle $\delta$ -függvényt tartalmazó jobboldal . . . . .	107
15.6	Véges térfogat módszer . . . . .	112
15.6.1	Bevezetés . . . . .	112
15.6.2	A diszkretizáció levezetése . . . . .	114
15.6.3	A véges térfogat módszer háromszögek esetén . . . . .	121

15.6.4	A véges térfogat módszer mátrixai . . . . .	126
15.6.5	Speciális kérdések . . . . .	129
15.7	A végeelem módszer II . . . . .	131
15.7.1	Bevezetés . . . . .	131
15.7.2	A peremfeltételekről . . . . .	136
15.7.3	A variációs feladat megoldhatósága . . . . .	139
15.7.4	Véges elemek 2- és 3-dimenziós feladatokban . . . . .	142
15.7.5	A végeelem módszer pontossága . . . . .	151
15.7.6	Numerikus integrálás . . . . .	160
15.7.7	Algoritmusok . . . . .	166
15.7.8	Modellfeladat végeelem megoldása . . . . .	170
15.7.9	Végelem és véges térfogat módszer . . . . .	177
15.7.10	Az izoparametrikus módszer *** 15.7.9. . . . .	181
15.7.11	Rácsszerkesztés *** 15.7.10. . . . .	185
15.8	A lineáris rendszerek megoldása . . . . .	190
15.8.1	Beágyazott végeelem terek . . . . .	190
15.8.2	A végeelem alapú többrácsos módszer konvergenciája *** 15.8.1. . . . .	193
15.8.3	A szimmetrikus eset . . . . .	197
15.8.4	Tartomány dekompozíciós és párhuzamos módszerek, bevezetés . . . . .	203
15.8.5	Alstruktúra-eljárás . . . . .	204
15.8.6	Dirichlet tartomány dekompozíciós módszer . . . . .	208
15.9	Elliptikus sajátérték feladatok . . . . .	218
15.9.1	Elméleti háttér . . . . .	219
15.9.2	A diszkrét Laplace-operátor sajátérték feladata . . . . .	223
15.9.3	Gradiens módszerek az általánosított sajátérték feladat megoldására . . . . .	229
15.9.4	A Rayleigh–Ritz- és a Lánczos-módszer . . . . .	239
15.10	Összefoglalás . . . . .	241
15.11	Feladatok . . . . .	242
<b>16</b>	<b>Parabolikus egyenletek</b>	<b>261</b>
16.1	Parabolikus egyenletek eredete . . . . .	261
16.1.1	Környezetvédelmi balesetek . . . . .	261
16.1.2	A Black–Scholes egyenlet . . . . .	262
16.1.3	A hővezetési egyenlet . . . . .	263
16.2	Elméleti tudnivalók . . . . .	265
16.3	Néhány hasznos fogás . . . . .	270
16.4	A súlyozott differenciaséma . . . . .	273
16.4.1	A súlyozott differenciaséma . . . . .	273

16.4.2	A súlyozott differenciaséma képlethibája . . . . .	277
16.4.3	A súlyozott differenciaséma stabilitása a szimmetrikus esetben . . . . .	279
16.4.4	A súlyozott differenciaséma konvergenciája . . . . .	288
16.4.5	Más peremfeltételek . . . . .	289
16.4.6	A végeelem séma . . . . .	297
16.4.7	Pozitivitástartás, maximumelv, konvergencia a maxi- mum normában . . . . .	301
16.4.8	A tömegmátrix kiszámításáról . . . . .	309
16.4.9	Vizsgálatok Fourier-módszer segítségével . . . . .	310
16.4.10	Stabilitás a nemszimmetrikus esetben . . . . .	317
16.4.11	További súlyozott differenciasémák . . . . .	319
16.5	Változó együtthatójú egyenletek . . . . .	320
16.5.1	Egydimenziós hővezetési egyenlet helytől függő együtt- hatóval . . . . .	321
16.5.2	Időtől is függő együttható esete . . . . .	323
16.5.3	A hővezetési egyenlet hengerszimmetriában . . . . .	325
16.5.4	A hővezetési egyenlet gömbszimmetriában . . . . .	332
16.6	Nemlineáris parabolikus egyenletek . . . . .	337
16.7	Többdimenziós parabolikus egyenletek . . . . .	345
16.7.1	A kétdimenziós eset : Peaceman–Rachford módszer . . . . .	346
16.7.2	A Peaceman–Rachford módszer mint simító eljárás . . . . .	354
16.7.3	Kettőnél több dimenziós parabolikus egyenletek megol- dása . . . . .	358
16.7.4	Többdimenziós egyenletek többrácsos megoldása . . . . .	363
16.8	Variációs eljárás . . . . .	364
16.8.1	Szemidiszkretizáció . . . . .	364
16.8.2	Teljes diszkretizáció . . . . .	372
16.9	Összefoglalás . . . . .	375
16.10	Feladatok . . . . .	375
<b>17</b>	<b>A Navier–Stokes egyenletek</b>	<b>387</b>
17.1	Bevezetés: fizikai háttér . . . . .	387
17.2	Numerikus problémák, az áramfüggvény . . . . .	393
17.2.1	Az áramfüggvény peremértékei és létezése . . . . .	395
17.2.2	Az $\omega - \psi$ -rendszer . . . . .	402
17.2.3	$\omega$ és $p$ peremfeltételei . . . . .	404
17.3	Lassú áramlások számítása . . . . .	409
<b>18</b>	<b>Kiegészítő információk, jelölések III</b>	<b>413</b>

<b>19 Irodalom III</b>	<b>417</b>
<b>20 Tárgymutató III</b>	<b>435</b>
20.1 Címszavak jegyzéke . . . . .	435
20.2 Tételék, lemmák jegyzéke . . . . .	444
20.3 Pszeudokódos algoritmusok jegyzéke . . . . .	449
20.4 Táblázatok jegyzéke . . . . .	449



# Fejezet 15

## Elliptikus egyenletek numerikus megoldása

### 15.1 Bevezetés: Elliptikus egyenletek alkalmazott feladatokban

Elliptikus egyenletek gyakran akkor szerepelnek fizikai jelenségek matematikai modelljeiben, ha eltekinthetünk az időtől, ha stacionárius állapotról van szó. Példa ilyen fizikai jelenségre az (időtől függő) hővezetés, amelynek egyenlete

$$c\rho\frac{\partial u}{\partial t} = \operatorname{div}(k \operatorname{grad} u) + f. \quad (15.1)$$

Itt  $u$  a hőmérséklet,  $c$  a fajhő,  $\rho$  a hővezető közeg sűrűsége,  $k$  a hővezetési tényező,  $f$  a hőforrások sűrűsége. Mind  $u, c, \rho, k, f$  az  $x$  és  $t$  függvényei. Az egyenlethez megfelelő kezdeti és peremfeltételeket is kell megadni (ld. részletesebben a 16.1. pontot). Amennyiben pl.  $c, \rho, k, f$  és a peremfeltételek nem függenek az időtől,  $t \rightarrow \infty$  esetén az  $u(x, t)$  megoldás tarthat (fizikailag: ha egyensúly áll be a hővezetési folyamatban) a következő egyenlet  $u(x)$  megoldásához:

$$0 = \operatorname{div}(k \operatorname{grad} u) + f. \quad (15.2)$$

Ezt *stacionárius hővezetési egyenletnek* hívjuk.

(15.1)-ben és (15.2)-ben a következő jelöléseket használtuk:

$$\operatorname{grad} u := \left( \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \frac{\partial u}{\partial x_3} \right)^T, \quad x := (x_1, x_2, x_3)^T,$$

és

$$\operatorname{div} w := \sum_{i=1}^3 \frac{\partial w_i}{\partial x_i}, \quad w := (w_1, w_2, w_3)^T.$$

A (15.1) ill. (15.2) *hővezetési egyenlettel* több más folyamat is leírható : így koncentráció eloszlást ír le (ha  $u$  a koncentráció,  $k$  a diffúziós tényező és  $f$  az anyagforrások sűrűsége), vagy pl. az áramlást porózus közegben. Ide tartozik a víz, valamint a kőolaj és földgáz mozgása a föld rétegeiben. Ekkor (15.2) a *szivárgási egyenlet* és benne  $u$  a nyomás, a  $k$  együttható a porózus közeg áteresztőképességének és a fluidum vizskóztításának a hányadosa, és mint ilyen a helykoordinátáktól függ (közvetlenül az áteresztőképesség révén, de közvetett módon a helyben változó hőmérsékleten keresztül is, amely a fluidum viszkozitását befolyásolja). Ezt említettük I. 4.1-ben, ahol az együtthatót akkor  $a(x, y, z)$ -vel jelöltük.

Az a tény, hogy mind a hővezetés, mind a porózus közegbeli áramlás a (15.2) egyenlettel leírható, azzal függ össze, hogy mindkét esetben először egy  $\operatorname{div} J = f$  alakú megmaradási tételre lehet hivatkozni, majd a hővezetési folyamatnál a Fick-törvény szerint, a porózus közeg esetén a Darcy-törvény szerint igaz  $J = -k \operatorname{grad} u$ . Mindkét törvény egy hasonló feltételezéssel kapcsolatos : az áram (a hőáram ill. a szűrődési sebesség) egy vékony rétegen keresztül arányos a hőmérsékletnek ill. nyomásnak a réteg két oldalán mért különbségével. Míg ez a különbség kicsi, a feltételezés jól adja vissza a tényleges viszonyokat.

Speciális esetben, ha a közeg homogénitása és a hővezetés izotrópiája miatt  $k = \operatorname{const}$ , (15.2)-ből kapjuk a legalapvetőbb elliptikus egyenletet, a *Poisson-egyenletet* :

$$\Delta u + f(x) = 0, \quad (15.3)$$

ahol

$$\Delta u := \sum_{i=1}^3 \frac{\partial^2 u}{\partial x_i^2}$$

a (háromdimenziós) Laplace-operátor.

A (15.3) egyenlettel máskor is találkozhatunk. Így a kétdimenziós, örvénymentes állandósult áramlás  $(u, v)^T$  sebességvektorának mindkét komponense eleget tesz a *Laplace-egyenletnek* (ehhez ld. a 17.2.1. pontot is) :

$$\Delta u = 0 \quad \text{és} \quad \Delta v = 0, \quad \text{ahol} \quad \Delta u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$$

a kétdimenziós Laplace-operátor.

A (15.3) egyenletnek még többféle fizikai értelmezése is van. Amikor (15.2)-ben  $k$  a dielektrikus tényező és  $f$  a töltések sűrűségét jelenti, akkor  $u$  adja az elektrosztatikus mező potenciálját. Ezért (15.3)-at potenciálegyenletnek is hívjuk. Amikor  $k$  a mágneses permeabilitás és  $f$  az áramsűrűség, akkor  $u$  lesz a mágneses tér potenciálja.

Parabolikus egyenleteket – és ezekhez tartozik (15.1), mivel fizikai okokból teljesül  $c\rho > 0$ ,  $k > 0$ , ld. a 16.1.3. és 16.2. pontokat – a numerikus megoldás céljából visszavezethetünk elliptikus egyenletekre : Approximálva (15.1)-ben az időderiváltat :

$$\left. \frac{\partial u}{\partial t} \right|_{t=t_n} \approx \frac{u^n - u^{n-1}}{\tau},$$

ahol  $u^n := u(x, t_n)$ ,  $t_n = n \cdot \tau$ ,  $\tau$  az időlépés, (15.1)-ből azt kapjuk, hogy

$$0 = \operatorname{div}(k \operatorname{grad} u^n) - \frac{c\rho}{\tau} u^n + f + \frac{c\rho}{\tau} u^{n-1},$$

ami megint egy elliptikus egyenlet  $u^n$ -re nézve.

Az elmondottakból világos, hogy az elliptikus egyenletek és különösen a Poisson–egyenlet megoldása alapvetően fontos feladat.

A Poisson–egyenlet segítségével vékony rudak csavarása is modellezhető :

$$\Delta \Phi = -1, \quad (x_1, x_2) \in \Omega, \quad \Phi|_{\Gamma} = 0. \quad (15.4)$$

Itt  $\Omega$  a rúd metszete, és  $\Gamma$  az  $\Omega$  tartomány pereme. A  $\Phi$  segédfüggvényből az eltolódások  $(u_1, u_2, u_3)$  vektorát kapjuk meg, feltéve, hogy  $\tau$ , az egység-hosszra vonatkoztatott csavarási szög, a rúd hosszának irányában konstans :

$$u_1 = -\tau x_3 x_2, \quad u_2 = \tau x_3 x_1, \quad \frac{\partial u_3}{\partial x_1} = 2\tau \frac{\partial \Phi}{\partial x_2} + \tau x_2, \quad \frac{\partial u_3}{\partial x_2} = -2\tau \frac{\partial \Phi}{\partial x_1} - \tau x_1.$$

A csavarásnak a térfogatra vonatkoztatott energiájára az

$$E = 2\mu\tau^2 |\operatorname{grad} \Phi|^2$$

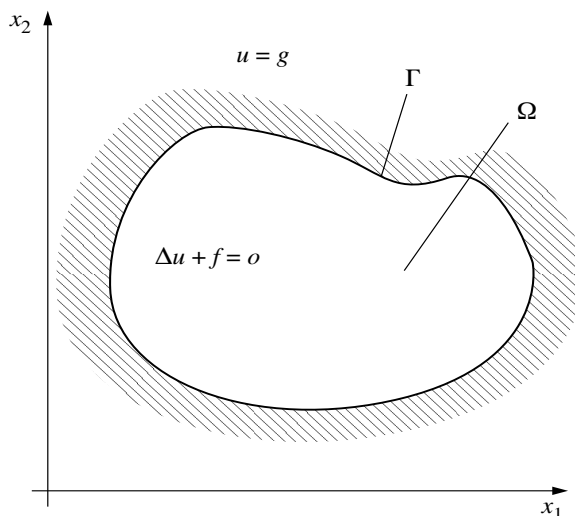
kifejezést lehet levezetni, ahol  $\mu$  az eltolódási modulus. Ekkor a rúd egység-hosszára vonatkoztatott csavarási energiája

$$\mathcal{E} = 2\mu\tau^2 \int_{\Omega} |\operatorname{grad} \Phi|^2 dx.$$

## 15.2 Elméleti háttér

Az elliptikus feladatokra tipikus a peremérték feladat, ld. Simon L. és E.A. Baderko könyvét.

Az adott  $\Omega$  tartomány  $\Gamma$  peremén meg kell adnunk a keresett megoldás értékét, a tartomány belsejében érvényes a differenciálegyenlet, és ez határozza meg a megoldást mindenütt. Megfelelő feltételek mellett az ilyen feladat megoldása létezik, egyértelmű és folytonosan függ a feladat adataitól (a



ábra 15.1: *Elliptikus peremérték feladat*

peremtől, a peremértékektől, a differenciálegyenlet együtthatóitól valamint a jobboldaltól). Ha mind ez a három tulajdonság (a létezés, az unicitás, a stabilitás) megvan, akkor azt mondjuk, hogy *korrekt kitűzésű* a peremérték feladat.

Mindjárt arra mutatunk rá, hogy a későbbi fejezetekben tárgyalásra kerülő parabolikus és hiperbolikus egyenleteknél (ld. a 16. ill. 18. fejezetet) a peremérték feladat nem korrekt kitűzésű : ezeknél az egyenleteknél nem szabad előírni a megoldást a megoldási tartomány egész peremén.

Annak érdekében, hogy az elliptikus peremérték feladat korrekt kitűzésűségét konkrétan kimondhassuk, most részletesebben írjuk le az ehhez elégséges feltételeket.

Legyen  $\Omega$  korlátos tartomány  $\mathbb{R}^d$ -ben, annak  $\Gamma$  pereme legyen *Lipschitz-folytonos* – egy fogalom, amellyel e könyvben többször találkozhatunk. Azt jelenti, hogy a peremnek van olyan véges lefedése nyílt  $\mathbb{R}^d$ -beli környezetekből, amelynek mindegyikében, alkalmasan forgatott lokális koordinátarendszerben, a perem egyenlete  $x_d = \phi(x_1, \dots, x_{d-1})$ , ahol  $\phi : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  Lipschitz-folytonos.

Ilyen peremmel rendelkeznek pl. a ( $d$ -dimenziós) gömb és kocka (gondoljunk arra, hogy kétdimenziós esetben a lokális koordinátarendszerben a perem egyenlete  $x_2 = \phi(x_1)$  alakú, és így  $x_2 = |x_1|$  jöhet szóba). Nem Lipschitz-folytonos viszont pl. a végtelen tartomány vagy az, amelynek peremén található egy zérus szöget alkotó (belső vagy külső) sarok (pl. az  $x_2 = \sqrt{|x_1|}$  vonal az origó környezetében nem Lipschitz-folytonos).

Az elliptikus differenciáloperátor legyen

$$\mathcal{L}u := - \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(x) \frac{\partial u}{\partial x_i} + c(x)u, \quad (15.5)$$

ahol  $c, b_i, a_{ij} \in C(\overline{\Omega})$ . Feltehetjük, hogy az  $a_{ij}$  együtthatókból összeállított  $A$  mátrix szimmetrikus legyen. A (15.5) operátort akkor hívjuk *egyenletesen elliptikusnak*  $\Omega$ -ban, ha az  $A = A(x)$  mátrix egyenletesen pozitív definit, azaz ha az euklideszi skalárszorzatban pozitív  $k_0$  konstanssal teljesül

$$(A(x)\xi, \xi) \geq k_0|\xi|^2, \text{ minden } \xi \in \mathbb{R}^d\text{-re és minden } x \in \Omega\text{-ra.} \quad (15.6)$$

Egyenletesen elliptikus operátor például  $\mathcal{L}u = -\Delta u$ , a Laplace-operátorral, hiszen akkor  $A = I$  és így (15.6) teljesül  $k_0 = 1$ -gyel. Az ilyen operátorral, továbbá adott  $f : \Omega \rightarrow \mathbb{R}$  és  $g : \Gamma \rightarrow \mathbb{R}$  függvényekkel, tekintsük a következő feladatot :

$$(\mathcal{L}u)(x) = f(x), \quad x \in \Omega; \quad (15.7)$$

$$u(x) = g(x), \quad x \in \Gamma. \quad (15.8)$$

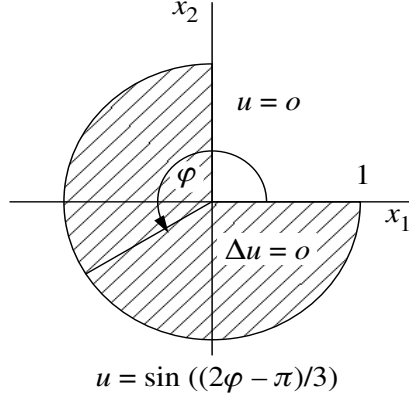
**Definíció.** Azt mondjuk, hogy ennek a feladatnak van *klasszikus megoldása*, ha van olyan  $u \in C^2(\Omega) \cap C(\overline{\Omega})$  függvény, amely a (15.7) egyenletet és a (15.8) peremfeltételt kielégíti.  $\square$

Az előzőekben említett eredmény az elliptikus peremérték feladatról a következő.

**15.1. Tétel** (J.H. Michael; egyenletesen elliptikus elsőfajú peremérték feladat korrekt kitűzésű). Legyen  $\Omega \subset \mathbb{R}^d$  Lipschitz-folytonos tartomány és  $\mathcal{L}$  egyenletesen elliptikus operátor, amelynek együtthatói folytonosak  $\Omega$ -ban. Ezek közül legyen  $c$  nemnegatív. Végül, legyen  $f$  is  $\Omega$ -n definiált, folytonos függvény és  $g$  definiált és folytonos az  $\Omega$  tartomány  $\Gamma$  peremén. Ekkor a (15.7), (15.8) peremérték feladat klasszikus megoldása létezik, egyértelmű és folytonosan függ a feladat adataitól.  $\square$

**Megjegyzés.** A  $c$  együttható nemnegativitása azt biztosítja, hogy nem az operátor spektrumán oldjuk meg a (15.7) feladatot.  $\square$

A klasszikus megoldás fenti definíciójában az  $u \in C^2(\Omega) \cap C(\overline{\Omega})$  követelmény  $u \in C^2(\overline{\Omega})$  helyett nem véletlen. Ezzel azt fejezzük ki, hogy a második deriváltakra csak a tartomány belsejében van szükség, míg a peremfeltételek miatt  $u \in C(\overline{\Omega})$ -ra számítunk. A (15.4) csavarási probléma adja a legegyszerűbb példát arra, hogy a (15.7), (15.8) megoldása még akkor sem feltétlenül



ábra 15.2:  $\Delta u = 0$  szinguláris megoldása

$u \in C^2(\bar{\Omega})$ , amikor  $f$  és  $g$  konstansok, vagyis tetszőlegesen sokszor differenciálhatók, de  $\Omega$  az  $x_1, x_2$ -sík egységnyezete (a lényeg az, hogy a peremnek derékszögű sarokpontja van, de Lipschitz-folytonos) :

Ekkor egyrészt tetszőleges sarokpontban minden  $\{(\Delta u)(x^{(m)})\}$  sorozat határértéke  $x^{(m)} \in \Omega$  esetén  $-1$ , másrészt az  $x_i$ -tengellyel párhuzamos peremen  $\frac{\partial^2 u}{\partial x_i^2} = 0$ ,  $i = 1, 2$ , ami szerint sarokpontban

$$\frac{\partial^2 u}{\partial x_1^2} = \frac{\partial^2 u}{\partial x_2^2} = (\Delta u)(x) = 0$$

kellene teljesülni.

Még arra sem számíthatunk általában, hogy legalább  $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$  legyen. Tekintsük ugyanis a Laplace-egyenlet következő elsőfajú peremérték feladatát :  $\Omega$  tehát az egységkör, amelyből a pozitív kvadransba eső rész ki van vágva, a peremérték  $u(x) = 0$  a nemnegatív  $x_1$ - és  $x_2$ -tengelyen, míg a perem többi részén a peremfeltétel, polárkoordinátákban,  $u(1, \varphi) := \sin((2\varphi - \pi)/3)$ . Az így megadott peremfüggvény folytonos. Vegyük most figyelembe, hogy a Laplace-operátor polárkoordinátákban felírva

$$\Delta_{r,\varphi} u = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} \quad (15.9)$$

alakú, akkor közvetlenül ellenőrizhető, hogy  $u(r, \varphi) := r^{2/3} \sin((2\varphi - \pi)/3)$  a Laplace-egyenlet megoldása – és a 15.1. tétel szerint az egyetlen – de első deriváltja nem folytonos az origóban.

Ezek a példák azért érdekesek numerikus módszerekkel kapcsolatban, mert a differenciasémák klasszikus kezelése (ld. 15.3. pont) olyan feltétellel történik, hogy teljesüljön pl.  $u \in C^4(\bar{\Omega})$ . Erre tehát korántsem számíthatunk

mindig, még az olyan a gyakorlatban előforduló egyszerű példánál sem, amelyen a (15.4) csavarási feladat, és ez indokolja a differenciasémák modernebb felépítésének a bemutatását, ld. 15.5.4. és 15.5.5. pontokat.

A példák mutatják, hogy a folytonos függvények tereivel nem igaz az, hogy  $f \in C(\bar{\Omega})$  esetén  $u \in C^2(\bar{\Omega})$ .

Viszont a  $H^k(\Omega)$  Szoboljev-terekkel (az általánosított, négyzetesen integrálható  $0 - k$ -adik deriváltakkal rendelkező függvények tereivel) ehhez közelebbre lehet jutni. Ott igaz az, hogy  $f \in H^0(\Omega) = L_2(\Omega)$  esetén  $u \in H^2(\Omega)$ , ld. 15.7.3-ban a 15.30. tételt. Ezt az irányt használjuk a véges elem módszernél. Speciálisan az  $u \in H^2(\Omega)$  már elegendő az elsőrendű konvergencia bizonyítására.

Az elliptikus egyenlet további jellegzetes tulajdonsága, hogy (klasszikus) megoldása eleget tesz a *Hopf-féle maximumelvnek*: ha  $\mathcal{L}$  az  $\Omega$  tartományban egyenletesen elliptikus operátor és  $u \in C^2(\Omega) \cap C(\bar{\Omega})$  a (15.7),  $f \equiv 0$  megoldása, akkor  $u$  veszi fel minimumát és maximumát a tartomány peremén, azaz

$$\min_{x \in \Gamma} g(x) = \min_{x \in \Gamma} u(x) \leq u(x) \leq \max_{x \in \Gamma} u(x) = \max_{x \in \Gamma} g(x) \quad (15.10)$$

minden  $x \in \bar{\Omega}$ -ra.

A maximumelv közel áll a termodinamika második főtételehez (durván megfogalmazva: a hő nem áramlik a hideg területekről a melegebbek felé). Így nem lehetséges, hogy egy test állandósult hőeloszlásának maximális vagy minimális hőmérsékletét annak belsejében találjunk, amikor nincsenek belső hőforrások.

A (15.7) feladat speciális  $\mathcal{L}u = -\Delta u = 0$ ,  $x \in \Omega$ ,  $u|_{\Gamma} = g$  esete egy hőforrások nélküli homogén test  $u$  hőeloszlását adja meg a  $g$  felületi hőmérséklet ismeretében. Ekkor a maximumelv azt mondja, hogy a hőmérséklet maximuma és minimuma a  $\Gamma$  felületén található.

A maximumelvvel szoros kapcsolatban áll az *összehasonlítási tétel*: amennyiben a fenti (15.5)-(15.6) egyenletesen elliptikus  $\mathcal{L}$  operátorral és  $u, v \in C^2(\Omega) \cap C(\bar{\Omega})$  függvényekkel érvényes

$$(\mathcal{L}u)(x) \leq (\mathcal{L}v)(x), \quad \text{minden } x \in \Omega\text{-ra,} \quad (15.11)$$

$$u(x) \leq v(x), \quad \text{minden } x \in \Gamma\text{-ra,} \quad (15.12)$$

akkor következik  $u(x) \leq v(x)$  minden  $x \in \bar{\Omega}$ -ra. Ezzel az összehasonlítási tétellel durva becsléseket lehet levezetni a megoldás értéktartományára. Legyen  $\mathcal{L} = -\Delta$  és  $u$  például a (15.4) egyenlet megoldása abban a fent említett esetben, amikor  $\Omega$  az egységnyégzet. Ekkor  $v \equiv 0$ ,  $w(x) := x(1-x) +$

$y(1-y)$  esetén

$$\begin{aligned} 0 &= (\mathcal{L}v)(x) \leq (\mathcal{L}u)(x) = 1 = \frac{1}{4}(\mathcal{L}w)(x), \quad x \in \Omega, \\ 0 &= v(x) \leq u(x) \leq \frac{1}{4}w(x), \quad x \in \Gamma \end{aligned}$$

miatt következik tehát  $0 \leq u(x) \leq [x(1-x)+y(1-y)]/4$  és ebből  $\|u\|_{C(\bar{\Omega})} \leq \frac{1}{8}$ .

Fent a másodrendű parciális differenciálegyenletek hagyományos felosztását elliptikus, parabolikus és hiperbolikus egyenletekre említettük meg. Vannak olyan –gyakorlatilag fontos – másodrendű egyenletek is, amelyeknek típusa változik a tartomány belsejében. Ilyen a transzsonikus (hangsebesség körüli) áramlást leíró úgynevezett *teljes potenciálegyenlet* :

$$\left(1 - \left(\frac{u_{x_1}}{c}\right)^2\right) u_{x_1 x_1} - 2 \frac{u_{x_1} u_{x_2}}{c^2} u_{x_1 x_2} + \left(1 - \left(\frac{u_{x_2}}{c}\right)^2\right) u_{x_2 x_2} = 0,$$

ahol  $u$  a sebesség potenciálja (azaz  $\vec{v} := \text{grad } u$  a sebesség vektora) és  $c$  a lokális hangsebesség – amely  $|\text{grad } u| := ((u_{x_1})^2 + (u_{x_2})^2)^{1/2}$  függvénye. Itt a diszkrimináns

$$D := \left(\frac{u_{x_1} u_{x_2}}{c^2}\right)^2 - \left(1 - \left(\frac{u_{x_1}}{c}\right)^2\right) \left(1 - \left(\frac{u_{x_2}}{c}\right)^2\right) = M^2 - 1,$$

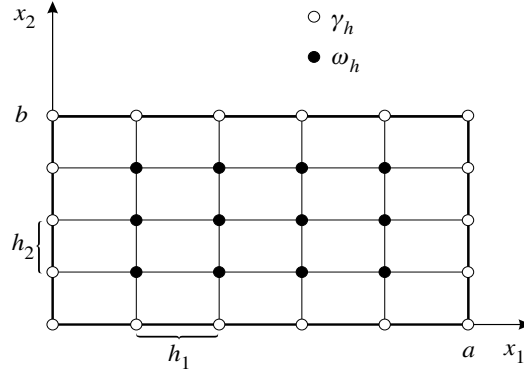
ahol  $M$  a Mach-szám :  $M := |\text{grad } u|/c$ .

Így az egyenlet elliptikus ott, ahol  $M < 1$  (szubszonikus áramlás) és hiperbolikus ott, ahol  $M > 1$  (szuperszonikus áramlás). A hangsebességhez közeli sebességgel járó repülőgép körüli áramlásban mindkét eset fordul elő.

### 15.3 Véges differencia eljárások II

A II. kötetben már foglalkoztunk differenciasémákkal, a differenciálegyenletekben előforduló deriváltakat differenciahányadosokkal felváltva. Erről az útról már ott derült ki, hogy az nem mindig megfelelő eredményekre vezet (ld. pl. a 11.4.8. pont elejét vagy 11.4.10-et), az alapvető Poisson-egyenlet esetén ez viszont a legalkalmasabb.





ábra 15.3: A kétdimenziós  $\omega_h$  rács

### 15.3.1 A Poisson–egyenlet approximációja differenciásémával

Elsőnek vizsgáljuk a következő kétdimenziós peremérték feladatot:

$$(\Delta u + f)(x) = 0, \quad x = (x_1, x_2) \in \Omega, \quad (15.13)$$

$$u(x) = g(x), \quad x \in \Gamma. \quad (15.14)$$

Itt  $\Omega := \{(x_1, x_2) \in (0, a) \times (0, b)\}$  téglalap és  $\Gamma$  annak pereme.

Bevezetünk egy ekvidisztáns rácsot („háló”-nak is hívhatjuk), ezzel a tartományt résztéglalapokra bontjuk fel:

$$\bar{\omega}_h := \{x^{ij} := (x_{1i}, x_{2j}) \in \bar{\Omega}, \quad x_{1i} = ih_1, \quad x_{2j} = jh_2, \\ i = 0, \dots, N_1, \quad j = 0, \dots, N_2\},$$

ahol  $h_1 := a/N_1$ ,  $h_2 := b/N_2$  a lépéstávolságok (ill. rácsállandók) és  $N_1, N_2 \geq 2$ . Az  $\omega_h$  rács indexe a  $h := (h_1, h_2)$  vektor.

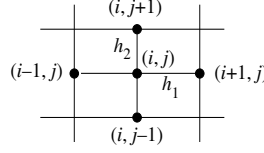
A rács  $\Omega$ -beli pontjait *belső pontoknak* hívjuk és  $\omega_h$ -val jelöljük :

$$\omega_h := \{x^{ij} = (x_{1i}, x_{2j}) \in \Omega, \quad x_{1i} = ih_1, \quad x_{2j} = jh_2, \\ i = 1, \dots, N_1 - 1, \quad j = 1, \dots, N_2 - 1\}, \quad (15.15)$$

míg a  $\Gamma$  peremen fekvő rácspontokat  $\gamma_h$ -val jelöljük,  $\gamma_h := \bar{\omega}_h \setminus \omega_h$ . Helyettesítve minden  $x^{ij}$  belső rácspontban a másodrendű deriváltakat a II. kötet 11.4.1. pontjában vizsgált másodrendű differenciahányadossal, a következő differencia-approximációt kapjuk (15.13)-(15.14)-ből :

$$(\Delta_h y)(x) + f(x) = 0, \quad x = x^{ij} \in \omega_h; \quad (15.16)$$

$$y(x) = g(x), \quad x = x^{ij} \in \gamma_h, \quad (15.17)$$



ábra 15.4: Az ötpontos differenciacsillag

ahol

$$\Delta_h y := y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2},$$

keresett  $x \in \bar{\omega}_h$ -ra az  $y$  diszkrét argumentumú avagy rácsfüggvény, és

$$\begin{aligned} y_{\bar{x}_1 x_1} &= y_{\bar{x}_1 x_1, ij} := (y_{i+1, j} - 2y_{ij} + y_{i-1, j})/h_1^2, \\ y_{\bar{x}_2 x_2} &= y_{\bar{x}_2 x_2, ij} := (y_{i, j+1} - 2y_{ij} + y_{i, j-1})/h_2^2. \end{aligned}$$

A (15.16)-(15.17) sémát *5-pontos differenciasémának* hívjuk ;  $\Delta u$  approximációjára az  $x^{ij}$  pontban az  $(x_{1, i\pm 1}, x_{2j})$ ,  $(x_{1i}, x_{2, j\pm 1})$ ,  $(x_{1i}, x_{2j})$  pontokat használja. Ezek alkotják az approximáció *differenciacsillagját* (avagy „stenciljét”). Emlékeztetünk a II. 11.4.1-ben bevezetett definícióra, amely szerint differenciaséma alatt a (15.16)-(15.17)-hez rendelhető egyenletrendszerek seregét értjük. Itt ennek a seregnek két paramétere van,  $N_1$  és  $N_2$  (vagy, ekvivalensen:  $h_1$  és  $h_2$ ), amelyekre szükség esetén a  $h = (h_1, h_2)$  vektorindexszel hivatkozunk.

A rögzített  $N_1$  és  $N_2$  mellett a (15.16)-(15.17) egyenletrendszerhez tartozó  $y_{ij}$  ismeretleneket sorfolytonosan  $\bar{y}_h$  vektorba rendezzük és az  $f_{ij}$ , ill  $g_{ij}$  értékeket ugyancsak sorfolytonosan  $\bar{\varphi}_h$  vektorba. Ezen vektorok dimenziója

$$\bar{n} := (N_1 + 1)(N_2 + 1),$$

az  $\bar{\omega}_h$  pontjai számának megfelelően. A sorfolytonosan átszámozott pontokat (és a hozzátartozó függvényértékeket)  $\ell$ -l fogjuk indexelni, az előző  $(i, j)$ -vel való indexelés helyett. Ekkor gyakran  $x^\ell := (x_{1\ell}, x_{2\ell})$  fog szerepelni  $x_{ij} = (x_{1i}, x_{2j})$  helyett.

A (15.16)-(15.17) differenciaséma mátrixalakja

$$\bar{A}_h \bar{y}_h = \bar{\varphi}_h, \tag{15.18}$$

ahol a  $\bar{\varphi}_h$  vektor komponenseit

$$\varphi_\ell = \begin{cases} f_\ell, & \text{ha } x^\ell \in \omega_h, \\ g_\ell, & \text{ha } x^\ell \in \gamma_h, \end{cases}$$

határozza meg, és ahol  $\bar{A}_h$  blokkmátrix. Blokkjainak dimenziója  $n_1 \times n_1$ , ahol  $n_k := N_k - 1$  a belső pontok száma  $x_k$  irányban ( $k = 1, 2$ ).

**Felhívjuk a figyelmet a következő jelölésbeli konvencióinkra:** amikor rácsfüggvényként tekintjük a diszkrét feladat megoldását vagy jobboldalát, akkor pl.  $y$  vagy  $y(x)$ ,  $x \in \bar{\omega}_h$  a jelölése és  $y_{ij}$  az  $x_{ij}$ -beli értéke, amikor viszont (oszlop-)vektornak értelmezzük (pl. az  $\bar{A}_h$  mátrix jobboldali szorzójaként), akkor  $\bar{y}_h$ , és ilyenkor a vektor komponense  $y_\ell$ . Ha tehát  $x^\ell := (x_{1\ell}, x_{2\ell})$  és  $x_{ij} = (x_{1i}, x_{2j})$  ugyanaz a pont a rácson, akkor  $y_\ell = y_{ij}$  és  $\varphi_\ell = \varphi_{ij}$ . Differencia kifejezésekben a rácsfüggvény szerinti értelmezés előnyesebb, pl.  $(y_{\bar{x}_1 x_1})_{ij} = (y_{i+1,j} - 2y_{ij} + y_{i-1,j})/h_1^2$  – és ekkor, ha nem okoz félreértést, inkább az indexmentes  $y_{\bar{x}_1 x_1}$  formát használjuk. A fenti  $\bar{A}_h \bar{y}_h$  vektor  $\ell$ -edik komponense pl.  $(\bar{A}_h \bar{y}_h)_\ell = -(\Delta_h y)(x^\ell)$ , ha  $x^\ell \in \omega_h$ , és az  $\ell$ -lel kapcsolatos sorfolytonos megszámozás nélkül kényelmes  $(\bar{A}_h \bar{y}_h)(x) = -(\Delta_h y)(x)$ , ha  $x \in \omega_h$ .

Az  $y$  rácsfüggvények tere az  $R^{\bar{n}}$ , ahogyan az  $\bar{y}_h$  vektoroké is, amely térben különböző normákat is fogunk használni.

**Végül, az  $u$  pontos megoldás  $x^\ell$  rácspontbeli  $u_\ell$  értékei vektorának jelölése  $\vec{u}_h$ .** Ez egy leképezés  $C(\bar{\Omega})$ -ből  $R^{\bar{n}}$ -be. Más folytonos függvények rácspontbeli értékeit és értékeinek vektorát is ennek megfelelően jelöljük.

Az  $\bar{A}_h$  mátrix felépítését abban az esetben mutatjuk meg, amikor  $h_1 = h_2 =: h$ . A mátrix főátlóján álló  $B \in \mathbb{R}^{n_1 \times n_1}$  blokkok képlete

$$B := \frac{1}{h^2} \begin{pmatrix} 4 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 4 \end{pmatrix} = \frac{1}{h^2} \text{tridiag}(-1, 4, -1). \quad (15.19)$$

Ezenkívül az  $\alpha := 1/h^2$  jelölést használjuk, a mátrix üres részei mind



### 15.3.2 A diszkrét maximumelv

Igazoljuk, hogy  $\bar{A}_h$  valóban M-mátrix. Ehhez tekintsük a következő  $w$  függvényt :

$$w(x) := 4 + x_1(a - x_1) + x_2(b - x_2).$$

Az  $\bar{\omega}_h$  rács  $(x_{1\ell}, x_{2\ell})$  tetszőleges pontjában (a sorfolytonos átszámozás szerint indexelve) vesszük  $w_\ell := w(x_{1\ell}, x_{2\ell})$ , amely számokból sorfolytonosan állítjuk össze a következő vektort :

$$\vec{w}_h := (w_1, \dots, w_{\bar{n}})^T, \quad \bar{n} = (N_1 + 1)(N_2 + 1).$$

A  $w$  függvényt – amelynek segítségével megmutatjuk az M-mátrix tulajdonságot – *majoráns függvénynek* hívjuk.

**15.2. Lemma** (az ötpontos differenciaséma tulajdonságai, 1).

a)  $\bar{A}_h$  M-mátrix;

b) Ha adott két (15.16)-(15.17) alakú feladat  $f^{(1)}, g^{(1)}$  ill.  $f^{(2)}, g^{(2)}$  adatokkal, és érvényes  $f^{(1)} \geq f^{(2)}$ ,  $g^{(1)} \geq g^{(2)}$ , akkor a hozzátartozó  $y^{(1)}, y^{(2)}$  megoldásokra igaz  $y^{(1)} \geq y^{(2)}$ .

**Megjegyzés.** Itt rácsfüggvényekkel fogalmaztuk meg a lemmát. Vektorok esetén a “ $\geq$ ” rendezés a vektoroknak az M-mátrixoknál szokásos komponensenkénti rendezése.  $\square$

**Bizonyítás.** a) Mivel a másodrendű differenciahányados a legfeljebb harmadfokú polinom pontos másodrendű deriváltját adja (ld. II. 11.4.1-ben a (30) képletet), ezért

$$(\bar{A}_h \vec{w}_h)_\ell = \begin{cases} -(\Delta w)_\ell = 4, & \text{amikor } x^\ell \in \omega_h, \\ w_\ell \geq 4, & \text{ha } x^\ell \in \gamma_h. \end{cases} \quad (15.20)$$

Ennek és az I. 1.3.4-beli definíciónak az alapján  $\bar{A}_h$  valóban M-mátrix. Így egyebek között  $\bar{A}_h$  reguláris, ami azt jelenti, hogy a (15.16)-(15.17) differenciaséma (és lejjebb a (15.24), (15.25) sémák) megoldása létezik és egyértelmű.

b) Az M-mátrix tulajdonságból következik, hogy  $(\bar{A}_h)^{-1} \geq 0$ . Legyenek most  $\bar{\varphi}_h^{(1)}$  és  $\bar{\varphi}_h^{(2)}$  a lemmában említett két feladat (15.18) mátrixalakjának jobboldali vektorai és  $\bar{y}_h^{(1)}$  ill.  $\bar{y}_h^{(2)}$  a hozzátartozó megoldások. Ekkor  $\bar{\varphi}_h^{(1)} \geq \bar{\varphi}_h^{(2)}$  ezen vektorok konstrukciója alapján, mindig komponensenként értelmezve az egyenlőtlenséget. Így

$$\bar{y}_h^{(1)} - \bar{y}_h^{(2)} = (\bar{A}_h)^{-1}(\bar{\varphi}_h^{(1)} - \bar{\varphi}_h^{(2)}) \geq 0. \quad \square$$

**Megjegyzések. 1.** A b) eredmény egyik gyakran alkalmazott formája az, hogy ( $|\bar{\varphi}_h|$ -vel jelölve a  $|\varphi_\ell|$  értékekből álló vektort)

$$\bar{A}_h \bar{y}_h = \bar{\varphi}_h, \quad \bar{A}_h \bar{v}_h = |\bar{\varphi}_h| \Rightarrow \bar{v}_h \geq 0 \text{ és } \bar{v}_h \geq |\bar{y}_h| \geq \pm \bar{y}_h.$$

**2.** A differenciaséma megőrzi a nemnegativitást : amennyiben a (15.13)-(15.14) feladat adatai nemnegatívak ( $f(x) \geq 0$ ,  $x \in \Omega$  és  $g(x) \geq 0$ ,  $x \in \Gamma$ ), akkor a pontos megoldás is nemnegatív és a diszkrét megoldás is az:  $y(x) \geq 0$ ,  $x \in \bar{\omega}_h$ , avagy  $\bar{y}_h \geq 0$ .  $\square$

Következőnek belátjuk, hogy a (15.16)-(15.17) differenciaséma stabil : megoldását becsüljük a feladat  $g$  és  $f$  adatainak a segítségével (v.ö. 10.2.2-vel, valamint 11.4.2-vel). Ehhez használjuk a maximum normát : ha  $v$  az  $S \subset \bar{\omega}_h$  halmazon adott rácsfüggvény, akkor a maximum normája :

$$\|v\|_{C(S)} := \max_{x^\ell \in S} |v(x^\ell)|.$$

**15.3. Tétel** (az ötpontos differenciaséma stabilitása). A (15.16)-(15.17) differenciaséma stabil a következő becslés értelmében :

$$\|\bar{y}_h\|_{C(\bar{\omega}_h)} \leq \left(1 + \frac{a^2 + b^2}{16}\right) \max(\|g\|_{C(\gamma_h)}, \|f\|_{C(\omega_h)}). \quad (15.21)$$

**Bizonyítás.** Az állítás azonnal következik az I. 1.8. lemmából, majoráns vektornak a fenti  $\vec{w}_h$  vektort használjuk. Eszerint

$$\begin{aligned} \|(\bar{A}_h)^{-1}\|_{C(\bar{\omega}_h)} &\leq \frac{\|\vec{w}_h\|_{C(\bar{\omega}_h)}}{\min_{x^\ell \in \bar{\omega}_h} (\bar{A}_h \vec{w}_h)_\ell} \\ &\leq \frac{1}{4} \|\vec{w}_h\|_{C(\bar{\omega}_h)} \leq \frac{1}{4} \|w\|_{C(\bar{\Omega})} = 1 + \frac{a^2 + b^2}{16}. \end{aligned} \quad (15.22)$$

Mivel tehát  $(\bar{A}_h)^{-1}$  (a megoldási operátor) egyenletesen korlátos  $h_1, h_2$ -ben, a differenciaséma stabil.

Az  $\bar{A}_h \bar{y}_h = \bar{\varphi}_h$  rendszer megoldása (15.22) alapján becsülhető :

$$\|\bar{y}_h\|_{C(\bar{\omega}_h)} \leq \|(\bar{A}_h)^{-1}\|_{C(\bar{\omega}_h)} \|\bar{\varphi}_h\|_{C(\bar{\omega}_h)} \leq \left(1 + \frac{a^2 + b^2}{16}\right) \|\bar{\varphi}_h\|_{C(\bar{\omega}_h)}.$$

Figyelembe véve  $\bar{\varphi}_h$  felépítését (ld. (15.18) alatt), megkapjuk a (15.21) becslést.  $\square$

Ezen eredmény birtokában már az ötpontos differenciaséma pontosságával tudunk foglalkozni.

A (15.13)-(15.14) feladat  $u$  pontos megoldásának értékeit sorfolytonosan az  $\vec{u}_h$  vektorba rendezve, a hibavektor a következő lesz :

$$\bar{z}_h := \bar{y}_h - \vec{u}_h.$$

Erre alkalmazzuk az  $\bar{A}_h$  mátrixot azzal a céllal, hogy egy  $\bar{z}_h$ -ra jellemző egyenletrendszert vezessünk le. Az  $\bar{A}_h \bar{y}_h$  kifejezés ismert (és egyenlő  $\bar{\varphi}_h$ -vel), de  $\bar{A}_h \bar{u}_h$  általában nem egyezik ezzel.

**Definíció.** A (15.16)-(15.17) differenciaséma  $\bar{\psi}_h$  képlethibájának nevezzük a

$$(\bar{\varphi}_h - \bar{A}_h \bar{u}_h)_\ell =: \psi_\ell := \begin{cases} (f + u_{\bar{x}_1 x_1} + u_{\bar{x}_2 x_2})_\ell, & \text{ha } x^\ell \in \omega_h, \\ 0, & \text{ha } x^\ell \in \gamma_h, \end{cases}$$

rácsfüggvényt. A  $\bar{\psi}_h$  vektort a  $\psi_\ell$  értékekből állítjuk össze sorfolytonos számozással.  $\square$

Amikor  $u \in C^4(\bar{\Omega})$ , akkor a képlethibára teljesül

$$\|\bar{\psi}_h\|_{C(\bar{\omega}_h)} \leq \frac{M_4}{12} (h_1^2 + h_2^2), \quad (15.23)$$

ahol  $M_4$  a negyedik deriváltak egy felső korlátja  $\bar{\Omega}$ -ban. Ehhez hivatkozunk a másodrendű differenciahányadossal kapcsolatos sorfejtésekre, amelyekkel már korábban 11.4-ben foglalkoztunk, ld. ott a (29)-(30) képleteket.

Ezekből világos, hogy  $u \in C^4(\bar{\Omega})$  esetén érvényes

$$\begin{aligned} (\Delta_h u)_{ij} = (\Delta u)_{ij} &+ \frac{h_1^2}{12} \frac{\partial^4 u}{\partial x_1^4} (x_{1i} + \vartheta_1 h_1, x_{2j}) \\ &+ \frac{h_2^2}{12} \frac{\partial^4 u}{\partial x_2^4} (x_{1i}, x_{2j} + \vartheta_2 h_2), \quad |\vartheta_{1,2}| < 1. \end{aligned}$$

Így a (15.23) becslés következik, amelynek alapján mondhatjuk, hogy a (15.16)-(15.17) séma *képlethibája másodrendű*, (avagy *approximációja másodrendű*), ha  $u$ -nak, a (15.13)-(15.14) peremérték feladat pontos megoldásának minden negyedik deriváltja folytonos  $\bar{\Omega}$ -n.

**15.4. Tétel** (az ötpontos differenciaséma konvergenciája). Legyen  $u \in C^4(\bar{\Omega})$  a (15.13)-(15.14) peremérték feladat pontos megoldása és  $\bar{y}_h$  a (15.16)-(15.17) differencia\_séma megoldásvektora.

Ekkor a (15.16)-(15.17) differenciaséma másodrendben konvergens :

$$\max_{x^\ell \in \bar{\omega}_h} |y_\ell - u(x^\ell)| = \|\bar{y}_h - \bar{u}_h\|_{C(\bar{\omega}_h)} \leq \left(1 + \frac{a^2 + b^2}{16}\right) \frac{M_4}{12} (h_1^2 + h_2^2).$$

**Bizonyítás.** A séma (15.18) mátrixalakjából indulunk ki. A  $\bar{z}_h$  hibavektor a következő egyenletrendszernek a megoldása:

$$\bar{A}_h \bar{z}_h = \bar{A}_h \bar{y}_h - \bar{A}_h \bar{u}_h = \bar{\varphi}_h - \bar{A}_h \bar{u}_h = \bar{\psi}_h.$$

Ezen rendszer megoldásának a becsléséhez nem a fenti  $w$  majoráns függvényt vesszük, hanem  $\frac{1}{4}\|\bar{\psi}_h\|_{C(\bar{\omega}_h)}w$ -t. Ezzel képezzük az  $\frac{1}{4}\|\bar{\psi}_h\|_{C(\bar{\omega}_h)}\vec{w}_h \in \mathbb{R}^{\bar{n}}$  vektort. Ekkor

$$\bar{A}_h \left( \frac{1}{4}\|\bar{\psi}_h\|_{C(\bar{\omega}_h)}\vec{w}_h \pm \bar{z}_h \right) \geq \|\bar{\psi}_h\|_{C(\bar{\omega}_h)}\bar{e}_h \pm \bar{A}_h\bar{z}_h \geq 0,$$

ahol  $\bar{e}_h := (1, \dots, 1)^T \in \mathbb{R}^{\bar{n}}$ . Ezért (15.21)-ből valamint (15.23)-ból kapjuk a

$$\|\bar{z}_h\|_{C(\bar{\omega}_h)} \leq \frac{1}{4}\|\bar{\psi}_h\|_{C(\bar{\omega}_h)}\|\vec{w}_h\|_{C(\bar{\omega}_h)} \leq \left(1 + \frac{a^2 + b^2}{16}\right) \frac{M_4}{12}(h_1^2 + h_2^2)$$

hibabecslést.  $\square$

Hasonlóan mint az egydimenziós esetben, jobb becslést úgy tudunk levezetni, ha a peremértékeket és a jobboldalt külön kezeljük, használva a feladat linearitását:

$$y = y_\gamma + y_\omega,$$

$$\begin{aligned} (y_\gamma)_{\bar{x}_1 x_1} + (y_\gamma)_{\bar{x}_2 x_2} &= 0, & x \in \omega_h, \\ y_\gamma(x) &= g(x), & x \in \gamma_h; \end{aligned} \quad (15.24)$$

$$\begin{aligned} (y_\omega)_{\bar{x}_1 x_1} + (y_\omega)_{\bar{x}_2 x_2} + f(x) &= 0, & x \in \omega_h, \\ y_\omega(x) &= 0, & x \in \gamma_h. \end{aligned} \quad (15.25)$$

A (15.24) homogén egyenlet egydimenziós megfelelőjének megoldását II. 11.4.3-ban explicite megadhattuk: ez megegyezett az eredeti homogén differenciálegyenlet megoldásával a rácspontokban, és így kaptuk az

$$\|y_\gamma\|_{C(\bar{\omega}_h)} \leq \|g\|_{C(\gamma_h)} \quad (15.26)$$

alakú becslést. Ez a becslés most is igaz (bár  $y_\gamma$  lényegesen eltérhet egy lineáris függvénytől), mivel a (15.24) differenciasémára – hasonlóan, mint a homogén (15.13)-(15.14) peremérték feladatra is – érvényes a (15.10) Hopf-féle maximumelv.

A *diszkrét maximumelv* fogalmának (egyik) definíciójához bevezetjük a következő halmazokat. Legyen  $A \in \mathbb{R}^{n \times n}$  adott mátrix és  $y \in \mathbb{R}^n$  rögzített vektor, ekkor

$$\begin{aligned} N^0 &= N^0(Ay) := \{ i; 1 \leq i \leq n, (Ay)_i = 0 \}, \\ N^\neq &= N^\neq(Ay) := \{ i; 1 \leq i \leq n, (Ay)_i \neq 0 \} = \{1, \dots, n\}/N^0. \end{aligned}$$



**Definíció.** Azt mondjuk, hogy  $A$  teljesíti a *(diszkrét) maximumelvet*, ha tetszőleges  $y \in \mathbb{R}^n$ -re igaz a következő egyenlőtlenség :

$$\max_{i \in N^0(Ay)} |y_i| \leq \max_{j \in N^\neq(Ay)} |y_j|. \quad (15.27)$$

Itt a  $\max_{k \in N} |y_k|$  értéke legyen nulla, ha az  $N$  halmaz üres.  $\square$

Eszerint a definíció szerint a maximumelv teljesüléséhez szükséges  $A$  regularitása. Ha ugyanis létezne olyan  $y \neq 0$ , hogy  $Ay = 0$ , akkor innen következne a  $0 < \max_{1 \leq i \leq n} |y_i| \leq 0$  ellentmondás.

**15.5. Tétel** (maximumelvvvel rendelkező mátrixosztály). Legyen  $A$  reguláris, (soronként) domináns főátlójú mátrix. Akkor teljesíti a maximumelvet.

**Bizonyítás.** Legyen  $y \neq 0$  olyan vektor, hogy  $N^0(Ay) \neq \emptyset$ . Az  $A$  regularitása miatt  $N^\neq(Ay) \neq \emptyset$ . Adott  $\varepsilon \geq 0$ -val tekintsük az  $A_\varepsilon := A + \varepsilon \text{diag}(\text{sign } a_{ii})$  mátrixot. Ha  $\varepsilon = 0$ , akkor  $A_\varepsilon = A$  reguláris. Ha  $\varepsilon > 0$ , akkor  $A_\varepsilon$  azért reguláris, mert főátlója szigorúan domináns :

$$|a_{ii}| + \varepsilon > \sum_{j \neq i} |a_{ij}| \quad \text{minden } i\text{-re.}$$

(Itt és lejjebb  $\sum_{\substack{j=1 \\ j \neq i}}^n$  helyett azt írjuk, hogy  $\sum_{j \neq i}$ .)

Legyen most  $\varepsilon > 0$  rögzített és  $z \neq 0$  tetszőleges olyan vektor, hogy  $N^0(A_\varepsilon z) \neq \emptyset$ . Legyen továbbá  $k$  tetszőleges olyan index, amellyel  $|z_k| = \max_{1 \leq i \leq n} |z_i| > 0$ . Ekkor, ha  $k \in N^0(A_\varepsilon z)$  lenne, akkor a

$$\begin{aligned} 0 = \text{sign}(a_{kk}z_k)(A_\varepsilon z)_k &= (|a_{kk}| + \varepsilon)|z_k| + \text{sign}(a_{kk}z_k) \sum_{j \neq k} a_{kj}z_j \\ &\geq \left( |a_{kk}| + \varepsilon - \sum_{j \neq k} |a_{kj}| \right) |z_k| > 0 \end{aligned}$$

ellentmondás keletkezne. Így tehát  $k \in N^\neq(A_\varepsilon z)$  és ( $z_k$  definíciója alapján)

$$|z_k| = \max_{j \in N^\neq(A_\varepsilon z)} |z_j| > \max_{i \in N^0(A_\varepsilon z)} |z_i|.$$

Ha most speciálisan  $z = z_\varepsilon$ -t abból választjuk meg, hogy  $A_\varepsilon z = Ay$  legyen, a fenti  $y$  vektorral, akkor  $N^\neq(A_\varepsilon z_\varepsilon) = N^\neq(Ay)$  és  $N^0(A_\varepsilon z_\varepsilon) = N^0(Ay)$ , és  $\varepsilon \rightarrow 0$  határátmenettel kapjuk meg a tétel állítását.  $\square$

Ehhez a tételhez ld. az 1. feladatot is.

Hasonlóan a fentihez *szigorú maximumelv* alatt érthetjük a következő mátrix-tulajdonságot : tetszőleges  $y \in \mathbb{R}^n$ -re teljesül

a)  $N^\neq(Ay) = \emptyset \Rightarrow y = 0$  ( $A$  reguláris);

b) ha  $N^\neq(Ay) \neq \emptyset$  és  $N^0(Ay) \neq \emptyset$ , akkor legyen igaz (15.27) szigorú egyenlőtlenséggel.

**15.6. Tétel** (szigorú maximumelvvvel rendelkező mátrixosztály).  $A \in \mathbb{R}^{n \times n}$  pontosan akkor teljesíti a szigorú maximumelvet, ha főátlója szigorúan domináns.

**Bizonyítás.** A fenti 15.5. tétel megmutatja, hogy a feltétel elégséges, hiszen ekkor még az  $\varepsilon$ -t sem kell bevezetni a bizonyításhoz. A szükségesség belátásához induljunk ki abból, hogy a) szerint a mátrix reguláris. Ezután tetszőleges rögzített  $i$ -re vizsgáljuk az  $y = e_i$  koordináta egységvektor esetét. Ez azonnal megmutatja, hogy  $i \in N^\neq(Ay)$ , azaz  $a_{ii} \neq 0$ , hiszen nem lehet  $N^\neq(Ay) = \emptyset$ , és  $\max_{1 \leq j \leq n} |y_j| = |y_i| = 1$  ismert.

Most válasszuk az  $y$  vektort a rögzített  $i$  függvényében úgy, hogy

$$y_j := \overline{\text{sign}}(a_{ij}), \text{ ha } j \neq i, \quad y_i := - \sum_{j \neq i} |a_{ij}|/a_{ii}.$$

Itt használtuk a

$$\overline{\text{sign}}(t) := \begin{cases} 1, & \text{ha } t \geq 0, \\ -1, & \text{ha } t < 0 \end{cases}$$

jelölést. Ekkor  $y \neq 0$ ,  $N^\neq(Ay) \neq \emptyset$ , és  $(Ay)_i = 0$ , tehát

$$|y_i| \leq \max_{k \in N^0} |y_k| < \max_{j \in N^\neq} |y_j| = 1,$$

vagyis  $\sum_{j \neq i} |a_{ij}| < |a_{ii}|$ .  $\square$

**Megjegyzés.** Ha  $A$  M-mátrix és  $Ae \geq 0$  (ill.  $Ae > 0$ ), akkor a 15.5. tétel (ill. a 15.6. tétel) feltételei teljesülnek és így igaz a (szigorú) maximumelv.  $\square$

Ehhez a témához ld. Borisov cikkét is.

Mielőtt a 15.5. és 15.6. tételekből következtéseket vonnánk le a (15.16)-(15.17), ill. (15.24), (15.25) differenciasémák megoldásaira nézve, még szükségünk van egy lemmára, amely az  $Ax = b$  egyenletrendszer megoldásának (az I-beli 1.8. lemmához képest) javított becslését teszi lehetővé, amikor  $A$  *monoton mátrix* (azaz reguláris, és inverze elemenként nemnegatív, mint pl. M-mátrix esetén), és a jobboldali vektornak nulla komponensei is vannak. Ilyen indexekre a  $g$  majoráns vektorral képzett  $Ag$  vektor is nulla lehet.

**15.7. Lemma** (monoton mátrixú egyenletrendszer megoldásának becslése). Legyen  $A \in \mathbb{R}^{n \times n}$  monoton mátrix és  $0 \neq b \in \mathbb{R}^n$  egy nulla komponensekkel is rendelkező vektor.

Továbbá, legyen  $g \geq 0$  majoráns vektor abban az értelemben, hogy  $Ag \geq 0$ , és ha valamelyik  $i$  indexre  $(Ag)_i = 0$ , akkor igaz  $b_i = 0$ . Ekkor az  $Ay = b$  megoldására teljesül a következő becslés :

$$|y_i| \leq g_i \frac{\max_{j \in N \neq(b)} |b_j|}{\min_{j \in N \neq(b)} (Ag)_j} \quad \text{minden } i\text{-re.}$$

**Bizonyítás.** Ez az idézett 1.8. lemma bizonyítása azzal a különbséggel, hogy az  $(A(mg \pm y))_i = m(Ag)_i \pm b_i \geq 0$  relációt az  $i \in N^0(b)$  indexekre nem kell vizsgálni.  $\square$

**15.8. Következmények** (az ötpontos differenciaséma tulajdonságai, 2).

1) A (15.16)-(15.17) feladat  $\bar{A}_h$  mátrixa (ld. (15.18)-at) teljesíti a maximumelvet;

2) A (15.24) feladat megoldása eleget tesz a (15.26) becslésnek.

3) A (15.25) feladat megoldása teljesíti a következő becslést :

$$\|y_\omega\|_{C(\bar{\omega}_h)} \leq \frac{a^2 + b^2}{16} \|f\|_{C(\omega_h)}; \quad (15.28)$$

4) A (15.16)-(15.17) differenciaséma megoldására érvényes a következő becslés :

$$\|y\|_{C(\bar{\omega}_h)} \leq \|g\|_{C(\gamma_h)} + \frac{a^2 + b^2}{16} \|f\|_{C(\omega_h)}. \quad (15.29)$$

**Bizonyítás.** 1) Az  $\bar{A}_h$  mátrixról már bebizonyítottuk, hogy M-mátrix. Ezenkívül

$$(\bar{A}_h \bar{e}_h)_\ell = \begin{cases} 1, & \text{ha } x^\ell \in \gamma_h \\ 0, & \text{ha } x^\ell \in \omega_h \end{cases} \geq 0,$$

ahol  $\bar{e}_h := (1, \dots, 1)^T \in \mathbb{R}^{\bar{n}}$ . Tehát  $\bar{A}_h$  főátlója domináns. Ezután alkalmazhatjuk a 15.5. tételt.

2) A (15.24) feladat egyenletrendszere,

$$\bar{A}_h(y_\gamma)_h = \bar{g}_h,$$

olyan  $\bar{g}_h$  jobboldallal rendelkezik, amelynek komponensei

$$g_\ell := \begin{cases} 0, & \text{ha } x^\ell \in \omega_h, \\ g(x^\ell), & \text{ha } x^\ell \in \gamma_h, \end{cases}$$

tehát  $N^0(\bar{g}_h) \supseteq \omega_h$ , és  $\gamma_h \supseteq N^\neq(\bar{g}_h)$ . Ezért következik

$$\max_{x^\ell \in \omega_h} |y_\gamma(x^\ell)| \leq \max_{\ell \in N^0} |(y_\gamma)_\ell| \leq \max_{\ell \in N^\neq} |(y_\gamma)_\ell| \leq \max_{x^\ell \in \gamma_h} |y_\gamma(x^\ell)|,$$

vagyis

$$\|y_\gamma\|_{C(\omega_h)} \leq \|y_\gamma\|_{C(\gamma_h)} = \|g\|_{C(\gamma_h)}. \quad (15.30)$$

Legkesőbb itt látszik a diszkrét maximumelv fenti definíciójának célszerűsége.

3) Ennek belátásához vesszük a

$$w(x) := x_1(a - x_1) + x_2(b - x_2) \quad (15.31)$$

majoráns függvényt. Értékeiből a  $\vec{w}_h(w(x^1), \dots, w(x^{\bar{n}}))^T$  vektort állítjuk össze. Ez nemnegatív, és

$$(\bar{A}_h \vec{w}_h)_\ell = \begin{cases} 4, & \text{ha } x^\ell \in \omega_h \\ w_\ell, & \text{ha } x^\ell \in \gamma_h \end{cases} \geq 0.$$

Tehát  $\vec{w}_h$ -mal együtt  $\bar{A}_h \vec{w}_h$  a téglalap sarokpontjaiban eltűnik, de ott (15.25) jobboldala is nulla, így a 15.7. lemma alkalmazható. Ez adja a (15.28) becslést, hiszen  $\min_{x^\ell \in \omega_h} (\bar{A}_h \vec{w}_h)_\ell = 4$  és  $\max_{x^\ell \in \omega_h} w_\ell = \frac{a^2+b^2}{4}$ .

4) Végző (15.29) becslésünk következik a háromszög egyenlőtlenség segítségével.  $\square$

A (15.25) feladat megoldásának (15.28) becslését másképpen is készíthetjük. Mivel most annak peremértékei homogének, kézenfekvő az, hogy tőlük teljesen megszabaduljunk. Ezen *homogén peremértékek eliminációja* céljából először a (15.25)-höz tartozó  $\bar{A}_h(\bar{y}_\omega)_h = \bar{b}_h$  alakú egyenletrendszer  $\ell \in \gamma_h$  számú sorait töröljük, majd egyrészt a visszamaradó egyenletrendszer mátrixából ugyanezeket a  $\gamma_h$ -val kapcsolatos oszlopokat, másrészt az  $(\bar{y}_\omega)_h$  vektorból a  $\gamma_h$  diszkrét peremre eső komponenseket is elhagyjuk. Az így kapott vektort  $y_h$ -val fogjuk jelölni, a pontos megoldás  $\omega_h$ -ra szűkített vektorát továbbra is  $\vec{u}_h$ -val jelöljük, és az  $y_h$ -hoz tartozó rácsfüggvény továbbra is  $y$ .

A homogén peremértékek eliminációja után az egyenletrendszer dimenziója

$$n := (N_1 - 1)(N_2 - 1).$$

Ez  $\omega_h$  elemeinek a száma, és  $\mathbb{R}^n$  a rácsfüggvények tere. Az  $f_\ell$ -ek lesznek a jobboldal komponensei az  $x^\ell \in \omega_h$  pontjaiban, amelyeknek (továbbra is  $\ell$ -lel indexelt) új sorfolytonos átszámozását 1-től  $n$ -ig elkészítjük. Az eredő vektort  $\vec{f}_h$ -val jelöljük, tehát  $\vec{f}_h := (f(x_{11}, x_{21}), \dots, f(x_{1n_1}, x_{2n_2}))^T$  és  $n_k := N_k -$

1,  $k = 1, 2$ . A hozzátartozó,  $A_h$ -val jelölt mátrix pedig blokk-tridiagonális és szimmetrikus. Négyzetes rács esetén alakja

$$A_h = \begin{pmatrix} B & -\alpha I & & & \\ -\alpha I & B & -\alpha I & & \\ & -\alpha I & \ddots & \ddots & \\ & & \ddots & \ddots & -\alpha I \\ & & & -\alpha I & B \end{pmatrix}, \quad (15.32)$$

ahol  $\alpha := 1/h^2$ , a blokkok  $N_1 - 1$ -dimenziósak és  $B$ -t (15.19) adja meg. A 15.3.1. pont végén ábrázolt speciális  $\bar{A}_h$  mátrix a peremfeltételek eliminációja után pl. a következő  $A_h$  alakot veszi fel :

$$A_h = \begin{pmatrix} \frac{2}{h_1^2} + \frac{2}{h_2^2} & -\frac{1}{h_1^2} \\ -\frac{1}{h_1^2} & \frac{2}{h_1^2} + \frac{2}{h_2^2} \end{pmatrix}.$$

Az egyenletrendszer ezután

$$A_h y_h = \vec{f}_h. \quad (15.33)$$

Amennyiben az  $\bar{w}_h$  pontjait úgy sorszámozzuk, hogy a perempontok után soroljuk a belső pontokat, akkor az  $\bar{A}_h$  és  $A_h$  mátrixok közti összefüggést a következő blokkalak szemlélteti :

$$\bar{A}_h = \begin{pmatrix} I_\gamma & 0 \\ A_\gamma & A_h \end{pmatrix} \in \mathbb{R}^{\bar{n} \times \bar{n}}, \quad \text{ahol } A_h \in \mathbb{R}^{n \times n}. \quad (15.34)$$

Itt  $I_\gamma$  az egységmátrix a perempontok  $\bar{n} - n$  számának megfelelő dimenzióban, az  $A_\gamma$  blokk tartalmazza a belső pontok kapcsolatait a peremmel.

Az  $A_h$  mátrix M-mátrix tulajdonsága, valamint a 15.8. következmény 3. részének megfelelő

$$\|A_h^{-1}\|_{C(\omega_h)} \leq \frac{a^2 + b^2}{16} \quad (15.35)$$

becslése beláthatók a (15.31) majoráns függvény segítségével (egy másik bizonyítás a 1\*. feladat tárgya). Ez ugyanis pozitív  $\omega_h$ -n, továbbá az  $\vec{w}_h = (w(x_{11}, x_{21}), \dots, w(x_{1n_1}, x_{2n_2}))^T$  ( $n$ -dimenziós) vektorra teljesül

$$(A_h \vec{w}_h)_i \geq 4, \quad 1 \leq i \leq n,$$

valamint

$$\|\vec{w}_h\|_{C(\omega_h)} \leq \|w\|_{C(\bar{\Omega})} = \frac{a^2 + b^2}{4}.$$

Az M-mátrix tulajdonságon alapuló technikával vizsgálható a Poisson-egyenlet (15.13)-(15.14) peremérték feladatának negyedrendű approximációját adó alábbi 9-pontos differenciaséma is :

$$x = x^{ij} \in \omega_h : \\ y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} + \frac{h_1^2 + h_2^2}{12} y_{\bar{x}_1 x_1 \bar{x}_2 x_2} + \varphi(x) = 0, \quad (15.36)$$

$$\varphi(x) = f(x) + \frac{h_1^2}{12} f_{\bar{x}_1 x_1} + \frac{h_2^2}{12} f_{\bar{x}_2 x_2}, \quad (15.37)$$

$$x = x^{ij} \in \gamma_h : \\ y(x) = g(x). \quad (15.38)$$

Ez az approximáció közvetlenül a II. 11.4.1-beli (15.33) sorfejtés, valamint a (15.13)-(15.14) egyenlet segítségével levezethető. A differenciaséma pontossági becslése miatt ld. a 2. feladatot.

A következő táblázat megmutatja a negyedrendű (15.36)-(15.38) eljárás előnyét a (15.16)-(15.17) másodrendűhez képest – ha  $u$  megfelelő simasággal rendelkezik. Az itt közelítőleg megoldott (15.13)-(15.14) feladat adatai :  $a = b = 1$ , és  $u(x) = x_1^6 + x_2^6$  a pontos megoldás, ennek megfelelően  $f$  és  $g$ . A táblázat tartalmazza a  $\delta(h) := \|\bar{u}_h - y_h\|_{C(\omega_h)}$  értékeit különböző  $h$ -kra, a számítás dupla pontosságban készült.

séma / h	1/2	1/4	1/8	1/16	1/32
(15.16)-(15.17)	0.25	$7.434 \cdot 10^{-2}$	$2.167 \cdot 10^{-2}$	$5.489 \cdot 10^{-3}$	$1.377 \cdot 10^{-3}$
(15.36)	0.0281	$1.728 \cdot 10^{-3}$	$1.079 \cdot 10^{-4}$	$6.744 \cdot 10^{-6}$	$4.216 \cdot 10^{-7}$

A fenti 15.5. és 15.6. tételekkel, valamint az azokból következő (15.30), ill. (15.28) becslésekkel még nem jellemeztük a (15.24), ill. (15.25) feladatok megoldásainak egy fontos tulajdonságát, a nemnegativitás megőrzését (amely a 15.2. lemmából következik, ha ott  $f^{(2)} = 0$ ,  $g^{(2)} = 0$ ). A diszkrét maximumelv fenti (15.27) definíciója csak az abszolútértékekre vonatkozik, ezenkívül ez az elv nem analóg a (15.10) maximumelvvvel. Ezért most megmutatjuk (az  $\bar{A}_h$  mátrix M-tulajdonsága és irreducibilitása alapján), hogy az utóbbi elv diszkrét hasonmása is igaz, sőt ennek a szigorúbb megfogalmazása is.

**15.9. Tétel** (szigorú diszkrét Hopf-féle maximumelv). Legyen  $\bar{A}_h$  a (15.18) mátrix. Ekkor

a) ha  $(\bar{A}_h \bar{y}_h)_\ell \geq 0$  minden  $x^\ell \in \omega_h$ -ra, akkor  $\bar{y}_h$  vagy konstans  $\bar{\omega}_h$ -n, vagy csak  $\gamma_h$ -n veheti fel a minimumát;

b) ha  $(\bar{A}_h \bar{y}_h)_\ell \leq 0$  minden  $x^\ell \in \omega_h$ -ra, akkor  $\bar{y}_h$  vagy konstans vagy csak  $\gamma_h$ -n veheti fel maximumát.

**Bizonyítás.** a) Tegyük fel, hogy  $\bar{y}_h$  nem konstans, és egy  $x^\ell = (x_{1\ell}, x_{2\ell}) \in \omega_h$ -beli minimumhelyéhez tartozó  $y_\ell$  értéknél szomszédpontjaiban,

$$\begin{aligned} x^{(\ell,1)} &:= (x_{1\ell} + h_1, x_{2\ell}), & x^{(\ell,3)} &:= (x_{1\ell} - h_1, x_{2\ell}), \\ x^{(\ell,2)} &:= (x_{1\ell}, x_{2\ell} + h_2), & x^{(\ell,4)} &:= (x_{1\ell}, x_{2\ell} - h_2), \end{aligned}$$

legalább egy nagyobb érték is van :  $y(x^{(\ell,j)}) > y_\ell$  valamelyik  $j \in \{1, 2, 3, 4\}$ -re. Ekkor

$$(\bar{A}_h \bar{y}_h)_\ell = -\frac{y(x^{(\ell,1)}) - 2y_\ell + y(x^{(\ell,3)})}{h_1^2} - \frac{y(x^{(\ell,2)}) - 2y_\ell + y(x^{(\ell,4)})}{h_2^2} < 0.$$

Ez ellentmondást jelent, így a feltevés nem lehet igaz és csak az a két lehetőség marad, hogy vagy  $\bar{y}_h$  konstans, vagy a minimumhely csak  $\gamma_h$ -n van.

Analóg módon igazoljuk a b) tulajdonságot is. Az a)-ból és b)-ből következik a (15.30)-nál pontosabb becslés

$$\min_{x^i \in \gamma_h} g(x^i) \leq y_\ell \leq \max_{x^i \in \gamma_h} g(x^i), \quad \text{minden } x^\ell \in \omega_h\text{-ra.} \quad \square$$

**15.10. Következmény** (pozitív elemű inverz mátrix). Az  $A_h^{-1}$  mátrix minden eleme pozitív.

**Bizonyítás.** Induljunk ki a (15.34) blokkalakból. Legyen  $\bar{y}_h^{(j)}$  az  $\bar{A}_h \bar{y}_h = \bar{e}_h^{(j)}$  rendszer megoldása, ahol  $\bar{e}_h^{(j)}$  a  $j$ -edik  $\bar{n}$ -dimenziós egységvektor. Ekkor  $\bar{y}_h^{(j)}$  az  $(\bar{A}_h)^{-1} = \begin{pmatrix} I_\gamma & 0 \\ -A_h^{-1} A_\gamma & A_h^{-1} \end{pmatrix}$  mátrix  $j$ -edik oszlopa és

$$x^j \in \gamma_h \text{ (azaz } j \leq \bar{n} - n) : \bar{y}_h^{(j)} = \begin{pmatrix} e_\gamma^{(j)} \\ y_h^{(j)} \end{pmatrix}, \quad x^j \in \omega_h : \bar{y}_h^{(j)} = \begin{pmatrix} 0 \\ y_h^{(j)} \end{pmatrix},$$

ahol  $e_\gamma^{(j)}$  a  $j$ -edik  $(\bar{n} - n)$ -dimenziós egységvektor. A 15.9. tétel most azt adja, hogy  $x^j \in \gamma_h$  esetén  $\bar{y}_h^{(j)}$  minimális (és maximális) komponense csak  $\gamma_h$ -n van, mert  $(\bar{A}_h \bar{y}_h^{(j)})_\ell = 0$ ,  $x^\ell \in \omega_h$ . Azaz  $y_h^{(j)} > 0$  (és minden komponense kisebb egynél). Amikor viszont  $x^j \in \omega_h$ , akkor  $\bar{y}_h^{(j)}$  minimális komponense csak  $\gamma_h$ -n van, tehát  $y_h^{(j)} > 0$ .  $\square$

Ezen és az előző pont konstrukcióit, vizsgálati módszereit nehézség nélkül alkalmazhatjuk a téglalapokból összeállított  $\Omega$  megoldási tartomány esetére (ld. a 3. feladatot), a *háromdimenziós Poisson-egyenlet* téglalakú tartományban megfogalmazott elsőfajú peremérték feladatára is (ld. a 4. feladatot), ill. más peremfeltételekre is (ld. 15.3.4.). Az általánosabb tartomány kérdése viszont több problémát vet fel (ld. 15.3.5.). Mielőtt ezzel foglalkoznánk, a következő pontban megmutatjuk, hogy a (15.33) egyenletrendszer megoldása már közepes rácsméret esetén sem mellékes számítási feladat.

### 15.3.3 A diszkrét Poisson–egyenlet megoldása; a Fourier–módszer

A (15.33) egyenletrendszer numerikus megoldása során a lényeges körülmény az, hogy az  $A_h$  tipikus ritkamátrix. Ha pl.  $N_1 = N_2 = 41$ , akkor  $n = 1600$ , és az összesen 2.56 millió mátrixelemből csak 7852 nemzérus. Általában az összesen  $n^2$  mátrixelemből kevesebb mint  $5n$  eleme nemzérus.

Mivel  $A_h$  szimmetrikus M-mátrix (tehát pozitív definit, ld. I. 3.2. tétel), a következő módszerek alkalmazhatók (v.ö. I. 1.3., 1.6.): Cholesky–módszer, Jacobi- és Gauss–Seidel iteráció, egyszerű iteráció, konjugált gradiens módszer, inkomplett Gauss–elimináció kombinált egyszerű iterációval, felső relaxáció (SOR) és szimmetrizált változata (SSOR), végül a Csebisev–iteráció.

A következőkben (feltéve, hogy  $N_1$  és  $N_2$  összemérhetők : vannak olyan pozitív  $c_1$  és  $c_2$  konstansok, hogy  $c_1 N_1 \leq N_2 \leq c_2 N_1$ ). Ekkor  $n = (N_1 - 1)(N_2 - 1) = O(N_1^2) = O(N_2^2)$  megadjuk a szükséges tár- és műveletigényt. Ehhez az iterációs módszereknél a pontosságot  $\varepsilon$ -nak választjuk, azaz, ha érvényes (a (15.33) rendszer  $y_h$  pontos megoldásával és  $y_h^{(m)}$   $m$ -edik iterációs közelítésével) az

$$\|y_h^{(m)} - y_h\| \leq \varepsilon \|y_h^{(0)} - y_h\| \quad (15.39)$$

becslés, akkor abba hagyjuk az iterációt. Az  $\varepsilon$  konkrét kiválasztása attól függ, hogy milyen normában vizsgáljuk az iterációs eljárás konvergenciáját. Pl. az egyszerű iteráció esetén ez I. 1.6.6-ban volt az euklidesi norma. Ekkor (15.39)-et megszorozva  $h_1 h_2$ -vel kapjuk ugyanazt a becslést az  $L_2(\omega_h)$ -normában. Mivel a diszkrétizációnk hibája  $O(h_1^2 + h_2^2)$  a maximum normában, így észszerű azt megkövetelni, hogy legyen

$$\varepsilon = (h_1^2 + h_2^2) \sqrt{h_1 h_2} = O(n^{-3/2}).$$

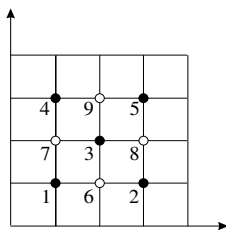
Ekkor ugyanis

$$\begin{aligned} \|\bar{y}_h^{(m)} - \bar{y}_h\|_{C(\bar{\omega}_h)} &= \|y_h^{(m)} - y_h\|_{C(\omega_h)} \leq \frac{1}{\sqrt{h_1 h_2}} \|y_h^{(m)} - y_h\|_{L_2(\omega_h)} \\ &\leq \frac{\varepsilon}{\sqrt{h_1 h_2}} \|y_h^{(0)} - y_h\|_{L_2(\omega_h)} = O(h_1^2 + h_2^2) \|y_h^{(0)} - y_h\|_{L_2(\omega_h)}, \end{aligned}$$

ami célszerű, tekintve, hogy a  $\bar{y}_h$  és  $\bar{u}_h$  távolságát is a maximum normában mértük.

Így a következő táblázatban  $\ln 1/\varepsilon = O(\ln n)$ . Még egyszer arra mutatunk rá, hogy mindig a diszkrét, kétdimenziós Poisson–egyenlet megoldásáról van szó, és  $n$  a rács belső pontjainak száma.





ábra 15.5: Sakktábla elrendeződés

eljárás	tárigény	műveletigény	megjegyzés
Cholesky módszer (telt mátrix)	$n^2/2 + O(n)$	$n^3/6 + O(n^2)$	
Cholesky (sávós mátrix)	$n^{3/2} + O(n)$	$n^2/2 + O(n)$	
szim. blokk-tridiag. Gauss-elimináció (ld. *** 18.10.3.)	$n^{3/2} + O(n)$	$\frac{n^2}{6} + 4n^{3/2} + O(n)$	
konjugált gradiens módszer	$5n$	$10n^2$	prekondicionálás nélkül
egyszerű iteráció	$2n$	$O(n^2 \log 1/\varepsilon)$	
SOR( $\omega_{opt}$ )	$2n$	$O(n^{3/2} \log 1/\varepsilon)$	$\omega_{opt}$ választásá- hoz ld. lent
2-réteg. Csebisev	$2n$	$O(n^{3/2} \log 1/\varepsilon)$	kb. $n = 225$ -től kezdve jobb mint SOR( $\omega_{opt}$ ); általánosítható
SSOR (Szamarszkij-féle iter. paraméterekkel)	$2n$	$O(n^{5/4} \log 1/\varepsilon)$	általánosítható
Fourier-módszer (FFT)	$n + (13/2 +$ $+1/2 \log_2 n)n^{1/2}$	$2.5n \log_2 n + n$	csak téglalpra általánosítható: $u_{xx} + a(y)u_{yy} +$ $+b(y)u_y + d(y)u = f$
MG	$8/3n$	$O(n)$	általán. ell. feladat

Itt az MG jelenti a 15.4-ben tárgyalásra kerülő többrácsos módszert (angolul „multigrid”) és FFT a gyors Fourier-transzformációt, ld. lejjebb. A háromdimenziós Poisson-egyenlet esetén az iterációs módszerek műveletigénye relatíve még jobb (ld. I. 1.6.1.), mint a sávós Cholesky-módszeré, amely akkor  $O(n^{7/3})$  műveletet követel. A konjugált gradiens módszer esetén abból indultunk ki, hogy az  $n$  lépés után a pontos megoldást adja, és hogy az  $A_h$  mátrixot nem kell tárolnunk. Ezen módszer igen fejlett prekondicionálására visszatérünk 15.8.5-ben.

Az SOR alkalmazásához először mégegyszer rámutatunk, hogy  $A_h = A_h^T > 0$  érvényes. Ezután lássuk be, hogy az ismeretlenek *sakktábla elrendeződése* után a mátrix blokk-tridiagonális lesz. Mégpedig világos (az ötpon-tos differenciacsillag miatt), hogy egy-egy fekete pont csupa „fehér” pontokkal van körülvéve – és ez biztosítja az  $A_h$  mátrix említett blokkalakját általános  $N_1, N_2$  esetén. Ha  $h_1 = h_2 = h$  és az  $A_h$  mátrixot még  $\frac{1}{4}h^2$ -tel szorozzuk, akkor főátlóján  $I$ -blokkok állnak – tehát az A-tulajdonság is megvan (v.ö. I.

1.6.4.), így  $\omega_{opt}$  kiszámítható. Legyen pl.  $N_1 = N_2 = 4$  :

$$\frac{h^2}{4}A_h = \begin{pmatrix} 1 & & & \cdot & \delta & \delta & & \\ & 1 & & \cdot & \delta & \delta & & \\ & & 1 & \cdot & \delta & \delta & \delta & \delta \\ & & & 1 & \cdot & \delta & \delta & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \delta & \delta & \delta & \cdot & 1 & & & \\ \delta & \delta & \delta & \cdot & & 1 & & \\ & \delta & \delta & \delta & \cdot & & 1 & \\ & & \delta & \delta & \delta & \cdot & & 1 \end{pmatrix} \begin{array}{l} \text{„fekete” ismeretlenek sorai} \\ \cdot \cdot \cdot \cdot \cdot \cdot \\ \text{„fehér” ismeretlenek sorai} \end{array}$$

ahol  $\delta := -\frac{1}{4}$ , az üres helyeken nullák állnak.

Foglalkozzunk most inkább a Fourier-módszerrel! Ez a fenti táblázatban szereplő módszerek közül a második legjobb, ráadásul ez iterációmentes és nemcsak (15.16)-(15.17) megoldásához hasznos. Az eljárás a *gyors Fourier-transzformáció*n alapszik (ld. I. 4.7.), valamint azon, hogy az

$$(\Delta_h v)(x) + \lambda^h v(x) = 0, \quad x \in \omega_h; \quad v(x) = 0, \quad x \in \gamma_h$$

diszkrét sajátérték feladat megoldása a sin függvény segítségével leírható. Ennek lineáris algebrai alakja

$$A_h y_h = \lambda^h y_h, \quad (15.40)$$

megoldásai a  $v^{(k\ell)}$  „diszkrét” saját-rácsfüggvények és a  $\lambda_{k\ell}^h$  sajátértékek :

$$v^{(k\ell)}(x) := \frac{2}{\sqrt{ab}} \sin \frac{k\pi x_1}{a} \sin \frac{\ell\pi x_2}{b}, \quad x \in \omega_h, \quad (15.41)$$

$$\lambda_{k\ell}^h := \frac{4}{h_1^2} \sin^2 \frac{k\pi h_1}{2a} + \frac{4}{h_2^2} \sin^2 \frac{\ell\pi h_2}{2b}, \quad (15.42)$$

$$1 \leq k \leq N_1 - 1, \quad 1 \leq \ell \leq N_2 - 1.$$

Az  $\omega_h$ -n definiált  $u, v$  rácsfüggvényekre vezessük be a

$$(u, v)_{(0,h)} := \sum_{x \in \omega_h} u(x)v(x)h_1 h_2, \quad (15.43)$$

skalárszorzatot! Ezzel a rácsfüggvények tere (az  $\mathbb{R}^n$ ) Hilbert-tér lesz, amelyet  $L_2(\omega_h)$ -val jelöljük. Az  $(u, v)_{(0,h)}$  skalárszorzatot  $(u, v)_{L_2(\omega_h)}$ -nak is írjuk.

Ebben a skalárszorzatban a  $\{v^{(k\ell)}\}$  sajátfüggvények ortonormáltak.

Ezenkívül ezek a sajátfüggvények a megfelelő egydimenziós sajátérték feladat (ld. I. 3.1.1. és II. 11.6.5.) sajátfüggvényeinek a szorzata (ill. mint  $v_h^{(k\ell)}$  sajátvektorok : ezen szorzat a sorfolytonosan rendezett belső rácspontokban felvett értékeiből összeállított vektor). A sajátértékek az egydimenziós sajátérték feladat sajátértékeinek az összege.

Régi  $2\sqrt{2}x \leq \sin \pi x$  becslésünket I. 3.1.1-ből alkalmazva és figyelembe véve, hogy  $h_1 N_1 = a$ ,  $h_2 N_2 = b$ , rögtön a

$$\lambda_{k\ell}^h \geq \lambda_{1,1}^h =: \lambda_{\min}^h(A_h) \geq \frac{8}{a^2} + \frac{8}{b^2} \quad (15.44)$$

eredmény adódik. Mivel  $A_h$  szimmetrikus, ez azt jelenti, hogy (15.35) mellett az is igaz, hogy az euklideszi normában, valamint a fenti  $L_2(\omega_h)$ -féle normában

$$\|A_h^{-1}\|_{(2)} = \|A_h^{-1}\|_{(0,h)} \leq \frac{a^2 b^2}{8(a^2 + b^2)}. \quad (15.45)$$

A (15.40) sajátérték feladat (15.41)-(15.42) megoldásához ld. az 5., a (15.45) becsléshez ld. a 6. feladatot.

A (15.16)-(15.17) diszkrét peremérték feladat megoldása a következőképpen vezethető vissza a gyors Fourier szinusz-transzformációra. Elsőnek a peremértékeket elimináljuk, így  $f$ -ből  $\varphi$  lesz,

$$\varphi_{ij} := \begin{cases} f_{ij}, & \text{ha } h_1 < x_{1i} < a-h_1, \quad h_2 < x_{2j} < b-h_2, \\ f_{ij} + \frac{1}{h_1^2} g_j, & \text{ha } x_{1i} = h_1 \text{ és } x_{1i} = a-h_1, \quad h_2 < x_{2j} < b-h_2, \\ f_{ij} + \frac{1}{h_2^2} g_i, & \text{ha } x_{2j} = h_2 \text{ és } x_{2j} = b-h_2, \quad h_1 < x_{1i} < a-h_1, \\ f_{ij} + \frac{1}{h_1^2} g_j + \frac{1}{h_2^2} g_i, & \text{ha } x_{1i} = h_1 \text{ és } x_{1i} = a-h_1, \quad x_{2j} = h_2 \text{ és } x_{2j} = b-h_2, \end{cases}$$

feltéve, hogy  $N_1, N_2 > 2$  (legalább négy belső rácspont van). Ezután az

$$A_h y_h = \varphi_h \quad (15.46)$$

rendszer marad vissza. A csak a belső rácspontoknak megfelelő  $y_{ij}$  ismeretleneket tartalmazó  $y_h$  vektort kiterjeszthetjük homogén peremértékekkel a  $\gamma_h$  diszkrét peremre. Az így keletkező rácsfüggvényt  $y$ -nal jelöljük. Ekkor (15.46) differencia-jelölésben

$$y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} + \varphi = 0, \quad (x_1, x_2) \in \omega_h; \quad y = 0, \quad (x_1, x_2) \in \gamma_h. \quad (15.47)$$

Ezen feladat megoldását az

$$y(x_1, x_2) = \sum_{k=1}^{N_1-1} v^{(k)}(x_1) c^{(k)}(x_2) \quad (15.48)$$

alakban keressük, ahol a  $c^{(k)}$  rácsfüggvényeket meg kell határoznunk,  $v^{(k)}$  pedig az  $(x_1$ -irányú, egydimenziós)

$$v_{\bar{x}_1 x_1} + \lambda^h v = 0, \quad x_1 \in \omega_{h_1}^{(1)} := \{ih_1; i = 1, \dots, N_1 - 1\}; \quad v(0) = v(a) = 0 \quad (15.49)$$

sajátérték feladat  $k$ -adik megoldása :

$$v^{(k)}(x_1) = \sqrt{\frac{2}{a}} \sin k\pi x_1, \quad x_1 \in \omega_{h_1}^{(1)}, \quad (15.50)$$

$$\lambda_k^{h_1} = \left( \frac{2}{h_1} \sin \frac{k\pi}{2N_1} \right)^2, \quad k = 1, \dots, N_1 - 1. \quad (15.51)$$

Az eddig nem használt  $\sqrt{\frac{2}{a}}$  szorzó normálttá teszi a  $v^{(k)}$  függvényeket :

$$\|v^{(k)}\|_{L_2(\omega_{h_1}^{(1)})}^2 = \|v^{(k)}\|_{(0, h_1)}^2 := \sum_{i=1}^{N_1-1} (v^{(k)}(x_{1i}))^2 h_1 = 1. \quad (15.52)$$

A különböző  $k$ -hoz tartozó sajátfüggvények ortogonálisak :

$$(v^{(\ell)}, v^{(m)})_{(0, h_1)} = \sum_{i=1}^{N_1-1} v^{(\ell)}(x_{1i}) v^{(m)}(x_{1i}) h_1 = 0, \quad \ell \neq m. \quad (15.53)$$

Ezután (15.48)-at (15.47)-be helyettesítjük be. Ekkor azt kapjuk (figyelembe véve a (15.49)-(15.53) relációkat), hogy

$$\begin{aligned} -\lambda_k^{h_1} c^{(k)} + c_{\bar{x}_2 x_2}^{(k)} + \varphi^{(k)} &= 0, \quad x_2 \in \omega_{h_2}^{(2)}, \\ c^{(k)}(0) = c^{(k)}(b) = 0, \quad \omega_{h_2}^{(2)} &:= \{jh_2, j = 1, \dots, N_2 - 1\}, \end{aligned} \quad (15.54)$$

ahol

$$\varphi^{(k)} := (\varphi, v^{(k)})_{(0, h_1)}, \quad k = 1, \dots, N_1 - 1,$$

a  $\varphi$  rácsfüggvény Fourier-együtthatói.

A (15.54) feladatok megoldását a rövidített Gauss-eliminációval (ld. I. 1.3.9.) számíthatjuk ki, ami után (15.48) alapján ismert a megoldás.

A számítás menete:

$$\begin{array}{c} f \\ g \end{array} \rightarrow \varphi \xrightarrow{\text{FA}} \{\varphi^{(k)}\} \xrightarrow{\text{GE}} \{c^{(k)}\} \xrightarrow{\text{FS}} \begin{array}{c} y \\ g \end{array}.$$

Itt FA a Fourier-analízist, FS a Fourier-szintézist és GE a Gauss-eliminációt jelöli.

Az eljárás során mindig ugyanazzal az  $(N_1 + 1)(N_2 + 1)$ -méretű tömbbel dolgozunk, ezenkívül még  $N_1(\frac{1}{2} \log_2 N_1 + \frac{3}{2})$  tárhely kell a gyors Fourier-transzformációk során, valamint egy  $N_2$ -méretű segédtömb a Gauss-elimináció alatt. A  $g$  peremértékek kezdettől végig változatlanul maradnak a nagy tömb szélén, míg ezen tömb belsejében  $f$  helyén áll egymás után minden felsorolt mennyiség és végül  $y$ . Az egész eljárás művelet- és tárigényét már a táblázatban megadtuk.

**Megjegyzés.** A Gauss-eliminációt lehet  $x_2$ -irányú Fourier-transzformációkkal helyettesíteni – de ez nem célszerű. Egyrészt ez több műveletbe kerülne ( $O(N_1 N_2)$  helyett  $O(N_1 N_2 \log_2 N_2)$ ), másrészt, ha Gauss-eliminációt vesszünk, akkor változatlan eljárással még az

$$a \frac{\partial^2 u}{\partial x_1^2} + b(x_2) \frac{\partial^2 u}{\partial x_2^2} + c(x_2) \frac{\partial u}{\partial x_2} + d(x_2)u + f(x_1, x_2) = 0, \quad a = \text{const}$$

alakú egyenletek elsőfajú peremérték feladatait is megoldhatjuk.  $\square$

Ehhez a ponthoz ld. a 7. feladatot is.

### 15.3.4 Harmadfajú peremfeltételek

Ezután a (15.13)-(15.14) peremérték feladat helyett a harmadfajú peremérték feladatot vizsgáljuk. A harmadfajú peremfeltétel értelmezéséről már II. 11.2-ben szóltunk.

$$(\Delta u + f)(x) = 0; \quad x \in \{0 < x_1 < a, \quad 0 < x_2 < b\} = \Omega, \quad (15.55)$$

$$\vec{n} \cdot \text{grad } u = \frac{\partial u}{\partial n} = \sigma(x)(u_0(x) - u), \quad x \in \Gamma, \quad (15.56)$$

ahol  $\vec{n}$  a  $\Gamma$  perem külső normálvektora,  $u_0$  a külső közeg hőmérséklete. Továbbá,  $f$  ill.  $\sigma$  adott függvények (a hőforrások sűrűsége ill. a hőátadási tényező), és teljesüljön

$$\sigma(x) \geq \sigma_0 > 0, \quad x \in \Gamma. \quad (15.57)$$

Feltételezzük, hogy ezek a függvények elég simák, úgy hogy  $u$ , a (15.55), (15.56) feladat megoldása,  $C^4(\bar{\Omega})$  eleme. Az  $\Omega$  téglalap diszkretizációjára használjuk a 15.3.1-ben bevezetett  $\bar{\omega}_h = \omega_h \cup \gamma_h$  téglalap alakú rácsot. Itt majd szükség lesz a sarokpontok megkülönböztetésére a többi peremponttól, ezért halmazukot  $\gamma_0$ -val jelöljük.

A (15.55) egyenlet approximációja a belső rácsponokban a megszokott :

$$y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} + f(x) = 0, \quad x \in \omega_h. \quad (15.58)$$

Ezen approximáció  $\psi = f + u_{\bar{x}_1 x_1} + u_{\bar{x}_2 x_2}$  képlethibájára (ahol  $u$  a (15.55)-(15.57) pontos megoldása) ismerjük a (15.23) becslést.

A (15.56) feltétel approximációjához használjuk az egydimenziós esetben (ld. II. 11.4.5.) tanultakat. Ha pl.  $x_1 = a$ ,  $0 < x_2 < b$ , akkor abból indulunk ki, hogy

$$u_{\bar{x}_1, N_1} = \left[ \frac{\partial u}{\partial x_1} - \frac{h_1}{2} \frac{\partial^2 u}{\partial x_1^2} \right]_{x_1=a} + \frac{h_1^2}{6} \frac{\partial^3 u}{\partial x_1^3}(a - \vartheta h_1, x_2), \quad 0 < \vartheta < 1,$$

azaz (15.55), (15.56) alapján

$$\begin{aligned} u_{\bar{x}_1, N_1} &= \left[ \sigma(u_0 - u) + \frac{h_1}{2} \left( f + \frac{\partial^2 u}{\partial x_2^2} \right) \right]_{x_1=a} + O(h_1^2) \\ &= \left[ \sigma(u_0 - u) + \frac{h_1}{2} (f + u_{\bar{x}_2 x_2}) \right]_{x_1=a} + O(h_1^2 + h_2^2). \end{aligned} \quad (15.59)$$

Itt azt használtuk fel, hogy a pontos megoldás feltételezett simasága mellett  $\frac{\partial^2 u}{\partial x_2^2}$  az  $x_1 = a$  peremszakaszon jól definiált, és

$$u_{\bar{x}_2 x_2} - \frac{\partial^2 u}{\partial x_2^2} = \frac{h_2}{6} \left[ \frac{\partial^3 u}{\partial x_2^3}(a, x_2 + \vartheta^+ h_2) + \frac{\partial^3 u}{\partial x_2^3}(a, x_2 - \vartheta^- h_2) \right]$$

érvényes. Ekkor (15.59)-ben képződik egy  $h_1 h_2$ -rendű hibatag, de  $h_1 h_2 \leq \frac{1}{2}(h_1^2 + h_2^2)$ .

Ezért a (15.56) peremfeltétel approximációjaként az  $x_1 = a$  szakaszon fekvő perempontokban a következőt vesszük :

$$-y_{\bar{x}_1} + \frac{h_1}{2} y_{\bar{x}_2 x_2} - \sigma y + \sigma u_0 + \frac{h_1}{2} f = 0, \quad x_1 = a, \quad 0 < x_2 < b. \quad (15.60)$$

Ennek képlethibája

$$\psi := -u_{\bar{x}_1} + \frac{h_1}{2} u_{\bar{x}_2 x_2} - \sigma u + \sigma u_0 + \frac{h_1}{2} f,$$

becslése a fentiek szerint

$$|\psi(x)| \leq \frac{h_1^2}{6} M_3 + \frac{h_1 h_2}{6} M_3 \leq \frac{h_1^2 + h_2^2}{4} M_3, \quad x \in \gamma_h / \gamma_0. \quad (15.61)$$

Teljesen analóg módon kapjuk a téglalap többi oldalán (kivéve a  $\gamma_0$  sarokpontokat) az alábbi differencia-approximációkat :

$$y_{x_1} + \frac{h_1}{2} y_{\bar{x}_2 x_2} - \sigma y + \sigma u_0 + \frac{h_1}{2} f = 0, \quad x_1 = 0, \quad 0 < x_2 < b, \quad (15.62)$$

$$y_{x_2} + \frac{h_2}{2} y_{\bar{x}_1 x_1} - \sigma y + \sigma u_0 + \frac{h_2}{2} f = 0, \quad x_2 = 0, \quad 0 < x_1 < a, \quad (15.63)$$

$$-y_{\bar{x}_2} + \frac{h_2}{2} y_{\bar{x}_1 x_1} - \sigma y + \sigma u_0 + \frac{h_2}{2} f = 0, \quad x_2 = b, \quad 0 < x_1 < a. \quad (15.64)$$

Ezen approximációk képlethibájára most is igaz a (15.61) becslés.

A (15.60), (15.62)-(15.64) approximációkkal kapcsolatban érdemes megemlíteni, hogy van olyan peremfeltétel, amelyben már eredetileg is szerepel a peremmenti irány szerinti második derivált :

Ha a hővezető test felületén jobban vezető vékony réteg van, pl.  $x_1 = 0$ ,  $0 < x_2 < b$ -nél, akkor

$$\frac{\partial u}{\partial x_1} + \kappa \frac{\partial^2 u}{\partial x_2^2} = \sigma u - g$$

a peremfeltétel, amelyben  $\kappa$  pozitív paraméter.

Az elsőfajú peremfeltételek tárgyalásánál a téglalapos megoldási tartomány sarokpontjainak nem jutott semmi szerepe. A harmadfajú peremfeltételek esetén a sarokpontbeli approximációkkal viszont kell foglalkozunk, példaként vesszük a  $(0, b)$  pontot. Itt használjuk azt, hogy

$$\begin{aligned} u_{\bar{x}_2, N_2} &= \left[ \frac{\partial u}{\partial x_2} - h_2 2 \frac{\partial^2 u}{\partial x_2^2} \right]_{x_2=b} + \frac{h_2^2}{6} \frac{\partial^3 u}{\partial x_2^3}(x_1, b - \bar{\eta}h_2), \\ u_{x_1, 0} &= \left[ \frac{\partial u}{\partial x_1} + \frac{h_1}{2} \frac{\partial^2 u}{\partial x_1^2} \right]_{x_1=0} + \frac{h_1^2}{6} \frac{\partial^3 u}{\partial x_1^3}(\bar{\vartheta}h_1, x_2) \end{aligned}$$

(ahol  $\bar{\eta}, \bar{\vartheta} \in (0, 1)$ ), tehát

$$\frac{2}{h_1} u_{x_1, 0} - \frac{2}{h_2} u_{\bar{x}_2, N_2} = \frac{2}{h_1} \frac{\partial u}{\partial x_1} - \frac{2}{h_2} \frac{\partial u}{\partial x_2} + \Delta u + O(h_1 + h_2). \quad (15.65)$$

A (15.56) peremfeltétel alapján igaz

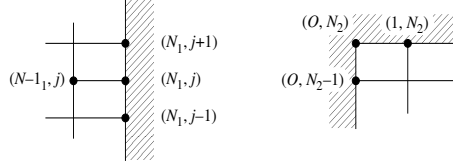
$$\frac{2}{h_1} \frac{\partial u}{\partial x_1} - \frac{2}{h_2} \frac{\partial u}{\partial x_2} = \frac{2\sigma}{h_1} (u - u_0) + \frac{2\sigma}{h_2} (u - u_0) = \frac{\sigma}{H} (u - u_0), \quad H := \frac{h_1 h_2}{2(h_1 + h_2)}.$$

Így a (15.55) differenciálegyenletet is figyelembe véve, (15.65)-ből következik, hogy a megfelelő approximáció :

$$H \left[ \frac{2}{h_1} y_{x_1, 0} - \frac{2}{h_2} y_{\bar{x}_2, N_2} + f \right] - \sigma (y - u_0) = 0, \quad \text{ha } (x_1, x_2) = (0, b) \in \gamma_0. \quad (15.66)$$

Ennek  $\psi$  képlethibája ugyanis a

$$|\psi(x)| \leq H \frac{h_1 + h_2}{3} M_3 \leq \frac{h_1^2 + h_2^2}{12} M_3, \quad x \in \gamma_0, \quad (15.67)$$



ábra 15.6: Harmadfajú peremfeltétel differenciacsillagjai

becslésnek tesz eleget. A többi sarokpontban érvényes approximációk hasonlók :

$$H \left[ \frac{2}{h_1} y_{x_1,0} + \frac{2}{h_2} y_{x_2,0} + f \right] - \sigma(y - u_0) = 0, \quad (x_1, x_2) = (0, 0), \quad (15.68)$$

$$H \left[ \frac{2}{h_1} y_{\bar{x}_1, N_1} + \frac{2}{h_2} y_{x_2,0} + f \right] - \sigma(y - u_0) = 0, \quad (x_1, x_2) = (a, 0), \quad (15.69)$$

$$H \left[ \frac{2}{h_1} y_{\bar{x}_1, N_1} - \frac{2}{h_2} y_{\bar{x}_2, N_2} + f \right] - \sigma(y - u_0) = 0, \quad (x_1, x_2) = (a, b). \quad (15.70)$$

Ehhez ld. a 15. feladatot is. (15.68)-(15.70) hibájára ugyancsak érvényes a (15.67) becslés. Ezzel elkészült a differenciaséma : Az  $\bar{\omega}_h$  pontjaihoz tartozó ismeretleneket (pl. sorfolytonosan)  $\bar{y}_h$  vektorba rendezve és az  $y$ -nal kapcsolatos kifejezéseket (15.58)-ban, (15.60)-ban, (15.62)-(15.64)-ben, (15.66)-ban és (15.68)-(15.70)-ben a másik oldalra átrendezve kapjuk az

$$\bar{A}_h \bar{y}_h = \bar{\varphi}_h \quad (15.71)$$

egyenletrendszer. Ennek  $\bar{n}$  dimenziója megint  $h_1$ -től és  $h_2$ -től függ, pontosabban  $\bar{n} := (N_1 + 1)(N_2 + 1)$ .

Az  $N_1 = N_2 = 2$  speciális esetben a mátrix így néz ki :

$$\bar{A}_h = \begin{pmatrix} \alpha_{00} & -\frac{2H}{h_1^2} & 0 & -\frac{2H}{h_2^2} & 0 & 0 & 0 & 0 & 0 \\ -\frac{h_2}{2h_1^2} & \alpha_{10}^{(2)} & -\frac{h_2}{2h_1^2} & 0 & -\frac{1}{h_2} & 0 & 0 & 0 & 0 \\ 0 & -\frac{2H}{h_1^2} & \alpha_{20} & 0 & 0 & -\frac{2H}{h_2^2} & 0 & 0 & 0 \\ -\frac{h_1}{2h_2^2} & 0 & 0 & \alpha_{01}^{(1)} & -\frac{1}{h_1} & 0 & -\frac{h_1}{2h_2^2} & 0 & 0 \\ 0 & -\frac{1}{h_2^2} & 0 & -\frac{1}{h_1^2} & \alpha & -\frac{1}{h_1^2} & 0 & -\frac{1}{h_2^2} & 0 \\ 0 & 0 & -\frac{h_1}{2h_2^2} & 0 & -\frac{1}{h_1} & \alpha_{21}^{(1)} & 0 & 0 & -\frac{h_1}{2h_2^2} \\ 0 & 0 & 0 & -\frac{2H}{h_2^2} & 0 & 0 & \alpha_{02} & -\frac{2H}{h_1^2} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{h_2} & 0 & -\frac{h_2}{2h_1^2} & \alpha_{12}^{(2)} & -\frac{h_2}{2h_1^2} \\ 0 & 0 & 0 & 0 & 0 & -\frac{2H}{h_2^2} & 0 & -\frac{2H}{h_1^2} & \alpha_{22} \end{pmatrix}.$$



Itt használtuk az  $\alpha := \frac{2}{h_1^2} + \frac{2}{h_2^2}$ , valamint a következő jelöléseket :

$$\alpha_{ij} := \frac{2H}{h_1^2} + \frac{2H}{h_2^2} + \sigma_{ij}, \quad \alpha_{ij}^{(1)} := \frac{h_1}{h_2^2} + \frac{1}{h_1} + \sigma_{ij}, \quad \alpha_{ij}^{(2)} := \frac{h_2}{h_1^2} + \frac{1}{h_2} + \sigma_{ij}.$$

Az  $N_1$  és  $N_2$  számok növekedésével nő a belső pontok száma, amelyekhez olyan sorok és oszlopok tartoznak a mátrixban, mint fent az ötödik.

Ez az  $\bar{A}_h$  mátrix akkor szimmetrikussá válik, ha a sarokpontokhoz tartozó egyenleteket  $h_1 h_2 / (4H) = (h_1 + h_2) / 2$ -vel szorozzuk meg, a (15.60) és (15.62) egyenleteket  $h_2$ -vel, a (15.63), (15.64) egyenleteket  $h_1$ -gyel és a belső pontok egyenleteit  $h_1 h_2$ -vel.

Az egyenletrendszer jobboldalának komponensei :

$$\varphi_\ell := \begin{cases} f_\ell, & \text{ha } x^\ell = (x_{1\ell}, x_{2\ell}) \in \omega_h, \\ \frac{h_1}{2} f_\ell + \sigma_\ell u_{0\ell}, & \text{ha } x^\ell \in \gamma_h, \quad 0 < x_{2\ell} < b, \\ \frac{h_2}{2} f_\ell + \sigma_\ell u_{0\ell}, & \text{ha } x^\ell \in \gamma_h, \quad 0 < x_{1\ell} < a, \\ H f_\ell + \sigma_\ell u_{0\ell}, & \text{ha } x^\ell \in \gamma_0. \end{cases} \quad (15.72)$$

A (15.71)-(15.72) differenciaséma stabilitási és konvergencia vizsgálata megint az M-mátrix elmélet segítségével történik.

**15.11. Tétel** (harmadfajú peremérték feladat differencia approximációjának stabilitása és konvergenciája). Ha igaz a (15.57) feltétel, akkor a (15.71)-(15.72) differenciaséma stabil : érvényes a

$$\|\bar{y}_h\|_{C(\bar{\omega}_h)} \leq \left( \frac{a^2 + b^2}{16} + \frac{4 + a + b}{4\sigma_0} \right) \|\bar{\varphi}_h\|_{C(\bar{\omega}_h)}$$

becslés. Ha a (15.55)-(15.57) feladat  $u$  pontos megoldása teljesíti az  $u \in C^4(\bar{\Omega})$  feltételt, akkor a séma másodrendben konvergens.

**Bizonyítás.** Vezessük be a következő majoráns függvényt :

$$w(x) := c + x_1(a - x_1) + x_2(b - x_2),$$

ahol a  $c$  egy pozitív konstáns, amelyet lejjebb alkalmasan választunk meg. A  $w_\ell := w(x^\ell)$  számokból a  $\vec{w}_h$  vektort állítjuk össze. A majoráns függvény és a differenciaséma képleteiből számítjuk ki azt, hogy

$$(\bar{A}_h \vec{w}_h)_\ell = \begin{cases} 4, & \text{ha } x^\ell \in \omega_h, \\ \sigma_\ell w_\ell - a + 2h_1, & \text{ha } x^\ell \in \gamma_h, \quad 0 < x_{2\ell} < b, \\ \sigma_\ell w_\ell - b + 2h_2, & \text{ha } x^\ell \in \gamma_h, \quad 0 < x_{1\ell} < a, \\ \sigma_\ell w_\ell + H \left[ 2 - 2\frac{a}{h_1} - 2\frac{b}{h_2} \right], & \text{ha } x^\ell \in \gamma_0. \end{cases}$$

Mivel

$$0 < H < \frac{h_1}{2}, \quad H < \frac{h_2}{2}, \quad w_\ell \geq c > 0 \quad \text{és} \quad \sigma_\ell \geq \sigma_0 > 0,$$

így minden  $x^\ell \in \gamma_h$ -ra

$$(\bar{A}_h \bar{w}_h)_\ell \geq \sigma_0 c - a - b \geq 4, \quad \text{ha} \quad c = \frac{4 + a + b}{\sigma_0}.$$

Ekkor minden  $1 \leq \ell \leq \bar{n}$ -re igaz  $(\bar{A}_h \bar{w}_h)_\ell \geq 4$ , és mivel  $0 < w_\ell \leq c + \frac{a^2 + b^2}{4}$ , ezért :

$$\|(\bar{A}_h)^{-1}\|_{C(\bar{\omega}_h)} \leq \frac{a^2 + b^2}{16} + \frac{4 + a + b}{4\sigma_0} =: c_1, \quad (15.73)$$

tehát (15.72)-t is figyelembe véve

$$\|\bar{y}\|_{C(\bar{\omega}_h)} \leq c_1 \|\bar{\varphi}\|_{C(\bar{\omega}_h)} = c_1 \max \left( \|f\|_{C(\bar{\Omega})}, \frac{a + b}{2} \|f\|_{C(\Gamma)} + \|\sigma u_0\|_{C(\Gamma)} \right).$$

Ezzel a tétel becslését beláttuk, amely a séma stabilitását jelenti.

Legyen most  $\bar{z}_h := \bar{y}_h - \bar{u}_h$  a hibavektor! Ez az

$$\bar{A}_h \bar{z}_h = \bar{\psi}_h$$

egyenletrendszernek tesz eleget, és (15.73), valamint (15.23), (15.61), (15.67) alapján következik

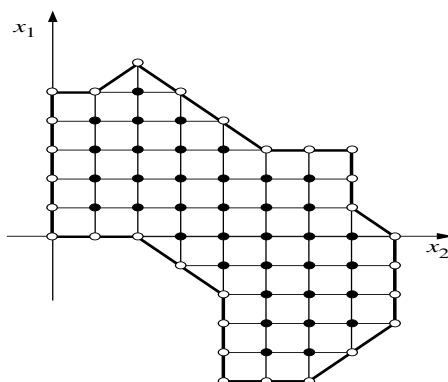
$$\|\bar{z}_h\|_{C(\bar{\omega}_h)} \leq c_1 \|\bar{\psi}_h\|_{C(\bar{\omega}_h)} \leq \frac{c_1}{4} (h_1^2 + h_2^2) \max \left( \frac{1}{3} M_4, M_3 \right) = O(h_1^2 + h_2^2). \quad \square$$

A (15.71) egyenletrendszer megoldása itt is a gyors Fourier–transzformációra vezethető vissza – ill. még hatékonyabban a többrácsos módszerrel végezhető el, ld. 15.4.

Ha (15.56) helyett adott a

$$\frac{\partial u}{\partial n} = g(x), \quad x \in \Gamma, \quad (15.74)$$

*másodfajú peremfeltétel* – ennek értelmezéséhez ld. II-ben a 11.2. pontot –, akkor a (15.60), (15.62)-(15.64) ill. (15.66), (15.68)-(15.70) approximációkban először  $\sigma u_0$ -t helyettesítjük  $g$ -vel, majd ezután vegyük  $\sigma = 0$ -t. Amennyiben a peremen mindenütt adott a (15.74) feltétel, akkor a (15.55), (15.74) feladat szinguláris, hiszen  $u \equiv \text{const}$  tetszőleges konstanssal elégíti ki az egyenleteket  $f \equiv 0$ ,  $g \equiv 0$  esetén (szinguláris másodfajú peremérték feladatot az egydimenziós esetben II. 11.4.11-ben tárgyaltuk). Ha a peremfeltételek nem csak másodfajúak, hanem a perem egyes szakaszán ahelyett elsőfajú vagy harmadfajú feltétel szerepel, akkor (megfelelő simaságot feltételezve) a feladat egyetlen megoldással rendelkezik és a leírt differencia-approximáció M-mátrix-eljárással vizsgálható (ld. pl. a 8. feladatot).



ábra 15.7: A rácsra illeszkedő tartomány

### 15.3.5 A Poisson–egyenlet általános tartományban

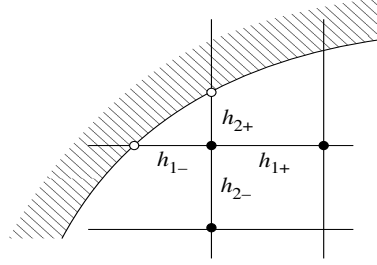
Következőnek tárgyaljuk azt a kérdést, amely többdimenziós feladatoknál (az egydimenziós feladatokhoz képest) új bonyodalmat jelent: Ha  $\Omega$  nem téglalap vagy téglalapokból összeállított olyan tartomány, amelyre ráborítható az ekvidisztáns  $h_1, h_2$  lépéstávolságú  $\omega_h$  rács, akkor hogyan approximáljuk a (15.13)-(15.14) differenciálegyenletet? Legyen  $\Omega$  véges, összefüggő,  $\Gamma$  pereme viszont szakaszonként sima.

Elsőnek vegyük észre, hogy a (15.16)-(15.17) approximáció addig értelmes, míg a perem csak rácspontokban metszi a rácsot alkotó egyeneseit és míg minden belső rácspontra igaz, hogy az öt négy szomszédjával összekötő szakaszok mind  $\bar{\Omega}$ -ban vannak. Speciális tartományokra mindkét feltétel teljesíthető, ld. a 7. ábrát. Általános tartomány esetén a második feltétel várhatóan a rács megfelelő finomításával biztosítható, de már az első feltétel kizár sok érdekes tartományt.

A probléma megoldására több ötlet merül fel. A legegyszerűbb az, ha  $\Omega$  belsejében mindenütt téglalap alakú ekvidisztáns rácsot használunk, és a rács valamelyik egyenesére illeszkedő perempontból visszük át az ott adott peremértéket abba a belső rácspontba, amely távolsága a peremhez kisebb mint  $h_1$ , ill.  $h_2$ . Sajnos ezzel általában csak elsőrendű konvergenciát tudunk biztosítani (ld. a 9.a) feladatot)! Egyszerűsége miatt mégis ez az approximáció eléggé elterjedt.

A második és harmadik lehetőség az, hogy vagy lemondunk a rács ekvidisztanciájáról, de a peremhez illesztjük, vagy pedig  $\bar{\omega}_h$ -hoz csatolhatjuk a  $\Gamma$  és a rács egyenesének a metszéspontjait.

Mint negyedik lehetőséget megemlíjtük azt, hogy transzformált téglalap alakú rácsra mehetünk át: az  $\Omega$  tartományt leképezzük egy téglalagra és ott használjuk a téglalapos rácsot. Ez gyakran eléggé bonyolítja a differenciál-



ábra 15.8: *Shortley–Weller approximáció a perem közelében*

egyenletet, amelyet ugyancsak kell transzformálni, de pl. a kör esetén igen hasznos (ld. a 10. feladatot).

Egyes tartományoknál nem nehéz a második lehetőség megvalósítása, azaz olyan rács megszerkesztése az adott peremből kiindulva, hogy egyeseinek metszéspontjai a peremmel mindig rácspontok legyenek.

De gyakran a kötöttség a peremhez kellemetlen programozási szempontból (ld. a 11. feladatot).

Foglalkozzunk ezután részletesen a harmadik ötlettel! Ilyenkor a perem közelében nemekvidisztáns approximációt kell bevetni (a II. 11.4.7.- pontban látottak mintájára), a tartomány belsejében továbbra is az ekvidisztáns rácsot használhatjuk. Annak lépéstávolságait ezután is  $h_1$ -gyel, ill.  $h_2$ -vel jelöljük. A konstrukció szerint ezek most a rács maximális lépéstávolságai. Az approximáció leírására használjuk a következő jelöléseket :

$$(u_{x_1})_{\tilde{x}_1} := \frac{(u_{x_1})_+ - (u_{x_1})_-}{\tilde{h}_{1i}}, \quad (u_{x_2})_{\tilde{x}_2} := \frac{(u_{x_2})_+ - (u_{x_2})_-}{\tilde{h}_{2j}}, \quad (15.75)$$

$$(u_{x_1})_{\pm} := \pm(u_{i\pm 1,j} - u_{ij})/h_1^{\pm}, \quad (u_{x_2})_{\pm} := \pm(u_{i,j\pm 1} - u_{ij})/h_2^{\pm}.$$

A  $h_1^{\pm}$  és  $h_2^{\pm}$  lépéstávolságok igenis függhetnek  $(i, j)$ -től, de ezt nem jelöljük. Képletei (v.ö. 11.4.7-tel) :

$$h_{1,i\pm 1/2} = h_1^{\pm} := |x_{1,i\pm 1} - x_{1,i}|, \quad h_{2,j\pm 1/2} = h_2^{\pm} := |x_{2,i\pm 1} - x_{2,i}|, \\ \tilde{h}_{1i} := \frac{1}{2}(h_1^+ + h_1^-), \quad \tilde{h}_{2j} := \frac{1}{2}(h_2^+ + h_2^-).$$

A Laplace-operátor approximációja, részletesen kiírva :

$$(\Delta u)_{ij} \approx \frac{2}{h_1^+ + h_1^-} \left( \frac{u_{i+1,j} - u_{ij}}{h_1^+} - \frac{u_{ij} - u_{i-1,j}}{h_1^-} \right) \\ + \frac{2}{h_2^+ + h_2^-} \left( \frac{u_{i,j+1} - u_{ij}}{h_2^+} - \frac{u_{ij} - u_{i,j-1}}{h_2^-} \right) \quad (15.76) \\ =: -(\bar{A}_{S,h} \bar{u}_h)_{ij}.$$

A peremre eső  $(i, j)$  pontokban az egységmátrix megfelelő sora adja az  $\bar{A}_{S,h}$  sorát.

A (15.76) képlet – amelyet *Shortley–Weller approximáció*nak hívunk – részben megoldja a felvetett problémát, ugyanis az a tény, hogy a perem közelében csak *elsőrendű az approximáció* (ez 11.4.7. alapján belátható), nem olyan jelentős: az eredő diszkrét megoldás *pontossága mégis másodrendű*. A képlettel viszont az a gond, hogy még az elsőfajú peremfeltételek eliminációja után az  $\bar{A}_{S,h}$ -ból kapható  $A_{S,h}$  mátrix sem lesz szimmetrikus. Lehet, hogy a mátrix a pozitív definittségét is elveszti.

Először lássuk be, hogy a (15.18)-nak megfelelő  $\bar{\varphi}_h$  jobboldallal felírt

$$\bar{A}_{S,h}\bar{y}_h = \bar{\varphi}_h \quad (15.77)$$

rendszer  $\bar{y}_h$  megoldása létezik, egyértelmű – és hibája másodrendű. Ennek bizonyítása során az  $\bar{\omega}_h$  rács  $(x_{1i}, x_{2j})$  pontjainak most is könnyen elkészíthető,  $\ell$ -lel indexelt sorfolytonos  $x^\ell = (x_{1\ell}, x_{2\ell})$  átszámozására hivatkozunk.

**15.12. Tétel** (a Shortley–Weller séma stabilitása és pontossága). A (15.77) rendszer mátrixa reguláris és stabil; érvényes a következő becslés :

$$\|(\bar{A}_{S,h})^{-1}\|_{C(\bar{\omega}_h)} \leq 1 + \frac{1}{4}(\text{diam}(\Omega))^2. \quad (15.78)$$

Amennyiben  $u \in C^4(\bar{\Omega})$ , akkor (15.77) megoldásának pontossága másodrendű,

$$\|\bar{u}_h - \bar{y}_h\|_{C(\bar{\omega}_h)} = O(h_1^2 + h_2^2).$$

**Bizonyítás.** Az  $\bar{A}_{S,h}$  mátrix rendelkezik az M-tulajdonsággal. Először ugyanis vegyük észre, hogy a mátrix elemeinek az előjeleloszlása megfelelő, ld. (15.76). Következőnek azt mutatjuk meg, hogy

$$w(x) := 4 + c - (x_1 - x_{1,0})^2 - (x_2 - x_{2,0})^2$$

majoráns függvény és majoráns vektor  $\vec{w}_h$ , a  $w$  értékeiből összeállított vektor, ha  $c := (\text{diam}(\Omega))^2$  és  $(x_{1,0}, x_{2,0}) \in \Omega$ . A  $c$  konstans választása azt eredményezi, hogy  $w_\ell = w(x^\ell) \geq 4$  minden  $\ell$ -re. Továbbá, mivel

$$\begin{aligned} \frac{w_{i+1,j} - w_{ij}}{h_1^+} &= \frac{\partial w}{\partial x_1}(x_{1i} + \frac{h_1^+}{2}, x_{2j}) = -2 \left( x_{1i} + \frac{h_1^+}{2} - x_{1,0} \right), \\ \frac{w_{ij} - w_{i-1,j}}{h_1^-} &= \frac{\partial w}{\partial x_1}(x_{1i} - \frac{h_1^-}{2}, x_{2j}) = -2 \left( x_{1i} - \frac{h_1^-}{2} - x_{1,0} \right) \end{aligned}$$

stb., így

$$(\overline{A}_{S,h}\vec{w}_h)_\ell = \begin{cases} 4, & \text{ha } x^\ell \in \omega_h \\ w_\ell, & \text{ha } x^\ell \in \gamma_h \end{cases} \geq 4.$$

Ezért  $\overline{A}_{S,h}$  valóban M-mátrix és az I. 1.8. lemma alapján a (15.78) becslés teljesül, hiszen  $\|\vec{w}_h\|_{C(\overline{\omega}_h)} \leq \|w\|_{C(\overline{\Omega})} \leq 4 + c$ .

Most legyen  $\vec{z}_h := \vec{y}_h - \vec{u}_h$  a hiba és  $\vec{\psi}_h := \overline{A}_{S,h}(\vec{y}_h - \vec{u}_h)$  a képlethiba :

$$\psi_\ell := \begin{cases} (f + (u_{x_1})_{\vec{x}_1} + (u_{x_2})_{\vec{x}_2})_\ell, & \text{ha } x^\ell \in \omega_h, \\ 0, & \text{ha } x^\ell \in \gamma_h. \end{cases}$$

Ezzel a  $\vec{\psi}_h$  vektorral a hibaegyenlet

$$\overline{A}_{S,h}\vec{z}_h = \vec{\psi}_h. \quad (15.79)$$

Ennek becslésére alkalmazhatjuk a 15.7. lemmát, de nem a fenti, hanem (változatlan  $c$  konstanssal) a

$$w(x) := c - (x_1 - x_{1,0})^2 - (x_2 - x_{2,0})^2$$

majoráns függvényből kiindulva, mert a  $\gamma_h$ -beli pontokat már nem kell figyelembe vennünk. A peremen a majoráns függvénnyel képzett  $\overline{A}_{S,h}\vec{w}_h$  vektor komponensei nullák is lehetnek. Ha viszont  $x^\ell \in \omega_h$ , akkor továbbra is  $(\overline{A}_{S,h}\vec{w}_h)_\ell = 4$ . Így (15.79) megoldására igaz a

$$\|\vec{y}_h - \vec{u}_h\|_{C(\omega_h)} = \frac{c}{4} \|\psi_h\|_{C(\omega_h)} \quad (15.80)$$

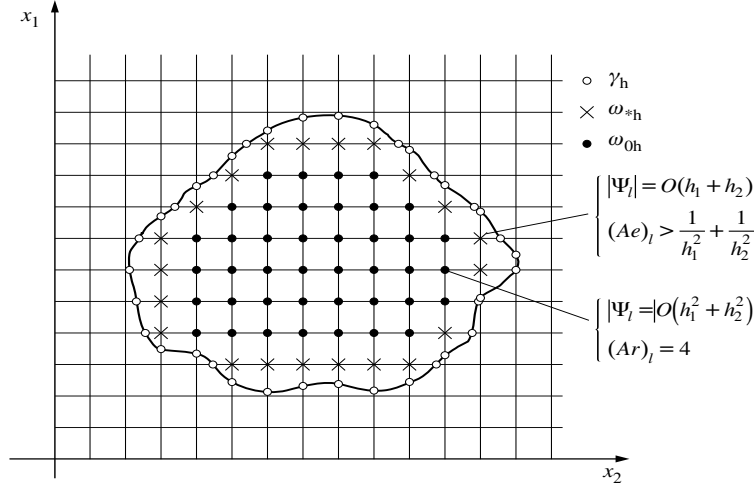
becslés. Ez még nem mutatja meg a másodrendű konvergenciát, hanem csak az elsőrendűt. Legyen ugyanis  $\omega_h^*$  azon  $(x_{1i}, x_{2j})$  rácspontok halmaza, ahol a rács nem ekvidisztáns, azaz  $h_{1,i+1/2} \neq h_{1,i-1/2}$  vagy  $h_{2,j+1/2} \neq h_{2,j-1/2}$ . Az  $\omega_h^*$ -beli  $(x_{1i}, x_{2j})$  pontokban a II. 11.4.7. pont (105), (106) sorfejtései szerint  $\psi_{ij} = O(h_1 + h_2)$ .

Az egyéb rácspontokban viszont teljesül (15.23).

Az említett sorfejtésekből, valamint a séma stabilitásából egyébként már az is következik, hogy abban az esetben másodrendű a pontosság, amikor a rács bár mindenütt nemekvidisztáns, de lépéstávolságai csak lassan változnak abban az értelemben, hogy

$$|h_1^+ - h_1^-| = O(h_1^2), \quad |h_2^+ - h_2^-| = O(h_2^2)$$

mindenütt. Az ilyen feltételek általában nem adnak lehetőséget a perem jó követéséhez, de elegendőek ahhoz, hogy a tartomány egyes kiválasztott



ábra 15.9: Shortley–Weller approximáció vizsgálata

részeiben vagy pl. a perem egyenes szakaszának közelében sűrítjük a rácsot megfelelő mértékben.

A továbbiakban csak azt az esetet vizsgáljuk – a Shortley–Weller approximáció szerkesztésénknél megfelelően – hogy  $\omega_h$ -nak minden olyan pontjában, amelynek nincs szomszédpontja  $\gamma_h$ -ből, a lépéstávolságok ekvidisztánsok. Ezen pontok halmazát  $\omega_h^0 := \omega_h / \omega_h^*$ -val jelölve, tekintsük a hibavektor következő felbontását :

$$\begin{aligned} \bar{z}_h &= \bar{z}_h^* + \bar{z}_h^0; & z_\ell &= z_\ell^0 = z_\ell^* = 0, & x^\ell &\in \gamma_h, \\ (\bar{A}_{S,h} \bar{z}_h^*)_\ell &= (\psi_h^*)_\ell = \begin{cases} O(h_1 + h_2), & \text{ha } x^\ell \in \omega_h^* \\ 0, & \text{ha } x^\ell \in \omega_h^0 \end{cases}, \\ (\bar{A}_{S,h} \bar{z}_h^0)_\ell &= (\psi_h^0)_\ell = \begin{cases} 0, & \text{ha } x^\ell \in \omega_h^* \\ O(h_1^2 + h_2^2), & \text{ha } x^\ell \in \omega_h^0 \end{cases}. \end{aligned}$$

(15.80)-ból következik az  $\bar{A}_{S,h} \bar{z}_h^0 = \bar{\psi}_h^0$  feladat megoldására, hogy  $\bar{z}_h^0 = O(h_1^2 + h_2^2)$ . Az

$$\bar{A}_{S,h} \bar{z}_h^* = \bar{\psi}_h^* \quad (15.81)$$

egyenlet megoldását külön becsüljük. A homogén peremértékek eliminációja után visszamaradó mátrixot  $A_{S,h}$ -vel jelöljük. Ez az M-mátrix teljesíti a  $w \equiv 1$  majoráns függvénynek megfelelő  $\vec{w}_h = e_h := (1, \dots, 1)^T$  vektorral az  $A_{S,h} e_h \geq 0$  feltételt. Ugyanis az  $\omega_h^0$ -beli  $x^\ell$  pontokban érvényes  $(A_{S,h} e_h)_\ell = 0$ . Ha egyébként például valamelyik  $\omega_h^*$ -beli  $x^\ell = (x_{1i}, x_{2j})$  pontban egyik  $x_1$ -

irányban ill. egyik  $x_2$ -irányban van szomszédpont  $\gamma_h$ -ból, akkor ott

$$(A_{S,h}e_h)_\ell = \frac{2}{(h_1^+ + h_1^-)h_1^+} \geq \frac{1}{h_1^2}$$

ill.

$$(A_{S,h}e_h)_\ell = \frac{2}{(h_2^+ + h_2^-)h_2^+} \geq \frac{1}{h_2^2},$$

tehát mindkét esetben

$$0 < \frac{1}{h_1^2 + h_2^2} < (A_{S,h}e_h)_\ell. \quad (15.82)$$

Itt előbb azt használtuk, hogy  $h_1$  és  $h_2$  a maximális lépéstávolságok, majd azt, hogy  $A_{S,h}$  M-mátrix.

A (15.82) becslés akkor is igaz, ha több mint egy szomszédpont  $\gamma_h$ -ból való. Ezért most a 15.7. lemmára hivatkozhatunk, amely szerint (15.81) megoldásának becslése

$$\|z_h^*\|_{C(\omega_h)} \leq \frac{\|\psi_h^*\|_{C(\omega_h)}}{\frac{1}{h_1^2 + h_2^2}} = \|\psi_h^*\|_{C(\omega_h)}(h_1^2 + h_2^2) = O(h_1^3 + h_2^3).$$

Ennek alapján  $\|z_h^*\|_{C(\omega_h)}$  még akkor is másodrendű lenne, ha csak  $\|\psi_h^*\|_{C(\omega_h)} = O(1)$  volna!  $\square$

**Megjegyzés.** A meglepő eredményt, hogy a képlethiba mindenütt a perem közelében lehet  $O(h_1 + h_2)$  ill. sőt  $O(1)$  a séma másodrendűségének csorbítása nélkül, azzal lehet magyarázni, hogy ott a diszkrét Green-féle függvény értékei (az  $\frac{1}{h_1 h_2} A_{S,h}^{-1}$  mátrix elemei, v.ö. II. 11.3-mal és 11.4.4-gyel) már közeliek nullához. A Poisson-egyenlet diszkrét Green-féle függvényével majd 15.5.1-ben foglalkozunk részletesebben.  $\square$

Tehát a Shortley–Weller approximáció pontosság szempontjából elfogadható (ehhez ld. a 12. feladatot is). A keletkező egyenletrendszer megoldása a többrácsos módszerrel hatékonyan kiszámítható.

Az approximációnak azon hátrányának kijavítására, hogy a peremközeli approximáció miatt általában megsérül az  $A_{S,h}$  mátrix szimmetriája, a fenti tétel után kézenfekvő a következő ötlet. Használjuk (15.76) helyett az alábbi approximációt :

$$\begin{aligned} (\Delta u)_{ij} &\approx \frac{1}{h_1} \left( \frac{u_{i+1,j} - u_{ij}}{h_1^+} - \frac{u_{ij} - u_{i-1,j}}{h_1^-} \right) + \frac{1}{h_2} \left( \frac{u_{i,j+1} - u_{ij}}{h_2^+} - \frac{u_{ij} - u_{i,j-1}}{h_2^-} \right) \\ &=: -(\bar{A}_{S,h} \bar{u}_h)_{ij}. \end{aligned} \quad (15.83)$$



Ekkor is a rácsegyenesek és a perem metszéspontjai hozzátartoznak a rácshoz, és így a pontos peremértékeket átvesszük – de a hozzátartozó  $A_{S,h}$  mátrix szimmetrikus lesz. A differenciálegyenlet approximációja viszont a peremközeli pontokban általában csak  $O(1)$ -rendű (13. feladat) – ami a 15.12. tétel bizonyításának alapján mégis másodrendű konvergenciát biztosít.

A tetszőleges tartományban megadott Poisson–egyenlet elsőfajú peremérték feladat approximációjára más ötletek is vannak, ld. a 9. feladatot.

Harmadfajú peremfeltételekre már nehezebb differencia-approximációt levezetni akkor, amikor a tartomány pereme általánosabb görbe (ld. a 14. feladat). Ennek kényelmes levezetését (általánosabb, változó együtthatójú elliptikus egyenletek esetén is) 15.6-ban adjuk meg, a boxmódszer keretében.

## 15.4 A többrácsos módszer

### 15.4.1 Az alapötlet

Ez a módszer (angolul: multigrid) az  $N$  ismeretlenű, diszkrét elliptikus egyenletrendszer közelítő megoldását a diszkretizációnak megfelelő pontossággal,  $O(N)$  művelettel, valamint  $O(N)$  tárigénnyel állítja elő, tehát nagyságrend szerint a módszer optimális.

Az alapötlet a következő : Ha a differenciálegyenlet és a peremfeltétel diszkretizációjával kapott

$$A_h y_h = b_h \quad (15.84)$$

egyenletrendszer  $y_h$  megoldásához  $y_h^0$  a tetszőleges közelítése, akkor

$$y_h^1 = y_h^0 + w_h \quad (15.85)$$

máris a pontos megoldás, amennyiben

$$A_h w_h = r_h, \quad r_h := b_h - A_h y_h^0. \quad (15.86)$$

Abból, hogy a (15.84) egyenletrendszer helyett a (15.86)-ot oldjuk meg, még semmi nyereség nem származik, mivel (15.86) ugyanolyan feladat, mint az eredeti (15.84). Viszont ha most  $w_h$ -t csak közelítőleg állítjuk elő olyan egyenletrendszerből, amely az eredeti peremérték feladat egy *durvább rácson* való diszkretizációjából ered, akkor  $w_h$  kiszámítása kevesebb műveletbe kerül, és számíthatunk arra is, hogy a (15.85) szerinti  $y_h^1$  jobb közelítés  $y_h$ -hoz, mint  $y_h^0$ .

Kézenfekvő ezt iteratív módon ismételni. Két rácson dolgozva, a finom  $\omega_h$  rácson és a durva  $\omega_H$  rácson, és a hozzátartozó vektorok különböző dimenzióját figyelembe véve, a következő iteráció adódik:

1.  $y_h^0$  adott,  $i := 0$
2.  $r_h := b_h - A_h y_h^i$
3.  $A_H w_H = r_H := R_h^H r_h$  ( $R_h^H$  a restrikciós vagy leszűkítő operátor)
4.  $w_h := P_H^h w_H$  ( $P_H^h$  interpolációs, más néven *prolongációs operátor*)
5.  $y_h^{i+1} := y_h^i + w_h$
6.  $i := i + 1, \rightarrow 2.$

Itt az  $R_h^H$  olyan ( $h$ -tól és  $H$ -tól függő) lineáris operátor (azaz téglalap alakú mátrix), amely az  $\omega_h$ -n definiált diszkrét argumentumú függvényeket ( $n(h)$ -dimenziójú vektorokat) az  $\omega_H$ -n definiált rácsgüggvényekre ( $n(H)$ -dimenziójú vektorokra) képezi le, és  $n(h) \gg n(H)$ .

A  $P_H^h$  lineáris operátor (ugyancsak téglalap alakú mátrix) fordított irányban hat, interpolációt hajt végre. (Az  $I$  betű kevésbé alkalmas a prolongációs operátor jelölésére, összetéveszhetőnek tenné azt az egységmátrixszal).

A nyereség tehát most az, hogy az  $A_h$ -val kapcsolatos egyenletrendszer helyett az  $A_H$ -val kapcsolatos egyenletrendszerrel kell foglalkoznunk, amiben kevesebb ismeretlen szerepel. Tipikusan  $n(h) \approx 4n(H)$  – amikor az eredeti peremérték feladat két-dimenziós és  $H = 2h$ . Ekkor – az  $A_h$  ritkamátrixú egyenletrendszer direkt megoldásának költségét  $O(n^2(H))$ -tel becsülve – a nyereség az, hogy a durva rácson már a finomrácsú egyenletrendszer költségének kb. 16-odára van csak szükség!

Viszont nincs okunk  $\omega_H$ -nál megállni és az  $A_H w_H = r_H$  egyenletet pontosan megoldani:  $\omega_H$ -ról még durvább rácshoz mehetünk át stb., így eljutva a többrácsos módszerhez.

Sajnos az így első nekifutásra szerkesztett iteráció általában nem konvergens.

Mivel ugyanis  $\omega_h$ -hoz több pont tartozik mint  $\omega_H$ -hoz, van olyan  $v_h \neq 0_h$  vektor, amelyre  $R_h^H v_h = 0_H$  (azaz  $R_h^H$  magtere nem triviális). Ha most  $r_h = v_h$ , akkor

$$r_H = R_h^H v_h = 0_H, \quad w_H = 0_H, \quad w_h = P_H^h w_H = 0_h,$$

és így  $y_h^1 = y_h^0$ . Az iteráció ekkor helyben topog, és  $y_h^0$  nem megoldás, hiszen  $r_h = b_h - A_h y_h^0 \neq 0_h$ .

Másképpen magyarázva a problémát:

A durva rácstról való visszatéréskor

$$y_h^{i+1} = y_h^i + P_H^h w_H$$

formában akarjuk a finomrácsú megoldást javítani, tehát  $P_H^h w_H$  akkor a legjobb, ha megegyezik  $y_h^i$  hibájával. A fenti példában az  $y_h^0$  közelítés  $y_h -$

$y_h^0 = w_h = A_h^{-1}r_h$  hibájának  $w_H = A_H^{-1}R_h^H r_h = A_H^{-1}R_h^H v_h$  közelítése egyszerűen nulla a durva rácson.

Tegyük fel, hogy  $A_h$  és  $A_H$  szimmetrikus pozitív definit mátrixok (amire számíthatunk, ha az eredeti elliptikus feladat önadjungált volt – különösen akkor, ha a végeelem módszerrel készítettük a (15.84) rendszert), ld. II. 11.6.1. és lejjebb 15.7.1. Akkor a tapasztalt jelenséget a mátrixok sajátvektoraival a következőképpen írhatjuk le, ezeket a vektorokat a hozzátartozó sajátértékek növekvő nagysága szerint rendezve :

Az  $A_H$  sajátvektorai többé-kevésbé közelítik az  $A_h$  első  $n(H)$  sajátvektorát (és mindkét rendszer közelíti az eredeti elliptikus differenciáloperátor első  $n(H)$  sajátvektorát). Az  $A_h$  utolsó  $n(h) - n(H)$  sajátvektorainak viszont nincsen megfeleltetése a durvább rácson (és  $n(h) - n(H) > n(H)$ ). Ezen magasabb sorszámú sajátvektorok számára nem nyerhetünk hasznos információt a durva rácson. Ugyanakkor – reményünk szerint –  $P_H^h w_H$  közel lesz  $y_H^i$  hibájához – ill. a fenti iteráció során – az  $y_h^i$  hibájához.

Ezért csak akkor lehet konvergens az eljárás, ha a magasabb sorszámú sajátvektorokkal kapcsolatos megoldási komponensek lényegében már  $y_h^i$ -ben jelen vannak és így  $y_h - y_h^i$ -ből hiányoznak, tehát ha  $y_h^i$  hibája *simá* abban az értelemben, hogy lényegében nincs benne magasabb sorszámú sajátvektor. Ezt a lejjebb tárgyalásra kerülő *simító iterációk*kal lehet elérni – míg az alacsonyabb sorszámú sajátvektoroknak megfelelő megoldási komponenseket a durva rácson megkapjuk kevés művelettel.

Így két egymást kiegészítő műveletet fogunk alkalmazni : a hiba közelítését a durva rácson ( $A_H w_H = R_h^H r_h$  alakú egyenletrendszerből) és a simító iterációkat a finom rácson. Ezzel a többrácsos módszernek az R.P. Fedorenko-ra visszamenő alapötletét máris elmondtuk.

### 15.4.2 A simító iterációk

Mint simító iterációk jó szolgálatot tesznek a klasszikus iterációs eljárások. Ezek bár lassan konvergálnak, de a megoldás hibáját gyorsan lesimítják az előző pont értelmében : a magasabb sorszámú sajátvektorokkal kapcsolatos megoldási komponenseket jóval gyorsabban pontosítják az iterációs vektorban, mint az alacsony sorszámúkat. Ezt illusztráljuk az egyszerű iteráció (v.ö. I. 1.6.6, más nevén a csillapított Jacobi-iteráció) példáján.

Legyen  $A_h \in \mathbb{R}^{n(h) \times n(h)}$  szimmetrikus és pozitív definit, sajátértékei :

$$0 < \lambda_{\min}^h \leq \lambda_k^h(A_h) \leq \lambda_{\max}^h, \quad k = 1, \dots, n(h),$$

és  $\{v_h^{(k)}\}_{k=1}^{n(h)}$  az ortonormált sajátvektor-bázis. Az iteráció

$$y_h^{(0)} \text{ adott, } y_h^{(m)} = (I_h - \omega A_h)y_h^{(m-1)} + \omega b_h, \quad m = 1, 2, \dots$$

alakú. Az  $y_h^{(m)}$  közelítések  $e_h^{(m)} := y_h^{(m)} - y_h$  hibájára érvényes

$$e_h^{(m)} = (I_h - \omega A_h)e_h^{(m-1)} = (I_h - \omega A_h)^m e_h^{(0)}.$$

Legyen

$$e_h^{(0)} = \sum_{k=1}^{n(h)} \alpha_k v_h^{(k)}$$

a hiba sorfejtése a sajátvektorok szerint (ahol  $\alpha_k v_h^{(k)}$  akkor egy, a fent említett magas sorszámú megoldási komponensek között, ha  $k$  összemérhető  $n(h)$ -val). Az  $m$ -edik közelítés hibája

$$e_h^{(m)} = \sum_{k=1}^{n(h)} \alpha_k (1 - \omega \lambda_k^h)^m v_h^{(k)},$$

és ennek euklideszi normája

$$\|e_h^{(m)}\|^2 = \sum_{k=1}^{n(h)} \alpha_k^2 (1 - \omega \lambda_k^h)^{2m}.$$

Ha

$$|1 - \omega \lambda_{\max}^h| \ll |1 - \omega \lambda_{\min}^h| \leq 1, \quad (15.87)$$

akkor a magasabb sorszámú sajátvektorokkal kapcsolatos komponensek gyorsan eltűnnek a megoldás  $e_h^{(m)}$  hibájából,  $m$  növekedésével. Ezek a sajátvektorok gyorsan oszcilláló rácsfüggvények, gondoljunk pl. az

$$u''(x) + f(x) = 0, \quad 0 < x < 1, \quad u(0) = u(1) = 0$$

peremérték feladat standard differencia-approximációjánál keletkező

$$A_h = \frac{1}{h^2} \text{tridiag}(-1, 2, -1)$$

mátrix (ld. II. 11.4.1. és II. 11.6.4.)  $v_h^{(k)} = \sqrt{2h}(\sin k\pi x_1, \dots, \sin k\pi x_{N-1})^T$  alakú sajátvektoraira (itt  $n(h) = N - 1$ ), ha  $k \gg 1$ .

Ha ismeretes  $\lambda_{\max}^h$ , akkor  $\omega$ -t közvetlenül megválaszthatjuk :

$$\omega = 1/\lambda_{\max}^h,$$

és ekkor  $v_h^{n(h)}$  komponense a hibából már az első iterációval eltűnik (tehát ezt a komponenst a megoldás közelítése tartalmazza már egzaktul), a többi magas sorszámú sajátvektor komponensei erősen lecsökkennek, a hiba „lesimul”.

Az előző  $\omega$ -érték helyett az általában könnyebben megszerezhető  $\|A_h\|_{C(\omega_h)}$ -val is dolgozhatunk,

$$\omega = 1/\|A_h\|_{C(\omega_h)},$$

hiszen  $0 < \lambda_{\min}^h \ll \lambda_{\max}^h \lesssim \|A\|_{C(\omega_h)}$  miatt ekkor (15.87) teljesül. (Az iterációs paraméternek ezen értékét I. 1.6.6-ban a „szegény ember iterációs paraméter”-ének neveztük. Itt kiderül, hogy ez az érték milyen hasznos.) A 15.14. lemmában térünk vissza arra, hogy az egész többrácsos módszer konvergenciájának a szempontjából is  $\omega$  megfelelő választásához elegendő az  $A_h$  maximális sajátértékét vagy felső becslését ismerni.

Végül is, ha az  $n(h)$ -pontú  $\omega_h$  rácsra és az  $n(H)$ -pontú  $\omega_H$  rácsra gondolunk, akkor az  $A_h$  mátrixnak csak a  $k \leq n(H)$  sorszámú sajátvektoroknak van megfelelője a durva rácson. Így célszerű az  $n_1 := n(H) + 1 \leq k \leq n(h)$  sorszámú sajátvektorok komponenseit erősen csillapítani a hibában azáltal, hogy az  $\omega$  iterációs paramétert a következő követelményből választjuk meg (v.ö. I. 1.6.6-gyel) :

$$1 > 1 - \omega \lambda_{n_1}^h = -(1 - \omega \lambda_{n(h)}^h) > -1. \quad (15.88)$$

Ez adja az

$$\omega = \frac{2}{\lambda_{n_1}^h + \lambda_{n(h)}^h}$$

simítás szempontjából optimális értéket. A tridiagonális mátrix és  $H = 2h$  esetén a sajátértékek explicit képletéből azonnal kiszámítható, hogy ekkor

$$\omega = \frac{h^2}{1 + 2 \cos^2(\pi h/2)} \gtrsim \frac{h^2}{3}, \quad (15.89)$$

és ez a választás azt eredményezi, hogy az  $n_1 \leq k \leq n(h)$  sorszámú sajátvektorok komponensei a hibában legalább  $|1 - \omega \lambda_k^h| \lesssim 1/3$ -ra csökkennek minden lépésben!

Az  $\omega$  paraméternek ilyen választása természetesen azt is eredményezi, hogy a hiba simább ( $k \leq n(H)$ ) része még lassabban konvergál, mint az  $\omega_{opt} = 2/(\lambda_1^h + \lambda_{n(h)}^h) = h^2/2$  optimális iterációs paraméter mellett. Ez utóbbi esetben a hibacsökkenés

$$\frac{\lambda_{n(h)}^h - \lambda_1^h}{\lambda_{n(h)}^h + \lambda_1^h} = 1 - \frac{h^2}{2} \lambda_1^h \approx 1 - \frac{\pi^2}{2} h^2$$

volt – tehát már ez is igen lassú. De a simító iteráció célja éppen a gyors simítás, és nem a gyors konvergencia. A kétdimenziós esetben hasonló lehet

elérni, ha a simítás kritériumát megfelelően választjuk meg, ld. a 16. feladatot is.

Most már ismerjük a többrácsos módszernek minden tartozékát, mert a simító iterációkkal kiegészített fenti iterációs algoritmus konvergál. Ennek bizonyítására később kitérünk. Először a módszer programozásával foglalkozunk.

### 15.4.3 Alapvető többrácsos algoritmusok

A többrácsos módszer legáttekinthetőbb programja rekurzív, a benne szereplő  $\gamma$  paraméterre (ld. lejjebb) teljesüljön  $1 \leq \gamma \leq 3$  kétdimenziós feladatok esetén. Ez a  $\gamma$  döntő szerepet játszik és megadja a leszállások számát a durvább rácsra. Érdekes csak a  $\gamma = 1$  eset (ilyenkor „V-ciklus”-ról beszélnek) vagy  $\gamma = 2$  („W-ciklus”) – amely elnevezések eredetére visszatérünk.

Az algoritmus megadásához a korábbi jelöléseinket csak annyiban változtatjuk meg, hogy a sok különböző rácsnak megfelelően nem  $A_{h_\ell} y_{h_\ell} = b_{h_\ell}$ , hanem  $A_\ell y_\ell = b_\ell$  alakban írjuk az  $\ell$ -edik rácson megoldandó feladatot,  $\ell = 1, \dots, \ell_{\max}$ . A legdurvább rács az  $\ell = 1$ -nek feleljen meg, a legfinomabb rács viszont  $\ell = \ell_{\max}$ -nak. „Rács” helyett most „szintet” is mondunk. Az  $\ell$ -edik szint pontszáma  $N_\ell := n(h_\ell)$  egyben az  $A_\ell$  mátrix dimenziója.

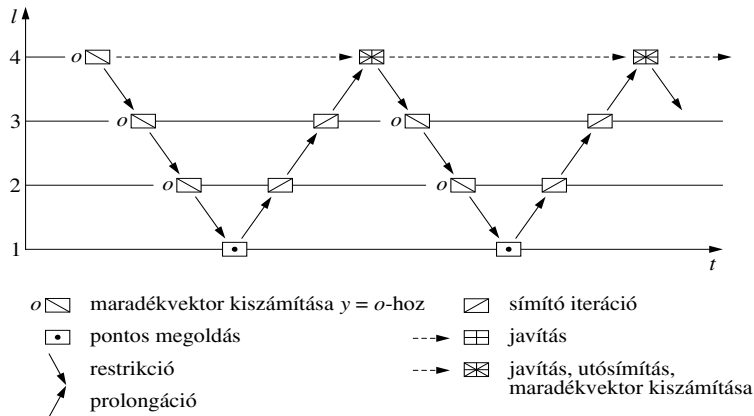
Az algoritmus célja tehát az  $A_{\ell_{\max}} y_{\ell_{\max}} = b_{\ell_{\max}}$  legmagasabbszintű egyenletrendszernek a megoldása.  $R_h^H$  ill.  $P_H^h$  helyett  $R_\ell^{\ell-1}$ , ill.  $R_\ell^{\ell-1}$  szerepel, de gyakran egyszerűen  $R$ -t ill.  $P$ -t írunk.

Először egy külön eljárást adunk meg, amelyet többféleképpen lehet hasznosítani.

```

procedure MG( $\ell$ , y, b);
begin
  if  $\ell=1$  then  $y:=A_1^{-1}*b$                                 direkt megoldás az első szinten
  else begin array r,w;
   $y:=S_h^{\nu_1}(A_\ell, b)y$ ;                                       elősimítás
   $r:=R(b-A_\ell y)$ ;                                           maradékvektor kiszámítása, restrikcója
   $\bar{\gamma}:=\gamma$ ; if  $\ell=2$  then  $\bar{\gamma}:=1$ ;
   $w:=0$ ; for  $i:=1$  step 1 until  $\bar{\gamma}$  do MG( $\ell-1, w, r$ );
   $y:=y+Pw$ ;                                                  prolongáció, javítás a finom rácson
   $y:=S_h^{\nu_2}(A_\ell, b)y$ ;                                       utósimítás
  end
end;

```



ábra 15.10: *Sima* többrácsos módszer, V-ciklus ( $\gamma = 1$ )

Ebben az eljárásban a direkt megoldást ritka mátrixra specializált LU-felbontással érhetjük el. Ennek többszörös hívását elkerülendő a  $\bar{\gamma}$  számot vezettük be  $\gamma$  helyett.

Az MG eljárást hasznosíthatjuk az  $A_{\ell_{\max}} y_{\ell_{\max}} = b_{\ell_{\max}}$  rendszer megoldására a következőképpen – ez az úgynevezett *sima többrácsos módszer*.

Adott  $\ell_{\max} \geq 1$ , az összes mátrix:  $\{A_\ell\}_{\ell=1}^{\ell_{\max}}$ , valamint a legmagasabb szint jobboldala:  $b_{\ell_{\max}}$ . Keresett az  $A_{\ell_{\max}} y_{\ell_{\max}} = b_{\ell_{\max}}$  rendszer közelítő megoldása.

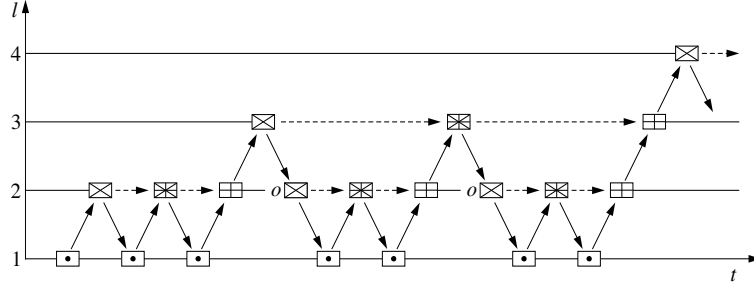
1.  $y_{\ell_{\max}} := 0$
2.  $i := 1(1)\mu \quad \text{MG}(\ell_{\max}, y_{\ell_{\max}}, b_{\ell_{\max}})$
3. stop [eredmény:  $y_{\ell_{\max}}$ ] Jóval hatékonyabb a következő *teljes többrácsos*

*módszer*, ugyanis csak itt érjük el a bevezetőben említett  $O(N_{\ell_{\max}})$  nagyságrendű műveletigényt :

Adott  $\ell_{\max} \geq 2$ , az összes mátrix és jobboldal :  $\{A_\ell\}_{\ell=1}^{\ell_{\max}}, \{b_\ell\}_{\ell=1}^{\ell_{\max}}$ .

1.  $\text{MG}(1, y_1, b_1)$
2.  $\ell := 2(1)\ell_{\max}$
3.  $[y_\ell := \tilde{P}_{\ell-1}^\ell y_{\ell-1}$
4.  $\mu := \mu_1, \quad ? \ell = \ell_{\max} ? [\mu := \mu_2]$
5.  $i := 1(1)\mu \quad \text{MG}(\ell, y_\ell, b_\ell) \quad ]_\ell$
6. stop [eredmény:  $y_{\ell_{\max}}$ ]

Itt  $\mu$  az MG eljárás hívásainak száma a mindenkori ( $\ell$ -edik) legmagasabb szinten. Gyakran elegendő  $\mu_1 = \mu_2 = 1$  – amire még visszatérünk. A  $\tilde{P}_{\ell-1}^\ell$  prolongációval még nem talákoztunk: nem a hiba approximációját viszi az új szintre, hanem a megoldását. Ezért érdemes,  $\tilde{P}_{\ell-1}^\ell$ -nek jobb interpolációt választani, mint  $P_{\ell-1}^\ell$ -nek. A következőkben megadjuk a *többrácsos módszer*



ábra 15.11: Teljes többrácsos módszer,  $W$ -ciklus ( $\gamma = 2$ )

nemrekurzív algoritmusát, amelyben  $\ell_{\max}$  továbbra is a maximális szintszám.

0. Rácsok, mátrixok, jobboldali vektorok kiszámítása;  $\ell := \ell_{\max}$
1. nulladik közelítés az  $\ell$ -edik szinten (pl.  $y_\ell := 0$ )
2. iterációs simítás:  $y_\ell := S_\ell^{\nu_1} y_\ell$
3. maradékvektor kiszámítása:  $r_\ell := b_\ell - A_\ell y_\ell$
4. leereszkedés az alacsonyabb szintre:  $b_{\ell-1} := R_\ell^{\ell-1} r_\ell$ ,  $\ell := \ell - 1$
5. ha  $\ell > 1$ :  $\rightarrow 1$ .  
ha  $\ell = 1$ :  $y_\ell := A_\ell^{-1} b_\ell$  (pl. ritkamátrixú LU-felbontással)
6. emelkedés a magasabb szintre:  $w_{\ell+1} := P_\ell^{\ell+1} y_\ell$ ,  $\ell := \ell + 1$
7. a megoldás javítása:  $y_\ell := y_\ell + w_\ell$
8. iterációs simítás:  $y_\ell := S_\ell^{\nu_2} y_\ell$
9. ha még nem végeztük el a  $\gamma$  iterációt ezen a szinten:  $\rightarrow 2$ .
10. ha ez a legmagasabb szint ( $\ell = \ell_{\max}$ ): stop [eredmény :  $y_\ell$ ]  
ha nem:  $\rightarrow 6$ .

Az itt leírt formában ez a sima többrácsos módszernek felel meg. A teljes többrácsos módszer algoritmusát viszont könnyen készíthető ebből (ld. a 17. feladatot).

Egyebek között ilyen típusú program is található W. Hackbusch "többrácsos" könyvében (a (15.13) feladat megoldására).

Következőnek a legelterjedtebb iterációkat, restriktciókat és interpolációkat tekintjük át (• ld. az „Ellipsz” program segédszövegeiben az ábrázolásokat is!).

### 1. simító iterációk :

- csillapított Jacobi (egyszerű iteráció) : elméleti vizsgálatra a legkényelmesebb; gyakorlatilag kevésbé javasolható,
- Gauss–Seidel (ez a standard simító iteráció; részleteket ld. lent),
- szimmetrikus Gauss–Seidel;



- SSOR (szimmetrikus felsőrelaxáció – váltakozó háromszög mátrixok módszere, Szamarszkij-féle paraméter választással, ld. Szamarszkij és Nyikolajev könyvét);

- konjugált gradiens iteráció;
- Csebisev-iteráció.

A Gauss–Seidel iteráció hatékonysága mint simító iteráció növelhető, ha különleges sorrendben vesszük az ismeretleneket, amikor téglalap alakú rácsra helyezkednek el (• tehát nemcsak akkor, amikor – mint az Ellipsz programban – a tartomány téglalap) : nem sorfolytonosan, hanem sakktable-szerű elrendezésben (ld. a 15.5. ábrát, pl. először a „fekete” majd a „fehér” pontokat iterálva), vagy négyzín elrendezésben. Továbbá, kedvelt a *zebra iteráció* is. Ennél a rács egyik irányában az egyenleteket úgy oldjuk meg pontosan, hogy a szomszéd rácsegyeneseken fekvő korábbi közelítő megoldási értékeket felhasználjuk Gauss–Seidel módra. Ez az iteráció attól „zebra”, hogy először a páros, majd a páratlan sorszámú rácsegyeneseket vesszük.

Ezek a „színes” Gauss–Seidel iterációk a háromdimenziós elliptikus standard feladatok esetén is jól működnek. Konvekció-diffúzió feladatokban a fekete-fehér iteráció viszont nem javasolható (már két dimenzióban sem). A többrácsos módszerben hasznos iterációkról a 16.7.2. és 17.5.5. \*\*\* pontokban is lesz szó.

Egy kézenfekvő programozási hiba elkerülésének az érdekében (a többrácsos algoritmus tesztelésénél hasznos sok tanácsot adnak egyébként Briggs, Henson és McCormick könyvében) hangsúlyozzuk, hogy a simító iterációkat (és a restrikiókat, interpolációkat) *másod- és harmadfajú, valamint periodikus peremfeltételek* esetén a peremen is kell alkalmazni! Az ilyen peremen fekvő pontok is számítanak belsőeknek.

Az *elsőfajú peremfeltétellel* viszont csak a legfelsőbb szinten kell foglalkoznunk : a  $\tilde{P}_{\ell_{max}-1}^{\ell_{max}} y_{\ell_{max}-1}$  interpoláció előtt betöltjük a finom rácsnak megfelelő peremértékeket, majd az interpoláció erre is támaszkodik. Az alacsonyabb szinteken mindig homogének az elsőfajú peremfeltételek, hiszen ott a hibaegyenletekről van szó, és még a  $P_{\ell_{max}-1}^{\ell_{max}} y_{\ell_{max}-1}$  elsőfajú peremfeltételei is homogének.

## 2. restrikiók :

- direkt átvitel (függvényérték szorozva  $\sigma$ -val; leginkább  $\sigma = 1, 0.5$ . Konvekció-diffúzió feladatoknál ennél kisebb, a 17.4.0-ban \*\*\* (42)-vel definiált  $q$  lokális cella-Reynolds-számtól függő  $\sigma$  is lehet célszerű),

- 5-pont restrikió;
- 7-pont (lináris háromszöges végeelem restrikió);
- 9-pont restrikió.

Most érdemes ellenőrizni, hogy a restriktciók magtere valóban a gyorsan oszcilláló rácsfüggvényeknek felel meg! Ehhez ld. a 18. feladatot.

### 3. interpolációk :

- bilineáris interpoláció;
- 7-pont (lineáris háromszögös végeelem interpoláció);
- harmadrendű interpoláció (perem közelében másodrendű).

További lehetőségeket ld. pl. Mohr és Wienands cikkében.

**Megjegyzés.** A felsorolt restriktciókra és interpolációkra nézve van bizonyos választási szabadság, de őket nem ajánlatos tetszőlegesen kombinálni. Egyrészt, van egy korlát ezen műveletek rendjére nézve, ld. 15.4.7. pont, másrészt, ha  $A_\ell$  szimmetrikus a  $D_\ell$  szimmetrikus és pozitív definit mátrix által definiált skalárszorzatra nézve :

$$(D_\ell A_\ell u_\ell, v_\ell)_\ell = (D_\ell u_\ell, A_\ell v_\ell)_\ell \text{ minden } u_\ell, v_\ell \in \mathbb{R}^{N_\ell}\text{-re,}$$

(ahol  $(\cdot, \cdot)_\ell$  az  $\ell$ -edik szint euklideszi normája) azaz ha  $D_\ell A_\ell$  szimmetrikus, akkor az – itt tárgyalt **véges differencia** alapú – többrácsos módszer hatékonysága nagyobb, ha teljesül

$$P_{\ell-1}^\ell = D_{\ell-1}^{-1} (R_{\ell-1}^{\ell-1})^T D_{\ell-1} \text{ avagy } R_{\ell-1}^{\ell-1} = D_{\ell-1}^{-1} (P_{\ell-1}^\ell)^T D_{\ell-1}. \quad (15.90)$$

Erre a kapcsolatra a 15.4.5, 15.4.6 és 15.8.1. pontokban térünk vissza. Az utóbbi pontban kiderül, hogy a **“beágyazott” végeelem** alapú többrácsos módszer esetén nincs választási szabadság, mind az  $A_\ell$ ,  $P_{\ell-1}^\ell$  és  $R_{\ell-1}^{\ell-1}$  mátrixokat a bázis határozza meg és érvényes

$$P_{\ell-1}^\ell = (R_{\ell-1}^{\ell-1})^T.$$

Ez viszont egyben egy példa (15.90) teljesülésére : ha a peremérték feladat elliptikus operátora önadjungált, akkor számíthatunk arra, hogy maga  $A_\ell$  szimmetrikus, azaz  $D_\ell = I_\ell$ .  $\square$

• Ehhez a ponthoz próbáljuk ki az „Ellipsz” programot és oldjuk meg a 19., valamint a 27. és 28. feladatot! Egy olyan algoritmus kidolgozása, amely (egyszerű, problemamentes elliptikus feladatoknál) a műveletigény szempontjából különösen előnyös, a 20. feladat tárgya.

#### 15.4.4 A többrácsos iterációk művelet- és tárigénye

Az alábbi elemzés adja a magyarázatot arra a fenti megállapításunkra, hogy a  $\gamma$  paramétertől érzékenyen függ a többrácsos módszerek műveletigénye.

Így a kétdimenziós peremérték feladatok diszkretizációjából származó lineáris rendszereknél  $\gamma = 3$  még lehetséges,  $\gamma = 4$  már tilos abban az értelemben, hogy ekkor már nem igaz, hogy a módszer műveletigénye az  $\ell$ -edik szinten  $Q_\ell = O(N_\ell)$ .

Tekintsük a sima módszernek egy iterációját ( $\mu = 1$ ) az  $\ell$ -edik szinten. Legyen  $\bar{Q}_\ell$  a következő lépések műveletigénye ( $\ell \geq 2$ ):

- a Gauss–Seidel-féle elősimítás az  $\ell$ -edik szinten,
- a maradékvektor kiszámítása,
- a bilineáris restrikció kiszámítása,
- a bilineáris interpoláció és a javítás kiszámítása, (az  $\ell - 1$ -edik szintről való visszatérés után),
- az utósimítás.

Ekkor

$$\bar{Q}_\ell \leq (2(s+1)\nu_1 + 2(s+1) + \frac{11}{4} + 104 + 2(s+1)\nu_2)N_\ell = O(N_\ell). \quad (15.91)$$

Itt egy művelet alatt bármely aritmetikai műveletet értünk, ezenkívül feltettük, hogy minden belső rácspontnak legfeljebb  $s$  szomszédja van (azaz soronként legfeljebb  $s + 1$  nemzérus eleme van  $A_\ell$ -nak). Továbbá azt is feltételeztük, hogy  $\nu_1$  és  $\nu_2$  nem nagyok, aminek alapozását 15.4.7-ben fogjuk látni (gyakran vesznek  $\nu_1 + \nu_2 \leq 2$ ). A Poisson-egyenlet ötpontos approximációjának az esetén pl.  $s = 4$ . Ha ekkor  $\nu_1 + \nu_2 = 2$ , akkor  $\bar{Q}_k < 36N_k$  adódik.

Ezután konkrétan feltehetjük, hogy  $\bar{Q}_k \leq c_0N_k$ ,  $2 \leq k \leq \ell$ . Ekkor

$$Q_k = \bar{Q}_k + \gamma Q_{k-1}, \quad k \geq 2,$$

lesz a (sima) többrácsos módszer egy iterációjának műveletigénye a  $k$ -edik szinten, ld. az eljárást. Ezt az eredményt azzal egészíthetjük ki, hogy  $Q_0 := 0$ , és  $\bar{Q}_1$  az első szinthez tartozó direkt megoldásnak a költsége.

A következőkben az egyszerűség kedvéért  $4N_{k-1} = N_k$ -val számítunk,  $1 \leq k \leq \ell$ , azaz a rács *finomítása felezéssel* történik (ehhez ld. a 21. feladatot). Ekkor  $Q_\ell$ -re azt kapjuk, hogy

$$\begin{aligned} Q_\ell &= \sum_{k=0}^{\ell-1} \gamma^k \bar{Q}_{\ell-k} \leq \gamma^{\ell-1} \bar{Q}_1 + c_0 \sum_{k=0}^{\ell-2} N_{\ell-k} \gamma^k \\ &\leq \max(c_1, c_0) \sum_{k=0}^{\ell-1} N_{\ell-k} \gamma^k \leq \max(c_1, c_0) \frac{N_\ell}{1 - \gamma/4} \leq 4 \max(c_1, c_0) N_\ell, \end{aligned}$$

amíg  $\bar{Q}_1 \leq c_1 N_1$  és  $1 \leq \gamma \leq 3$ .

**Megjegyzés.** A  $\bar{Q}_1 \leq c_1 N_1$  egyenlőtlenség érvényessége néha problémát jelent. Ugyanis  $c_1 := \bar{Q}_1/N_1$  nagy is lehet (és ezzel az itt vázolt eljárás nehézkes), ha a legalacsonyabb, első szinten a peremérték feladat sajátosságai (pl. a peremvonal erős tagoltsága) miatt már viszonylag sok pontra van szükség. Ekkor  $N_1 \gg 1$  és  $\bar{Q}_1/N_1 \lesssim \frac{1}{3}N_1^2$ , a legrosszabb esetben teltmátrixú Gauss-eliminációra gondolva. Ugyanemiatt (és különösen akkor, ha még a memória is eléggé kicsi) várható, hogy csak néhányszor finomítható a rács – tehát  $(\frac{\gamma}{4})^{\ell-1}$  nem kicsi.

Ilyen esetben nehéz a többrácsos módszer alkalmazása, ld. a 15.4.10. pontot. Ez a probléma viszont a  $Q_\ell \leq \text{const } N_\ell$  eredményünket nem érinti.  $\square$

A fenti költségekhez hozzájön még  $\{A_k\}_{k=1}^\ell$  és  $b_\ell$  kiszámításának műveletigénye, ami becsülhető mint  $c_2 \sum_{k=1}^\ell N_k < 4/3 c_2 N_\ell$ , tekintve, hogy  $A_k$  összes nemzérus elemeinek száma  $(s+1)N_k$ .

Ha a sima többrácsos módszer egész költségét akarjuk becsülni, akkor előbb még azt kellene tisztáznunk, hány iterációra van szükség egy adott pontosság eléréséhez. Erre visszatérünk a konvergencia tétel után, ld. a 15.17. tételhez fűzött 2. megjegyzést.

Most a teljes többrácsos módszer  $\tilde{Q}_\ell$  költségének vizsgálatára térünk rá, ebbe beleszámítva minden aritmetikai műveletet addig, míg  $\ell$  az ideiglenesen legmagasabb szint és az algoritmus be nem fejezte ott a  $\mu_1 = \mu_2 =: \mu$  iterációt.  $\tilde{Q}_\ell$ -re az előbbiek alapján (ld. az eljárást is) a következő becslést kapjuk :

$$\begin{aligned} \tilde{Q}_\ell &= \bar{Q}_1 + \sum_{k=2}^\ell (\mu Q_k + c_3 N_k) \leq \bar{Q}_1 + \sum_{k=1}^\ell (c\mu + c_3) N_k \\ &\leq c_1 N_1 + (c\mu + c_3) \frac{4}{3} N_\ell \leq \text{const } N_\ell. \end{aligned}$$

Itt a  $\tilde{P}_k^{k-1}$  interpoláció költségét  $c_3 N_k$ -val számítottuk. Újabb feltételre nincs szükség és  $\tilde{Q}_\ell = O(N_\ell)$  biztosított  $1 \leq \gamma \leq 3$  esetén – kivéve azt, hogy  $\mu$  korlátosnak választható legyen. Ez majd következik a konvergencia tételből, ld. a 15.18. tétel 2. megjegyzését.

Következőnek vizsgáljuk a tárigényt olyan kétdimenziós feladatra, ahol megfelelő egy ötponos approximáció. Ilyenkor a következőt kell eltárolni, ha egyszerű a tartomány (pl. téglalap) :

$$\{A_k, b_k, y_k\}_{k=1}^\ell \rightarrow (5 + 1 + 1) \sum_{k=1}^\ell N_k < \frac{7 \cdot 4}{3} N_\ell \quad \text{tárhely,}$$

ugyanis  $Rr_k$  elfoglalja  $b_{k-1}$  helyét (és  $Rr_k$  kiszámítható  $b_k$ ,  $A_k$ , valamint  $y_k$  ismeretében úgy, hogy maga  $r_k$  nem is készül egész vektorként, csak az  $Rr_k$  kiszámításához éppen szükséges komponensei),  $w_{k-1} = y_{k-1}$ , stb. Kényelmesebb persze  $r_k$  külön kiszámítása és tárolása. Ekkor az előbbi  $\frac{28}{3}N_\ell$  helyett  $\frac{32}{3}N_\ell$  tárhely az eredmény. Valójában viszont az  $r_k$  tárolására szolgáló ( $v$ -vel jelölt) segéd tömböt elegendő csak a legmagasabb szinten bevezetni, hiszen ugyanazt használhatjuk az alacsonyabb szinteken is.

Ha bonyolult a tartomány, akkor  $A_k$  mellett még a kompakt tárolását biztosító **ia**, **ja** mutatóvektorokat is el kell tárolni (ld. I. 1.3.9., ezek egyben a szomszédsági viszonyokat is adják, v.ö. 15.7.7-tel). Ezenkívül  $A_k$  kiszámításához és az eredmények kirajzolásához az  $(x_1, x_2)$  koordináták előnyösek. Ha most pl. a valós számok nyolcbájtosak, **ia** komponensei négybájtosak, **ja** komponensei kétbájtosak, akkor

$$v, \{\mathbf{ia}, \mathbf{ja}, A_k, b_k, y_k, x_1^k, x_2^k\}_{k=1}^\ell \rightarrow N_\ell + \left(\frac{1}{2} + \frac{5}{4} + 5 + 4\right) \sum_{k=1}^\ell N_k < 16N_\ell$$

nyolcbájtos tárhely elegendő.

Ehhez a ponthoz ld. a 22. feladatot is.

### 15.4.5 A kétrácsos módszer konvergenciája

A többrácsos módszer konvergencia bizonyítását megkönnyíti, ha először a kétrácsos módszert vizsgáljuk. Ennek egy iterációja az  $A_h y_h = b_h$  rendszer megoldására felírható a következőképpen ( $\nu_1 = \nu$ ,  $\nu_2 = 0$ . Legyen  $T_h$  az iteráció folyamán a jobboldali vektorra ható mátrix) :

$$\begin{aligned} \bar{y}_h^{(0)} &:= y_h^0, \quad \bar{y}_h^{(m)} = S_h \bar{y}_h^{(m-1)} + T_h b_h, \quad m = 1, \dots, \nu, \\ r_h &:= b_h - A_h \bar{y}_h^{(\nu)}, \\ A_H w_H &= R_h^H r_h, \quad (\text{pontos megoldás}) \\ y_h^1 &:= \bar{y}_h^{(\nu)} + P_H^h w_H. \end{aligned}$$

Már I. 1.6.2-ből tudjuk, hogy az iterációs eljárások konvergencia vizsgálatához lényeges az  $e_h^{(m)} := y_h^{(m)} - y_h$  hiba egyenletének az előállítása

$$e_h^{(m)} = K_h e_h^{(m-1)}$$

alakjában, valamint a  $K_h$  iterációs mátrix spektrálsugarának (ill. normájának) a becslése. Mivel a pontos megoldás a simító iteráció fixpontja kell, hogy legyen, érvényes

$$\bar{e}_h^{(m)} = S_h \bar{y}_h^{(m-1)} + T_h b_h - y_h = S_h \bar{e}_h^{(m-1)} = \dots = S_h^m \bar{e}_h^{(0)},$$

ahol  $\bar{e}_h^{(m)} := \bar{y}_h^{(m)} - y_h$ . Ezzel a  $b_h$ -val kapcsolatos tagok eltűntek (ld. a 22\*. feladatot \*\*\* is). Most

$$\begin{aligned} r_h &= b_h - A_h \bar{y}_h^{(\nu)} = A_h (y_h - \bar{y}_h^{(\nu)}) = -A_h \bar{e}_h^{(\nu)} = -A_h S_h^\nu \bar{e}_h^{(0)} = -A_h S_h^\nu e_h^0, \\ w_H &= A_H^{-1} R_h^H r_h = -A_H^{-1} R_h^H A_h S_h^\nu e_h^{(0)}, \\ e_h^1 &= \bar{e}_h^{(\nu)} - P_H^h A_H^{-1} R_h^H A_h S_h^\nu e_h^0 = (I_h - P_H^h A_H^{-1} R_h^H A_h) S_h^\nu e_h^0 \\ &= (A_h^{-1} - P_H^h A_H^{-1} R_h^H) A_h S_h^\nu e_h^0. \end{aligned}$$

Tehát az iterációs mátrix

$$K_h = K_h(\nu) = (A_h^{-1} - P A_H^{-1} R) (A_h S_h^\nu) =: C_h A_h S_h^\nu.$$

$K_h$  alakjának meghatározása után rögtön adódik a következtetés, hogy a pontos megoldás a kétrácsos (és hasonlóképpen a többrácsos) módszer fixpontja is :  $e_h^0 = 0 \implies e_h^1 = 0$ .

**Megjegyzés.** Ha  $A_h$  szimmetrikus a  $D_h$  szimmetrikus és pozitív definit mátrix által definiált skalárszorzatra nézve, akkor  $C_h = A_h^{-1} - P A_H^{-1} R$  is ilyen értelemben szimmetrikus, amennyiben igaz (15.90). Hogy milyen értelemben lesz ekkor  $K_h$  szimmetrikus, azt majd 15.8.3-ban vizsgáljuk.  $\square$

$K_h$  szorzatalakjának megfelelően a konvergencia vizsgálatot két részre bontjuk, valamilyen indukált mátrixnormát választva :

a) megmutatjuk, hogy az *approximációs tulajdonság* megvan :

$$\|C_h\| = \|A_h^{-1} - P A_H^{-1} R\| \leq c_a; \quad (15.92)$$

b) megmutatjuk, hogy a *simítási tulajdonság* megvan :

$$\|A_h S_h^\nu\| \leq c_s, \quad \text{ahol } c_s = c_s(\nu) \rightarrow 0, \quad \nu \rightarrow \infty. \quad (15.93)$$

Ezután következik

$$\|K_h(\nu)\| \leq c_a c_s(\nu) < 1$$

elég nagy  $\nu$ -re. Világos, hogy (15.92), (15.93) helyett az is megfelelő, ha érvényes

$$\|C_h\| = \|A_h^{-1} - P A_H^{-1} R\| \leq c_a h^\alpha, \quad (15.94)$$

és

$$\|A_h S_h^\nu\| \leq c_s h^{-\alpha}, \quad \text{ahol } c_s = c_s(\nu) \rightarrow 0, \quad \nu \rightarrow \infty. \quad (15.95)$$

Ebben a formában (bizonyos  $\alpha \geq 0$ -val) gyakran adódik a simítási és approximációs tulajdonság, a normától függően néha  $\alpha = 0$ -val, néha  $\alpha = 2$ -vel stb. Lényeges csak az, hogy (15.94) és (15.95) ugyanazt az  $\alpha$  számot tartalmazzák, egyszer plusz, egyszer mínusz előjellel.

### 15.4.6 Az egydimenziós eset

A többrácsos módszer alkalmazása az egydimenziós esetre nemlineáris feladatoknál lehet indokolt. Lineáris egydimenziós feladatok esetén viszont nincs szükség a többrácsos módszerre (mert ekkor a sávós Gauss-elimináció műveletigénye  $O(N)$ ). De egyszerűbb és tanulságos ezen eset elemzése. Ezért vizsgáljuk először részletesen az idevonatkozó legegyszerűbb modelfeladatot (v.ö. II. 11.4.2-vel) :

$$u'' + f = 0, \quad 0 < x < 1; \quad u(0) = u_1, \quad u(1) = u_2,$$

és annak megoldását a szokásos sémával :

$$y_{\bar{x}x} + f = 0, \quad x \in \omega_h; \quad y_0 = u_1, \quad y_n = u_2, \quad h = 1/n.$$

Itt  $n = 2N$ . Ugyanezt a sémát az  $\omega_H$  rácson is vizsgáljuk, ahol  $H = 2h = 1/N$ . Ekkor

$$\begin{aligned} A_h &= \frac{1}{h^2} \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{(n-1) \times (n-1)}, \\ A_H &= \frac{1}{H^2} \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{(N-1) \times (N-1)}. \end{aligned}$$

Mint simító iterációt választjuk az egyszerű iterációt  $\omega$  paraméterrel :

$$\begin{aligned} \bar{y}_h^{(m+1)} &= (I_h - \omega A_h) \bar{y}_h^{(m)} + \omega b_h, \\ b_h &:= \left( f_1 + \frac{u_1}{h^2}, f_2, \dots, f_{n-2}, f_{n-1} + \frac{u_2}{h^2} \right)^T. \end{aligned}$$

A restriktio legyen három szomszédpont súlyozott átlaga :

$$R = R_H^h := \frac{1}{4} \begin{pmatrix} 1 & 2 & 1 & & & & & & 0 \\ & & 1 & 2 & 1 & & & & \\ & & & & \ddots & & & & \\ & & & & & & 1 & 2 & 1 \\ 0 & & & & & & & & 1 & 2 & 1 \end{pmatrix} \in \mathbb{R}^{(N-1) \times (n-1)},$$

az interpoláció a szakaszonkénti lineáris interpoláció :

$$P = P_h^H := \frac{1}{2} \begin{pmatrix} 1 & & & & 0 \\ 2 & & & & \\ 1 & 1 & & & \\ & & 2 & \ddots & \\ & & 1 & \ddots & \\ & & & & 1 \\ 0 & & & & 2 \\ & & & & 1 \end{pmatrix} \in \mathbb{R}^{(n-1) \times (N-1)}.$$

Nyilvánvalóan ezeket a mátrixokat nem fogjuk tárolni, a velük való szorzást két-két programsor intézi el. Egyébként azt látjuk, hogy  $P = 2R^T$ ,  $R = \frac{1}{2}P^T$ . Ez egy példa (15.90) teljesülésére : Az  $A_h$  mátrix szimmetrikus az egydimenziós  $(0, h)$ -skalárszoratra nézve (ld. II 11.4.6., (88)), vagyis itt  $D_h = hI_h$ . Ekkor  $D_H = HI_H = 2hI_H$  és  $D_h^{-1}R^TD_H = 2R^T = P$ .

Számítsuk ki most az  $RA_hP$  szorzatot! Először kapjuk, hogy

$$A_hP = \frac{1}{2h^2} \begin{pmatrix} 0 & \dots & \dots & \dots & \dots & 0 \\ 2 & -1 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 \\ & & & \ddots & \ddots & \ddots \\ 0 & & & & -1 & 2 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix} \in \mathbb{R}^{(n-1) \times (N-1)},$$

majd

$$RA_hP = \frac{1}{8h^2} \begin{pmatrix} 4 & -2 & 0 & \dots & \dots & 0 \\ -2 & 4 & -2 & 0 & \dots & 0 \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & -2 \\ 0 & & & & -2 & 4 \end{pmatrix} = A_H. \quad (15.96)$$

Ebből következik, hogy  $C_hA_h$  és  $Q_h := A_h^{1/2}C_hA_h^{1/2}$  (ahol  $C_h := (A_h^{-1} - PA_H^{-1}R)$ ) projektorok. Vegyük pl. a  $Q_h = I_h - A_h^{1/2}PA_H^{-1}RA_h^{1/2}$  mátrixot :

$$Q_h^2 = I_h - 2A_h^{1/2}PA_H^{-1}RA_h^{1/2} + A_h^{1/2}PA_H^{-1}(RA_hP)A_H^{-1}RA_h^{1/2} \quad (15.97)$$

$$= I_h - A_h^{1/2}PA_H^{-1}RA_h^{1/2} = Q_h. \quad (15.98)$$

Ezenkívül  $Q_h$  szimmetrikus és következésképpen pozitív szemidefinit is :

$$(Q_hx, x) = (Q_h^2x, x) = \|Q_hx\|^2 \geq 0.$$

Mivel  $Q_h \neq 0$ , így normája 1. Az a tény, hogy  $Q_h$  projektor, várhatóan előnyös a konvergencia vizsgálatnál (ld. 15.8.2-ben a 15.40. lemmát). De ha csak ezt használjuk, akkor a

$$\|K_h\|_{A_h} = \|A_h^{1/2}K_hA_h^{-1/2}\| = \|A_h^{1/2}C_hA_hS_h^\nu A_h^{-1/2}\| = \|Q_hS_h^\nu\| \leq \|S_h^\nu\|$$

összefüggést kapjuk, amely nem ad valós képet  $K_h$  hatásáról. Hogy ezt belásuk, számítsuk ki a  $K_hv_h^{(j)}$  szorzatokat az  $A_h$  mátrix  $v_h^{(j)}$  sajátvektoraira



( $j = 1, \dots, n-1$ ), tudva, hogy  $\lambda_j^h = \left(\frac{2}{h} \sin \frac{j\pi h}{2}\right)^2$  és  $(v_h^{(j)})_i := \sin ji\pi h$ . Ezek a vektorok így nem normalizáltak, de ez egyelőre nem lényeges. Ezen vektorok mellett az  $A_H$  mátrix  $v_H^{(j)}$  sajátvektorai és  $\lambda_j^H$  sajátértékei is előfordulnak. Érvényesek a következő összefüggések (vegyük figyelembe, hogy mindig  $n = 2N$ ) :

$$(v_h^{(j)})_{2i} = \begin{cases} (v_H^{(j)})_i, & \text{ha } 1 \leq j \leq N-1, \\ 0, & \text{ha } j = N, \\ -(v_H^{(n-j)})_i, & \text{ha } N+1 \leq j \leq 2N-1, \end{cases} \quad i = 1, \dots, \frac{n}{2},$$

$$\lambda_j^H = \lambda_j^h \cos^2 \frac{j\pi h}{2}, \quad 1 \leq j \leq N-1,$$

$$\lambda_{n-j}^H = \lambda_j^h \cos^2 \frac{j\pi h}{2}, \quad N+1 \leq j \leq 2N-1,$$

$$(v_h^{(j)})_{i-1} + (v_h^{(j)})_{i+1} = 2 \cos j\pi h (v_h^{(j)})_i,$$

$$(v_H^{(j)})_i + (v_H^{(j)})_{i-1} = 2 \cos j\pi h (v_h^{(j)})_{2i-1}.$$

Ezekből a relációkból következik

$$\begin{aligned} (Rv_h^{(j)})_i &= \frac{1}{4} \left[ (v_h^{(j)})_{2i-1} + 2(v_h^{(j)})_{2i} + (v_h^{(j)})_{2i+1} \right] = \frac{\cos j\pi h + 1}{2} (v_h^{(j)})_{2i} \\ &= \cos^2 \frac{j\pi h}{2} (v_h^{(j)})_{2i} = \cos^2 \frac{j\pi h}{2} \begin{cases} (v_H^{(j)})_i, & \text{ha } 1 \leq j \leq N-1, \\ 0, & \text{ha } j = N, \\ -(v_H^{(n-j)})_i, & \text{ha } N+1 \leq j \leq 2N-1. \end{cases} \end{aligned}$$

Ezért

$$A_H^{-1} Rv_h^{(j)} = \cos^2 \frac{j\pi h}{2} \begin{cases} (\lambda_j^H)^{-1} v_H^{(j)}, & \text{ha } 1 \leq j \leq N-1, \\ 0, & \text{ha } j = N, \\ -(\lambda_{n-j}^H)^{-1} v_H^{(n-j)}, & \text{ha } N+1 \leq j \leq 2N-1, \end{cases}$$

és

$$(PA_H^{-1} Rv_h^{(j)})_{2i} = \begin{cases} (\lambda_j^h)^{-1} (v_h^{(j)})_{2i}, & \text{ha } 1 \leq j \leq N-1 \\ 0, & \text{ha } j = N \\ (\lambda_j^h)^{-1} (v_h^{(j)})_{2i}, & \text{ha } N+1 \leq j \leq 2N-1 \end{cases} = (\lambda_j^h)^{-1} (v_h^{(j)})_{2i}.$$

Itt azt is felhasználtuk, hogy  $(v_h^{(N)})_{2i} = \sin 2iNh\pi = \sin i\pi = 0$ . Továbbá

következik

$$\begin{aligned} \left( PA_H^{-1} Rv_h^{(j)} \right)_{2i-1} &= \frac{1}{2} \left[ \left( A_H^{-1} Rv_h^{(j)} \right)_i + \left( A_H^{-1} Rv_h^{(j)} \right)_{i-1} \right] \\ &= \cos^2 \frac{j\pi h}{2} \cdot \begin{cases} (\lambda_j^H)^{-1}, & 1 \leq j \leq N-1 \\ 0, & j = N \\ (\lambda_{n-j}^H)^{-1}, & N+1 \leq j \leq 2N-1 \end{cases} \frac{(v_h^{(j)})_{2i} + (v_h^{(j)})_{2i-2}}{2} \\ &= (\lambda_j^h)^{-1} \cos j\pi h (v_h^{(j)})_{2i-1}. \end{aligned}$$

Összesítve :

$$\left( C_h A_h v_h^{(j)} \right)_k = \left( (I_h - PA_H^{-1} RA_h) v_h^{(j)} \right)_k = (v_h^{(j)})_k \cdot \begin{cases} 0, & k = 2i, \\ 1 - \cos j\pi h, & k = 2i-1. \end{cases} \quad (15.99)$$

Ahogy látjuk, kicsi  $j$ -re a  $C_h A_h$  erőteljesen csillapítja a  $v_h^{(j)}$  vektort, de pl.

$$(C_h A_h v_h^{(N)})_k = (v_h^{(N)})_k, \text{ ha } k = 2i - 1.$$

Most figyelembe vesszük az elősimító iterációnak egy lépését is. A következőkben tehát a  $K_h(1)$  mátrix normáját fogjuk becsülni. Abból indulunk ki, hogy mivel

$$1 - \cos j\pi h = 2 \sin^2 j\pi h/2$$

és

$$(v_h^{(n-j)})_k = \begin{cases} -(v_h^{(j)})_k, & \text{ha } k = 2i, \\ (v_h^{(j)})_k, & \text{ha } k = 2i - 1, \end{cases}$$

a fenti (15.99) eredményt így is lehet írni :

$$C_h A_h v_h^{(j)} = s_j \cdot \left( v_h^{(j)} + v_h^{(n-j)} \right), \quad s_j := \sin^2 \frac{j\pi h}{2} = \frac{h^2}{4} \lambda_j^h, \quad (15.100)$$

és a  $j \leftrightarrow n - j$  cserével következik (felhasználva a  $s_{n-j} = \cos^2 \frac{j\pi h}{2} =: c_j$  relációt)

$$C_h A_h v_h^{(n-j)} = c_j \cdot \left( v_h^{(j)} + v_h^{(n-j)} \right).$$

Ez azt jelenti, hogy a  $\text{span}\{v_h^{(j)}, v_h^{(n-j)}\}$  altér invariáns  $C_h A_h$ -ra (valamint  $K_h = K_h(1) = C_h A_h S_h$ -ra) nézve. Emiatt vezessük be a  $j' = n - j$  jelölést és vizsgáljuk az

$$\alpha v_h^{(j)} + \beta v_h^{(j')} \quad (1 \leq j \leq N)$$

vektor sorsát. Az egyszerű iteráció egy lépése alatt ebből

$$(1 - \omega \lambda_j^h) \alpha v_h^{(j)} + (1 - \omega \lambda_{j'}^h) \beta v_h^{(j')}$$

lesz és a durva rácsú pontos megoldás után kapjuk

$$[(1 - \omega\lambda_j^h)\alpha s_j + (1 - \omega\lambda_{j'}^h)\beta c_j] (v_h^{(j)} + v_h^{(j')}).$$

Most használjuk (15.100) második relációját, valamint a következőt :

$$c_j = s_{j'} = \frac{h^2}{4} \lambda_{j'}^h.$$

Ennek alapján  $K_h(1) = C_h A_h S_h$ -ra következik

$$K_h(1) (\alpha v_h^{(j)} + \beta v_h^{(j')}) = \frac{h^2}{4} [(1 - \omega\lambda_j^h)\lambda_j^h \alpha + (1 - \omega\lambda_{j'}^h)\lambda_{j'}^h \beta] (v_h^{(j)} + v_h^{(j')}).$$

Ezen összefüggés mindkét oldalát megszorozva  $\sqrt{2}$ -vel, a  $v_h^{(j)}$  normált sajátvektorokhoz megyünk át. Ekkor hivatkozhatunk az

$$\|y\| = \|y\|_{L_2(\omega_h)} := \left( \sum_{i=1}^{2N-1} y_i^2 h \right)^{1/2}$$

súlyozott euklideszi normában érvényes

$$\left\| \sum_{k=1}^{2N-1} \alpha_k v_h^{(k)} \right\|^2 = \sum_{k=1}^{2N-1} \alpha_k^2$$

relációra. Mivel nálunk csak két sajátvektor van jelen, így

$$\begin{aligned} \|K_h(1) (\alpha v_h^{(j)} + \beta v_h^{(j')})\|^2 &= \frac{h^4}{16} [(1 - \omega\lambda_j^h)\lambda_j^h \alpha + (1 - \omega\lambda_{j'}^h)\lambda_{j'}^h \beta]^2 2 \leq \\ &\leq \frac{h^4}{8} \sum_{k=j, j'} ((1 - \omega\lambda_k^h)\lambda_k^h)^2 (\alpha^2 + \beta^2), \end{aligned}$$

a Cauchy-egyenlőtlenség segítségével. Ez igaz  $j = j' = N$  esetén is, amikor  $\alpha = 1$ ,  $\beta = 0$ . Eredményünket a következő formában is felírhatjuk :

$$\|K_h(1) (\alpha v_h^{(j)} + \beta v_h^{(j')})\|^2 \leq \frac{h^4}{8\omega^2} [(1-t)^2 t^2 + (1-(2\rho-t))^2 (2\rho-t)^2] (\alpha^2 + \beta^2). \quad (15.101)$$

Itt a következő jelöléseket vezettük be :  $\rho := \frac{2}{h^2}\omega$ ,  $t := \omega\lambda_j^h \in [0, 2\rho]$ , és azt használtuk fel, hogy

$$\omega\lambda_{j'} = 2\rho \cos^2 \frac{j\pi h}{2} = 2\rho - \omega\lambda_j = 2\rho - t.$$

Ezekkel a jelölésekkel (15.101)-ben az  $\alpha^2 + \beta^2$  szorzója

$$f_1(\rho, t) := \frac{1}{2\rho^2} [(1-t)^2 t^2 + (1-2\rho+t)^2 (2\rho-t)^2].$$

Ha

$$y = \sum_{j=1}^{2N-1} \alpha_j v_h^{(j)},$$

akkor (15.101) szerint

$$\begin{aligned} \|K_h(1)y\|^2 &\leq \alpha_N^2 \|K_h(1)v_h^{(N)}\|^2 + \sum_{j=1}^{N-1} \|K_h(1)(\alpha_j v_h^{(j)} + \alpha_{j'} v_h^{(j')})\|^2 \\ &\leq \max_{0 \leq t \leq 2\rho} f_1(\rho, t) \sum_{j=1}^{2N-1} \alpha_j^2 = \max_{0 \leq t \leq 2\rho} f_1(\rho, t) \|y\|^2. \end{aligned}$$

Mivel  $f_1(\rho, t) \geq 0$ , a konvergencia attól függ, vajon  $\max_{0 \leq t \leq 2\rho} f_1(\rho, t) < 1$ . Nem nehéz kiszámítani, hogy ez akkor áll fenn, amikor

$$(2 - \sqrt{2})/4 < \rho < (2 + \sqrt{2})/4,$$

vagyis ha

$$(2 - \sqrt{2}) \frac{h^2}{8} < \omega < (2 + \sqrt{2}) \frac{h^2}{8}.$$

A  $\rho$ -intervallum középpontja  $\rho = 1/2$ , azaz

$$\omega = \frac{h^2}{4} = \frac{1}{\|A_h\|_{C(\omega_h)}},$$

ld. (15.101), és ekkor  $\max_{0 \leq t \leq 2\rho} f_1(\rho, t) = \frac{1}{4}$ .

De van optimális iterációs paraméter is. Ennek meghatározásához vegyük észre, hogy

$$f_1(\rho, t) = \frac{\rho^2}{8} \left( T_4\left(\frac{t-\rho}{\rho}\right) - 1 \right) + (1-\rho)^2,$$

ha  $\rho = \rho_{opt} := (3 + \sqrt{2})/7$ . Itt  $T_4$  jelöli a negyedfokú Csebisev-polinomot, ld. I. 294. o. Mivel az a  $[-1, 1]$  intervallumban veszi fel maximális 1 értékét pl. -1-nél (ami  $t = 0$ -nak felel meg), így - újra  $f_1$  nemnegativitását is figyelembe véve - végülis

$$\min_{\rho > 0} \max_{0 \leq t \leq 2\rho} f_1(\rho, t) = f_1(\rho_{opt}, 0) = (1 - \rho_{opt})^2 = \frac{18 - 8\sqrt{2}}{49} = 0.13645 \dots,$$

és  $\omega_{opt} = \frac{h^2}{2} \rho_{opt} := (3 + \sqrt{2}) \frac{h^2}{14}$ .

Akár ezt az értéket vesszük, akár  $\omega = \frac{h^2}{4}$ -et, a konvergencia sebessége nem függ  $h$ -tól! Az optimális  $\omega$  esetén pl. 4 iteráció után a hiba már kisebb mint az eredetinek 50-öd része, és

$$\|K_h(1)\| \leq \frac{(18 - 8\sqrt{2})^{1/2}}{7} = 0.36939 \dots$$

Ha egynél több simító iterációt végzünk, akkor még jobb lehetőségeink is vannak.

Ha pl. két simító iterációt veszünk, változatlan  $\omega$ -val, akkor az egyébként megegyező előbbi jelölésekkel legyen

$$f_2(\rho, t) := \frac{1}{2\rho^2} [(1-t)^4 t^2 + (1-2\rho+t)^4 (2\rho-t)^2].$$

Ekkor  $K_h(2) := C_h A_h S_h^2$  becsülhető a következőképpen :

$$\|K_h(2) (\alpha v_h^{(j)} + \beta v_h^{(j')})\|^2 \leq f_2(\rho, t) (\alpha^2 + \beta^2). \quad (15.102)$$

Amennyiben most  $\rho = \frac{2}{3}$ , akkor a  $T_6$  Csebisev-polinommal teljesül

$$f_2(\rho, t) = \frac{1}{162} \left( T_6\left(\frac{t-\rho}{\rho}\right) + 3 \right),$$

tehát

$$\min_{\rho > 0} \max_{0 \leq t \leq 2\rho} f_2(\rho, t) = f_2\left(\frac{2}{3}, 0\right) = \frac{2}{81} = 0.02469 \dots \quad (15.103)$$

A  $\rho = \frac{2}{3}$  azt jelenti, hogy  $\omega = \frac{h^2}{3}$ , amivel már előbb a 15.4.2. pont (15.89) összefüggésében találkoztunk. Ilyenkor (15.102) és (15.103) szerint

$$\|K_h(2)\| \leq \frac{\sqrt{2}}{9} = 0.15713 \dots$$

Az előbbi gondolatsorhoz hasonlóan azt az esetet is lehet vizsgálni, hogy az első és a második iterációs lépésben más-más iterációs paramétert választunk (ld. a 23. feladatot).

Térjünk most vissza a (15.100) azonossághoz! Azt úgy is lehet érteni, hogy igaz

$$C_h v_h^{(j)} = \frac{h^2}{4} \left( v_h^{(j)} + v_h^{(j')} \right),$$

vagyis

$$\begin{aligned} C_h \left( \alpha v_h^{(j)} + \beta v_h^{(j')} \right) &= (\alpha + \beta) \frac{h^2}{4} \left( v_h^{(j)} + v_h^{(j')} \right), \\ \|C_h \left( \alpha v_h^{(j)} + \beta v_h^{(j')} \right)\|^2 &= (\alpha + \beta)^2 \frac{h^4}{16} \left( \|v_h^{(j)}\|^2 + \|v_h^{(j')}\|^2 \right) \leq (\alpha^2 + \beta^2) \frac{h^4}{4}. \end{aligned}$$

Innen, megint a  $\{v_h^{(j)}\}$  rendszer ortogonalitása miatt, következik, hogy

$$\frac{h^2}{2} \geq \|C_h\| = \|A_h^{-1} - PA_H^{-1}R\|.$$

Ez nyilván a (15.94) approximációs tulajdonság speciális esete :  $\alpha = 2$ ,  $c_a = \frac{1}{2}$ . Másképpen kifejezve: a  $PA_H^{-1}R$  operátor másodrendben approximálja  $A_h^{-1}$ -t.

Ez utóbbi tulajdonságot az általános, többdimenziós esetben is be lehet bizonyítani, megfelelő feltételek mellett.

Ehhez a ponthoz ld. a 24. feladatot is.

### 15.4.7 A simítási és az approximációs tulajdonság

Az approximációs tulajdonság a következőt jelenti az általános, akár többdimenziós esetben. Az

$$A_h y_h = f_h \tag{15.104}$$

diszkrét elliptikus feladat mellett (ahol  $f_h$ -t a differenciálegyenlet  $f$  jobboldalából és a peremértékekből számítjuk ki) vizsgáljuk az

$$A_H y_H = R_h^H f_h \tag{15.105}$$

egyenletet is, és keressük egy

$$\|y_h - P_H^h y_H\|_h \leq ch^2 \|f_h\|_h \tag{15.106}$$

alakú becslés az  $\omega_h$  rácson definiált  $\|\cdot\|_h$  normában és a  $h$ -tól független  $c$  konstanssal. Ha ez teljesül minden  $f_h$ -ra, akkor ugyanis

$$\|(A_h^{-1} - P_H^h A_H^{-1} R_h^H) f_h\|_h = \|y_h - P_H^h y_H\|_h \leq ch^2 \|f_h\|_h,$$

vagyis ekkor érvényes (15.94),  $\alpha = 2$ ,  $c_a = c$ . A (15.106) becslés bizonyíthatósága nem csak az  $A_h$  és  $A_H$  mátrixoktól (azaz a megfelelő diszkrétizációtól), valamint a restrikció és interpoláció kiválasztásától függ, hanem az eredeti feladat simaságától is, azaz a differenciálegyenlet együttthatóinak, jobboldalának, peremértékeinek, és a tartomány peremének differenciálhatóságától.

A differenciasémáknál szokásosak az

$$\|y_h - R_h u\|_h = O(h^2)$$

alakú becslések, ahol  $\|\cdot\|_h$  pl. a  $C(\omega_h)$ - vagy a  $L_2(\omega_h)$ -norma és  $R_h u$  a folytonos pontos megoldás restrikiója a rácsfüggvények terére. Ezen becslés klasszikus módszerekkel való levezetése esetén (ld. a 15.3. pontot) eléggé erős simasági feltételek kellene, mert az  $O(h^2)$  tag tartalmazza az  $u$  deriváltjainak normáját (míg végeselem approximáció esetén ismeretes, hogy  $\|u - u_h\|_0 \leq Mh^2\|f\|_0$ , ami megkönnyíti az approximációs tulajdonság bizonyítását, ld. 15.8.1.).

Ezután viszont

$$\|w_H - R_H u\|_H = O(H^2)$$

is érvényes, ahol  $A_H w_H = f_H$ , és

$$\|y_H - R_H u\|_H = \|w_H - R_H u\|_H + \|y_H - w_H\|_H = O(H^2) = O(h^2)$$

is kézenfekvő, ha

$$\|R_h^H f_h - f_H\|_H = O(H^2).$$

Végül, ha

$$\|R_h u - P_H^h R_H u\|_h = O(h^2) \text{ és } \|P_H^h\| \leq c_0,$$

(ahol gyakran  $c_0 = 1$ -re számíthatunk,) akkor

$$\begin{aligned} \|y_h - P_H^h y_H\|_h &\leq \|y_h - R_h u\|_h + \|R_h u - P_H^h R_H u\|_h + \|P_H^h R_H u - P_H^h y_H\|_h \\ &\leq O(h^2) + O(h^2) + c_0 \|R_H u - y_H\|_H = O(h^2). \end{aligned}$$

Az itt mutatott elvi becslési lehetőség aláhúzza, hogy a többrácsos módszer egyes elemei (az approximáció ( $A_h$ ), az interpoláció és a restrikió) között szoros összefüggés kell, hogy legyen. Hátránya, hogy túl sok simaságot kell feltételezni – és hogy *a normák nem összehangoltak* a simító tulajdonság bizonyításánál alkalmazott normákkal a következő értelemben. Ha a célunk (15.94)-nek és (15.95)-nek megfelelően a

$$\|K_h(\nu)\|_{0,h} = \|C_h \cdot A_h S_h^\nu\|_{0,h}$$

mátrix normájának becslése pl. a  $\|\cdot\|_{0,h}$  diszkrét  $L_2$ -féle normában, akkor ehhez ugyan nem kell, hogy  $K_h(\nu)$ -vel együtt  $C_h$ -t is és  $A_h S_h^\nu$ -t is mint  $L_2(\omega_h) \rightarrow L_2(\omega_h)$  leképezést értelmezzük. De amennyiben bizonyos  $W(\omega_h)$  diszkrét függvényterrel vesszük

$$A_h S_h^\nu : L_2(\omega_h) \rightarrow W(\omega_h),$$

akkor kell, hogy legyen

$$C_h : W(\omega_h) \rightarrow L_2(\omega_h).$$

Ennek megfelelően (15.95) azt jelenti, hogy a

$$\|A_h S_h^\nu f_h\|_{W(\omega_h)} \leq c_s h^{-\alpha} \|f_h\|_{L_2(\omega_h)} \quad (15.107)$$

becslésnek kell teljesülnie, míg (15.94) azt, hogy

$$\|C_h y_h\|_{L_2(\omega_h)} \leq c_a h^\alpha \|y_h\|_{W(\omega_h)}. \quad (15.108)$$

A  $W(\omega_h)$  térrel együtt változnak a (15.107)-ben és (15.108)-ban szereplő normák is.

Ezt a problémát meg lehet oldani : az ilyen vizsgálat N.Sz. Bahvalov, M. Dryja, W. Hackbusch és más szerzőknek az érdeme, és végülis olyan követelmény mellett adja a differenciasémák esetén az approximációs tulajdonság érvényességét, hogy *a restrikció rendje plusz a prolongáció rendje legyen nagyobb mint az elliptikus operátor  $2m$  rendje*. Ennek levezetésére nem térünk ki. Itt a prolongáció  $p$  rendje azzal definiált, hogy igaz

$$\|P_H^h R_H u - R_h u\|_h \leq O(h^p)$$

minden elég sima  $u$  függvényen. A restrikció transzponáltját (ld. 15.4.6.) értelmezhetjük mint interpolációt, és ennek  $q$  rendje adja a restrikció rendjét :

$$\|(R_h^H)^T P_H u - P_h u\|_h \leq O(h^q).$$

Mint példát említjük, hogy másodrendű feladat esetén az előbbi eredmények értelmében kell, hogy  $p+q \geq 3$  legyen. Ekkor megfelelő pl. a bilineáris interpoláció, valamint a bilineáris restrikció használata, mert mindkettőnek a rendje 2. Amennyiben viszont mint restrikciót a direkt átvitelt választjuk (amelynek rendje 0), akkor az interpoláció rendje kell legyen legalább 3 (amit a köbös interpoláció teljesít). Ezt az esetet vizsgálta N.Sz. Bahvalov, és erre bizonyította be elsőnek a többrácsos módszer konvergenciáját általános elliptikus operátorra.

Ha (15.108) bizonyítása sikerült, mondjuk

$$\|C_h\| = \|A_h^{-1} - P A_H^{-1} R\| \leq c_a h^2$$

alakjában, akkor emellett az approximációs tulajdonság mellett a  $K_h(\nu)$  mátrix normájának konvergencia becsléséhez elégséges az

$$\|A_h S_h^\nu\| \leq \frac{c_s}{h^2(1+\nu)} \quad (15.109)$$



alakú becslés, amelyben most  $c_s$  független  $\nu$ -tól. Ezzel a simítási tulajdonsággal most foglalkozunk olyan feltételek mellett, amelyek tipikusak másodrendű differenciálegyenletek differencia-approximációjánál.

**15.14. Lemma** (az egyszerű iteráció simító tulajdonsága). Legyen  $A_h = A_h^T > 0$ ,  $\|A_h\|_{C(\omega_h)} \leq c_0/h^2$  és a  $\|\cdot\|$  norma az euklideszi. Az egyszerű iterációnak

$$h^2/c_0 \leq \omega \leq 1/\|A_h\|$$

esetén megvan a (15.109) simító tulajdonsága  $c_s = c_0/2$  konstanssal.

[Kommentár: Ha pl.  $\omega = 1/\|A_h\|_{C(\omega_h)}$ , akkor a feltétel teljesül, ld. a 25. feladatot is.]

**Bizonyítás.** Feltételeink mellett hasonló optimalizációs problémához jutunk, mint az előző pontban  $f_k(\rho, t)$ -vel kapcsolatban ( $k = 1, 2$ ).

Most

$$\begin{aligned} S_h^\nu &= (I_h - \omega A_h)^\nu, \\ \|A_h S_h^\nu\| &= \max_{\lambda_j^h = \lambda_j^h(A_h)} |\lambda_j^h (1 - \omega \lambda_j^h)^\nu| \leq \frac{1}{\omega} \max_{0 \leq t \leq t_{\max}} |f_\nu(t)|, \end{aligned}$$

ha

$$t := \omega \lambda_j^h \in [\omega \lambda_{\min}^h, \omega \lambda_{\max}^h] \subset [0, t_{\max}], \quad f_\nu(t) := t(1-t)^\nu.$$

$[0, 1]$ -ben  $f_\nu$  felveszi maximumát  $t = t_\nu = \frac{1}{1+\nu}$ -nél ( $\nu \geq 1$ ), ami feltételeink mellett mérvadó, mert  $t_{\max} := \omega \lambda_{\max}^h = \omega \|A_h\| \leq 1$ . Ekkor

$$\|A_h S_h^\nu\| \leq \frac{1}{\omega(1+\nu)} \left( \frac{\nu}{1+\nu} \right)^\nu \leq \frac{c_0}{h^2} \frac{1}{2(1+\nu)}. \quad (15.110)$$

Itt egyrészt  $1/\omega$  becslését alkalmaztuk, másrészt azt, hogy

$$2 \leq \left( 1 + \frac{1}{\nu} \right)^\nu \rightarrow e, \quad \nu \rightarrow \infty.$$

Tehát aszimptotikusan

$$\max_{[0,1]} |f_\nu(t)| \sim \frac{1}{e(\nu+1)},$$

vagyis : nem sokat veszítettünk a (15.110) becslésnél.  $\square$

**Megjegyzések. 1.** Ha alkalmazható a lemma és érvényes  $\|C_h\| \leq c_a h^2$ , akkor most az következik, hogy

$$\|K_h\| \leq c_a h^2 \frac{c_0}{h^2} \frac{1}{2(1+\nu)} = \frac{c_a c_0}{2(1+\nu)} \ll 1$$

ha  $\nu$  elég nagy, függetlenül  $h$ -tól.

**2.** Vegyük észre, hogy  $\omega$  megfelelő választásához elegendő az  $A_h$  maximális sajátértékét vagy annak egy  $c_0/h^2$  alakú felső becslését ismerni!

**3.** Az egydimenziós esetben igaz  $\|A_h\| = \|A_h\|_{L_2(\omega_h)} = \frac{4}{h^2} \cos^2 \frac{\pi h}{2} < \|A_h\|_{C(\omega_h)} = \frac{4}{h^2}$  és  $\|C_h\|_{L_2(\omega_h)} \leq \frac{h^2}{2}$ , ami azt jelenti, hogy  $\omega = \frac{h^2}{4}$  esetén

$$\|K_h(\nu)\|_{L_2(\omega_h)} \leq \frac{h^2}{2} \cdot \frac{2}{h^2(1+\nu)} = \frac{1}{1+\nu},$$

tehát a fenti elemzésünkkel egybehangzóan  $\|K_h(\nu)\|_{L_2(\omega_h)} \leq \frac{1}{2}$ , mihelyt  $\nu \geq 1$ . Ha viszont az  $\omega$  paramétert a  $[\frac{h^2}{12}, \frac{h^2}{4}]$  intervallumból választjuk és  $\nu \geq 5$ , akkor  $\|K_h(\nu)\|_{L_2(\omega_h)} \leq \frac{3}{1+\nu}$ .

**4.** Felhívjuk a figyelmet a következő körülményre : Az  $\omega = 1/\|A_h\|$  esetén adódó

$$\|A_h S_h^\nu\| \leq \|A_h\| \frac{1}{2(1+\nu)}$$

alakú becslést nem lehet levezetni az

$$\|A_h S_h^\nu\| \leq \|A_h\| \|S_h^\nu\|$$

egyenlőtlenség alapján: ha szimmetrikus  $A_h$ , a norma az euklideszi norma (mint ahogyan feltételeztük), akkor I. 1.6.6. szerint

$$\begin{aligned} \|S_h^\nu\| &= \max_{\lambda^h = \lambda_j^h(A_h)} |(1 - \omega \lambda^h)^\nu| = \left( \frac{\text{cond } A_h - 1}{\text{cond } A_h + 1} \right)^\nu \\ &= \left( 1 - \frac{2\nu}{(1 + \text{cond } A_h) \nu} \right)^\nu \lesssim \exp \left( -\frac{2\nu}{1 + \text{cond } A_h} \right). \end{aligned}$$

Ekkor az egész becslés használhatatlan lesz, természetesen csak azt mutatja, hogy a simító iteráció nagyon lassan konvergál.

A  $h$ -tól (ill.  $\text{cond } A_h$ -tól) nem függő, gyors konvergenciának a bizonyítása annak köszönhető, hogy

$$(1 - \omega \lambda_j^h)^\nu \quad \text{helyett} \quad \lambda_j(1 - \omega \lambda_j^h)^\nu.$$

szerepelt.

**5.** A lemma csak a  $\nu_1 = \nu$ ,  $\nu_2 = 0$  esetről szól, de valójában az általános esetről is ad információt, amikor utóiterációt is alkalmazunk. Legyen ugyanis

$$C_h = A_h^{-1} - P A_H^{-1} R, \quad \text{és} \quad K_h(\nu_1, \nu_2) := S_h^{\nu_2} C_h A_h S_h^{\nu_1},$$

akkor (ehhez ld. a 26. feladatot is)  $\nu_1 + \nu_2 =: \nu$ -vel

$$\rho(S_h^{\nu_2} C_h A_h S_h^{\nu_1}) = \rho(C_h A_h S_h^{\nu_1 + \nu_2}) \leq \|C_h\| \|A_h S_h^\nu\| \leq \frac{c_a c_0}{2(1+\nu)},$$

másrészt az iteráció során úgylis az utósimító iterációt követi az elősimító iteráció :

$$K_h^\mu(\nu_1, \nu_2) = S_h^{\nu_2} C_h(A_h S_h^{\nu_1}) C_h(A_h S_h^{\nu_2}) \dots C_h A_h S_h^{\nu_1},$$

ahol  $\mu - 1$  darab  $C_h A_h S_h^\nu$  szorzó áll. Tehát

$$\|K_h^\mu(\nu_1, \nu_2)\| \leq \|S_h^{\nu_2}\| \|C_h\| \|A_h S_h^{\nu_1}\| \dots \|C_h\| \|A_h S_h^{\nu_2}\| \|C_h\| \|A_h S_h^{\nu_1}\|,$$

és megint  $\|A_h S_h^\nu\|$  becslése döntő.

**6.** (V.V. Sajdurov) Legyen rögzített az iterációk  $\nu$  száma és engedjük meg, hogy minden iterációs lépésben más paramétert alkalmazunk – úgy mint I. 1.6.7-ben a Csebisev-iterációnál. Ezt a lehetőséget az előző pontban is említettük. Ekkor optimálisan választva az  $\omega$  paramétert elérhető, hogy

$$\|A_h S_h^\nu\| \leq \frac{c_0}{h^2(1 + \nu)^2}$$

teljesüljön.

**7.** A simító tulajdonság bizonyítása birtokában érdemes visszatekinteni a 15.4.4-beli, a műveletigényre vonatkozó becsléseinkre. A  $\nu_1$  és  $\nu_2$  iterációszámok ott feltételezett korlátossága megfelelően bizonyult, hiszen biztosíthatja, hogy  $\|K_h(\nu_1, \nu_2)\| \ll 1$ , viszont a gyakorlatilag elegendőnek említett  $\nu_1 + \nu_2 \leq 2$  (bár az „Ellipsz” programmal ellenőrizhető de) az itt bemutatott elmélettel nem garantálható. Mindenesetre a (15.91) műveletigényben a döntő rész az iterációkkal kapcsolatos.  $\square$

Mint következő simító iterációt tekintünk a szimmetrikus Gauss–Seidel iterációt! Ehhez felidézünk ezen módszer képleteit (v.ö. I. 119. o.) abban az esetben, amikor feladatunk az  $Ay = b$  egyenlet megoldása, ahol  $A$  szimmetrikus és pozitív definit mátrix :

$$\begin{aligned} A &= L + D + U, & L^T &= U, \\ & & y^{(0)} &\text{adott,} \\ (D + L)y^{(m+1/2)} &= -Uy^{(m)} + b, \\ (D + U)y^{(m+1)} &= -Ly^{(m+1/2)} + b, & m &= 0, 1, \dots \end{aligned}$$

Eliminálva az  $y^{(m+1/2)}$ -t azt kapjuk (feltéve  $D$  regularitását, amely  $D + U = (D + L)^T$  regularitását is jelenti), hogy

$$(D + U)y^{(m+1)} = L(D + L)^{-1}Uy^{(m)} + b - L(D + L)^{-1}b.$$

Innen azt olvassuk le, hogy

$$S = (D + U)^{-1}L(D + L)^{-1}U$$

az iterációs mátrix. Ezt az I. 1.27. tétel bizonyításában látott módon (ahol  $W^{-1}$ -et  $V$ -vel jelöltük) alakítjuk át :

$$S = I - W^{-1}A, \quad \text{ahol} \quad W := (D + L)D^{-1}(D + U).$$

Ez a  $W$  a prekondicionálási mátrix, amelynek segítségével az iterációt az I. 1.6.3. pont (1.80) képlet alakjában írhatjuk. Mostani jelöléseinkkel ez

$$W(y^{(m+1)} - y^{(m)}) + Ay^{(m)} = b, \quad m = 0, 1, \dots \quad (15.111)$$

Vegyük észre, hogy  $W$  inverzét nem kell explicite kiszámítanunk és  $A$ -n felül nem keletkezik további tárigény! Ha valamilyen  $b$  vektorra érdekel  $W^{-1}b$ , akkor  $y^{(0)} = 0$ -val a szimmetrikus Gauss-Seidel iterációt beindítjuk és kapjuk  $y^{(1)} = W^{-1}b$ .

Néhány, a simító iterációkkal kapcsolatban érdekes esetet (15.111)-alakú iterációra sorolunk az alábbiakban :

$$W := \begin{cases} \frac{1}{\omega}I, & \text{az egyszerű iteráció,} \\ D, & \text{a Jacobi-iteráció,} \\ D + L, & \text{a Gauss-Seidel módszer,} \\ (D + L)D^{-1}(D + U), & \text{a szimmetrikus Gauss-Seidel módszer.} \end{cases}$$

Jelen esetünkben  $W$  nem csak szimmetrikus mátrix (ez rögtön  $L^T = U$ -ból következik), hanem pozitív definit is :

$$\begin{aligned} W &= (D + L)(I + D^{-1}U) = D + L + U + LD^{-1}U = A + LD^{-1}U, \\ (LD^{-1}Ux, x) &= (D^{-1}Ux, Ux) \geq 0, \end{aligned}$$

hiszen  $A$ -val együtt  $D$  is pozitív definit. Innen

$$W = A + LD^{-1}U \geq A > 0.$$

Maga az  $S$  mátrix nem feltétlenül szimmetrikus, de

$$X := W^{-1/2}AW^{-1/2}$$

igen. Most tekintsük a (15.84) egyenletet, amikor ott  $A_h$  szimmetrikus és pozitív definit, és az  $A_h$ -ből levezethető mátrixokat  $h$  indexszel látjuk el :  $A_h = L_h + D_h + L_h^T$ ;  $W_h = A_h + L_h D_h^{-1} L_h^T$ ,  $X_h = W_h^{-1/2} A_h W_h^{-1/2}$ .  $A$  szimmetrikus és pozitív definit  $W_h$  mátrixhoz norma rendelhető hozzá :

$$\|x\|_{W_h} := (W_h x, x)^{1/2} = \|W_h^{1/2} x\|,$$

és ha a kétrácsos módszer  $K_h = C_h A_h S_h^\nu$  iterációs mátrixát ebben a  $W_h$ -normában vesszük :

$$\|K_h\|_{W_h} = \max_{x \neq 0} \frac{\|K_h x\|_{W_h}}{\|x\|_{W_h}} = \|W_h^{1/2} K_h W_h^{-1/2}\|,$$

akkor

$$\|K_h\|_{W_h} = \|W_h^{1/2} K_h W_h^{-1/2}\| \leq \|W_h^{1/2} C_h W_h^{1/2}\| \cdot \|W_h^{-1/2} A_h S_h^\nu W_h^{-1/2}\|.$$

Vegyük észre, hogy ez egy konkrét példa összehangolt normákra! Mégpedig az  $A_h S_h^\nu$  mátrixot tekintjük mint az

$$\{\mathbb{R}^n, W_h\} \rightarrow \{\mathbb{R}^n, W_h^{-1}\}$$

leképezést (ahol  $\{\mathbb{R}^n, W_h\}$  jelenti a  $W_h$ -normával ellátott  $\mathbb{R}^n$ -et stb.), és  $C_h$ -t tekintjük mint az

$$\{\mathbb{R}^n, W_h^{-1}\} \rightarrow \{\mathbb{R}^n, W_h\}$$

leképezést. Ekkor az  $A_h S_h^\nu$  normája éppen  $\|W_h^{-1/2} A_h S_h^\nu W_h^{-1/2}\|$ , hiszen

$$\begin{aligned} \|A_h S_h^\nu x\|_{W_h^{-1}} &= \|W_h^{-1/2} A_h S_h^\nu x\| = \|(W_h^{-1/2} A_h S_h^\nu W_h^{-1/2}) W_h^{1/2} x\| \\ &\leq \|W_h^{-1/2} A_h S_h^\nu W_h^{-1/2}\| \|W_h^{1/2} x\| \\ &= \|W_h^{-1/2} A_h S_h^\nu W_h^{-1/2}\| \|x\|_{W_h} \end{aligned}$$

javíthatatlan becslés. Hasonlóan  $C_h$  normája  $\|W_h^{1/2} C_h W_h^{1/2}\|$ .

Ezek miatt az approximációs tulajdonságot most a

$$\|W_h^{1/2} (A_h^{-1} - P_H^h A_H^{-1} R_H^H) W_h^{1/2}\| \leq c_a$$

alakjában kell bizonyítani, a simítási tulajdonság viszont teljesüljön így :

$$\|W_h^{-1/2} A_h S_h^\nu W_h^{-1/2}\| \leq \frac{c_0}{2(1+\nu)},$$

(15.92)-nek és (15.93)-nak megfelelően.  $h$ -hatványok nem keletkeznek, mert a normabeli kifejezésekben  $A_h$  és  $W_h \geq A_h$  egyensúlyban vannak.

**15.15. Lemma** (a szimmetrikus Gauss–Seidel iteráció simítási tulajdonsága). Legyen  $A_h$  szimmetrikus és pozitív definit. Ekkor a szimmetrikus Gauss–Seidel iteráció rendelkezik a simítási tulajdonsággal abban az értelemben, hogy igaz a

$$\|W_h^{-1/2} A_h S_h^\nu W_h^{-1/2}\| = \|X_h (I_h - X_h)^\nu\| \leq \frac{1}{2(1+\nu)}$$

becslés.

**Bizonyítás.** Elsőnek lássuk be, hogy  $X_h$  pozitív definit. Pontosabban igaz  $0 < X_h \leq I_h$ . Definíciója szerint ugyanis egyrészt

$$(X_h x, x) = (A_h W_h^{-1/2} x, W_h^{-1/2} x) \geq 0,$$

és itt az egyenlőség csak  $x = 0$  esetén igaz, másrészt  $0 < A_h \leq W_h$  alapján az

$$(A_h W_h^{-1/2} x, W_h^{-1/2} x) \leq (W_h W_h^{-1/2} x, W_h^{-1/2} x) = \|x\|^2$$

relációból következik  $X_h \leq I_h$ . Ekkor

$$W_h^{1/2} S_h W_h^{-1/2} = I_h - X_h \geq 0 \quad \text{és} \quad W_h^{1/2} S_h W_h^{-1/2} = I_h - X_h \leq I_h.$$

A simítási tulajdonság ezután abból adódik, hogy mivel

$$W_h^{1/2} S_h W_h^{-1/2} = I_h - X_h \quad \text{és} \quad W_h^{1/2} S_h^\nu W_h^{-1/2} = (I_h - X_h)^\nu,$$

így

$$\|W_h^{-1/2} A_h S_h^\nu W_h^{-1/2}\| = \|X_h (I_h - X_h)^\nu\| \leq \frac{1}{2(1+\nu)},$$

ld. a 15.14. lemma bizonyítását.  $\square$

Amennyiben az  $A_h$  mátrix nem feltétlenül szimmetrikus, akkor is használhatjuk a (15.111) iterációt. Ekkor a következő eredmény biztosítja a simítási tulajdonságot.

**15.16. Lemma** (Reusken; Ecker–Zulehner; simítási tulajdonság a nem-szimmetrikus esetben). Legyen  $A_h := P_h - Q_h$ , ahol  $\|P_h\| \leq c_P \|A_h\|$  és  $\|I_h - P_h^{-1} A_h\| \leq 1$ . Legyen  $W_h := \frac{1}{\omega} P_h$  a (15.111) prekondicionálási mátrixa az  $\omega \in (0, 1)$  iterációs paraméterrel.

Ekkor az  $S_h := I_h - \omega P_h^{-1} A_h$  iterációs mátrixszal teljesül a következő becslés:

$$\|A_h S_h^\nu\| \leq c_P \|A_h\| \eta(\omega, \nu),$$

ahol

$$\eta(\omega, \nu) := \begin{cases} 2 \max(1 - \omega, \omega), & \text{ha } \nu = 1, \\ 2 \max((1 - \omega)^2, 2\omega(1 - \omega), \omega^2), & \text{ha } \nu = 2, \\ (2/(\pi\nu(\omega(1 - \omega) - \frac{\nu}{(1+\nu)^2})))^{1/2}, & \text{ha } \nu \geq 2, \quad \frac{1}{1+\nu} < \omega < \frac{\nu}{1+\nu}. \quad \square \end{cases}$$

**Megjegyzések. 1.** Nagy  $\nu$ -re ezek szerint  $\eta(\omega, \nu) = O(\nu^{-1/2})$ . Páratlan  $\nu$ -re az  $\omega = \frac{1}{2}$  érték előnyös, páros  $\nu$ -re az  $\omega = \frac{1}{2} \pm \frac{1}{2(1+\nu)}$ , ld. Ecker és Zulehner dolgozatát.

**2.** Az  $\|I_h - P_h^{-1} A_h\| \leq 1$  feltétel teljesíthető, ha  $P_h$ -t az inkomplett Gauss-eliminációval állítjuk elő (ld. I. 1.6.5.), és ennek során elegendően sok nemnulla elemet  $A_h$ -ból veszünk figyelembe.  $\square$

### 15.4.8 A többrácsos módszer konvergenciája

A kétrácsos módszer vizsgálata után a többrácsos módszerrel foglalkozunk. Ekkor, mint már az algoritmusoknál is, nem a  $h$  és  $H$  indexeket használjuk a szintek megkülönböztetésére, hanem ehelyett a szint sorszámát adjuk meg (az  $\ell$  indexszel). A többrácsos módszer iterációs mátrixát  $M_\ell$ -vel jelöljük. Elsőnek ezt számítjuk ki.

Amikor az  $(\ell - 1)$ -edik szinten ( $\ell > 2$ ) végrehajtjuk a  $\gamma$  iterációt, akkor a cél az  $A_{\ell-1}y_{\ell-1} = d_{\ell-1}$  egyenlet megoldása. Az iterációt

$$y_{\ell-1}^0 = 0, \quad y_{\ell-1}^{j+1} = M_{\ell-1}y_{\ell-1}^j + N_{\ell-1}d_{\ell-1}, \quad j = 0, \dots, \gamma - 1, \quad (15.112)$$

alakjában tudjuk leírni. Itt az  $N_{\ell-1}$  mátrix  $M_{\ell-1}$ -ből számítható ki, hiszen ismert, hogy – tetszőleges  $d_{\ell-1}$  esetén is – az iteráció fixpontja a pontos megoldás, azaz  $A_{\ell-1}^{-1}d_{\ell-1}$  :

$$A_{\ell-1}^{-1}d_{\ell-1} = M_{\ell-1}A_{\ell-1}^{-1}d_{\ell-1} + N_{\ell-1}d_{\ell-1}.$$

Tehát, mivel az előbbi reláció tetszőleges  $d_{\ell-1}$ -re áll fenn,

$$N_{\ell-1} = (I_{\ell-1} - M_{\ell-1})A_{\ell-1}^{-1}. \quad (15.113)$$

(Hadd mutassunk rá, hogy a többrácsos módszer végrehajtásához nem szükségeltetik az  $A_{\ell-1}^{-1}$  mátrix. (15.113) a módszer során használt  $N_{\ell-1}$ -nek csak egy felírásmódja.) (15.112)-ből kapjuk

$$y_{\ell-1}^\gamma = \sum_{j=0}^{\gamma-1} (M_{\ell-1})^j N_{\ell-1} d_{\ell-1} = (I_{\ell-1} - (M_{\ell-1})^\gamma) A_{\ell-1}^{-1} d_{\ell-1}.$$

Ezután térünk vissza az  $\ell$ -edik szintre,

$$y_\ell := y_\ell + P_{\ell-1}^\ell y_{\ell-1}^\gamma,$$

és  $\nu_2$  utósimító iteráció után a ciklust az  $\ell$ -edik szinten befejezzük,

$$y_\ell := S_\ell^{\nu_2}(A_\ell, d_\ell)y_\ell = S_\ell^{\nu_2} (y_\ell + P_{\ell-1}^\ell (I_{\ell-1} - (M_{\ell-1})^\gamma) A_{\ell-1}^{-1} d_{\ell-1}) + d_\ell\text{-tagok}.$$

Ehhez a számításhoz ld. a 22\*. feladatot \*\*\*. Figyelembe véve, hogy

$$d_{\ell-1} = R_\ell^{\ell-1}(d_\ell - A_\ell y_\ell) = -R_\ell^{\ell-1} A_\ell y_\ell + d_\ell\text{-tag},$$

ahol az  $y_\ell$ -t a ciklus elején rendelkezésére álló közelítésből  $\nu_1$  elősimító iterációval nyertük, végülis a teljes,  $y_\ell$ -re ható mátrix

$$M_\ell = S_\ell^{\nu_2} (I_{\ell-1} - P_{\ell-1}^\ell (I_{\ell-1} - M_{\ell-1}^\gamma) A_{\ell-1}^{-1} R_\ell^{\ell-1} A_\ell) S_\ell^{\nu_1}.$$

Ezt a képletet úgy alakítjuk át, hogy  $M_\ell$  lényeges része, a korábban vizsgált  $K_\ell = K_\ell(\nu_1, \nu_2)$  kétrácsos iterációs mátrix látszódjék :

$$\begin{aligned} M_\ell &= S_\ell^{\nu_2} (I_{\ell-1} - P_{\ell-1}^\ell A_{\ell-1}^{-1} R_\ell^{\ell-1} A_\ell) S_\ell^{\nu_1} + S_\ell^{\nu_2} P_{\ell-1}^\ell M_{\ell-1}^\gamma E_\ell \\ &= K_\ell + S_\ell^{\nu_2} P_{\ell-1}^\ell M_{\ell-1}^\gamma E_\ell, \\ E_\ell &:= A_{\ell-1}^{-1} R_\ell^{\ell-1} A_\ell S_\ell^{\nu_1}. \end{aligned} \quad (15.114)$$

Ha  $\ell = 2$ , akkor valóban a kétrácsos módszer működik : tehát

$$M_2 = K_2 \quad (\text{vagyis } M_1 := 0 \text{ vehető}).$$

A 15.14. és 15.15. lemmákból tudjuk, hogy  $\|K_\ell\|$  tetszőlegesen kicsi lesz, ha  $\nu$  elég nagy. A következő tétel megmutatja, hogy  $\|K_\ell\|$ -nél nem lényegesen nagyobb  $\zeta_\ell := \|M_\ell\|$ , és így a sima többrácsos módszer konvergens, ha  $\gamma = 2$ .

**15.17. Tétel** (a W-ciklusos sima többrácsos módszer konvergenciája).  
Legyen  $\gamma = 2$  és  $\eta_\ell := \|E_\ell\|$ , teljesüljön függetlenül  $\ell$ -től

$$\|P_{\ell-1}^\ell\| \leq 1, \quad \|P_{\ell-1}^\ell x\| \geq c_p \|x\|$$

minden  $x$  vektorra, és  $\|S_\ell^{\nu_2}\| \leq 1$ . Ha ekkor  $\|K_\ell\| \leq \zeta = \zeta(\nu) < \frac{1}{2} \min(1, \frac{c_p}{4})$ , akkor  $\|M_\ell\| = \zeta_\ell \leq 2\zeta < 1$  következik, pontosabban :

$$\zeta_2 \leq \zeta, \quad \zeta_\ell \leq \zeta + \zeta_{\ell-1}^2 \eta_\ell \quad \ell > 2,$$

függetlenül  $\ell$ -től.

**Bizonyítás.** Ha  $\ell = 2$ , akkor  $\zeta_2 = \|M_2\| = \|K_2\| = \zeta < \frac{1}{2}$ . Legyen most  $\gamma = 2$ . A (15.114)-beli  $S_\ell^{\nu_2} P_{\ell-1}^\ell M_{\ell-1}^\gamma E_\ell$  tag becslése kézenfekvő :

$$\|S_\ell^{\nu_2} P_{\ell-1}^\ell M_{\ell-1}^\gamma E_\ell\| \leq \|S_\ell^{\nu_2}\| \|P_{\ell-1}^\ell\| \|M_{\ell-1}\|^2 \|E_\ell\| \leq (\zeta_{\ell-1})^2 \eta_\ell.$$

Az  $\eta_\ell$  becsléséhez használjuk azt, hogy  $P_{\ell-1}^\ell E_\ell = S_\ell^{\nu_1} - K_\ell(\nu_1, 0)$  :

$$\eta_\ell \leq \frac{1}{c_p} \|P_{\ell-1}^\ell A_{\ell-1}^{-1} R_\ell^{\ell-1} A_\ell S_\ell^{\nu_1}\| = \frac{1}{c_p} \|S_\ell^{\nu_1} - K_\ell(\nu_1, 0)\| \leq \frac{2}{c_p},$$

mivel sem  $S_h^{\nu_1}$ , sem  $K_\ell$  nem növelik a normát. Akkor (15.114)-ből kapjuk

$$\zeta_\ell \leq \zeta + (\zeta_{\ell-1})^2 \eta_\ell \leq \zeta + (2\zeta)^2 \eta_\ell = \zeta(1 + 4\zeta \eta_\ell) < 2\zeta < 1,$$

hiszen  $\zeta_{\ell-1} \leq 2\zeta$ ,  $4\eta_\ell \zeta \leq \frac{8}{c_p} \zeta < 1$  és  $\zeta < \frac{1}{2}$ .  $\square$



**Megjegyzések. 1.** A  $\gamma > 2$  eset gyakorlatilag kevésbé érdekes, de erre általánosítható a bizonyítás. A fenti indukcióval mutathatjuk meg azt, hogy  $\zeta_\ell < \frac{\gamma}{\gamma-1}\zeta < 1$ , ha

$$\zeta < \frac{\gamma-1}{\gamma} \min\left(1, \left(\frac{c_p}{2\gamma}\right)^{1/(\gamma-1)}\right).$$

A  $\gamma = 1$  esetre 15.8.2-ben térünk vissza.

**2.** Azon eredmény birtokában, hogy a sima többrácsos módszer  $h$ -tól függetlenül konvergál, a megoldás  $\varepsilon$  pontossággal való kiszámításának teljes műveletigényét is megadhatjuk. Az  $\ell$ -edik szinten, mivel  $e_\ell^{(m)} = M_\ell^m e_\ell^{(0)}$ , tételünkből következik

$$\|e_\ell^{(m)}\| \leq \zeta_\ell^m \|e_\ell^{(0)}\| \leq (2\zeta)^m \|e_\ell^{(0)}\| \leq \varepsilon \|e_\ell^{(0)}\|,$$

ha  $m = \ln \frac{1}{\varepsilon} / \ln \frac{1}{2\zeta}$  iterációt hajtunk végre. 15.4.4. szerint egy iteráció műveletigénye  $O(N_\ell)$ , tehát összesen  $O(N_\ell \ln \frac{1}{\varepsilon})$  aritmetikai művelet szükséges a sima többrácsos módszer megvalósításához.  $\square$

Azután, hogy beláttuk, a (sima) többrácsos módszer  $h$ -tól függetlenül konvergál, már csak kevés kell annak a bizonyításához, hogy a teljes többrácsos módszer  $O(N_\ell)$  műveletigénnyel állítja elő a megoldást  $O(h^\kappa)$  pontossággal – ahol  $\kappa$  a diszkretizáció rendje.

A következő tétel ezt bizonyítja. Az eredmény a végeelem diszkretizáció esetén a következők miatt kézenfekvő :

Ha már az  $\ell$ -edik közbülső rácson a  $O(h_\ell^\kappa)$ -pontosságú (végeelem) megoldást számítottuk ki, akkor ez egyben az  $\ell + 1$ -dik rácson is a peremérték feladatnak egy (ugyanekkora hibájú) megoldása (a „kanonikus interpoláció” révén, ld. 15.8.1).

Legyen pl.  $2h_{\ell+1} = h_\ell$  és  $\kappa = 2$ , az  $\ell$ -edik szintű végeelem megoldás hibája legyen  $c_\ell h_\ell^2$ . Akkor az  $\ell + 1$ -edik szint szempontjából ez a hiba  $4c_\ell h_{\ell+1}^2$ , és az ott beindított sima többrácsos módszernek (mint a teljes többrácsos módszer részének) a feladata most az, hogy ezt a kiindulási hibát legalább a negyed részére csökkentse. Mivel a sima többrácsos módszer egy lépés alatt – a szintszámától függetlenül – a hibát a  $\zeta$ -ad részére tudja csökkenteni ( $\zeta = \text{const} < 1$ ), úgy a sima módszert elegendő  $\mu$ -szer alkalmazni, ahol  $\mu$  teljesíti a  $4\zeta^\mu \leq 1$  feltételt.

Legyen most  $A_\ell y_\ell = b_\ell$  az  $\ell$ -edik szinthez tartozó egyenletrendszer és  $\tilde{P}_{\ell-1}^\ell$  az az interpoláció, amelynek segítségével a teljes többrácsos módszer során magát a megoldást az  $\ell - 1$ -edik szintről az  $\ell$ -edik szintre visszük (ld. az algoritmust : a  $P_{\ell-1}^\ell$  interpoláció viszi át a megoldás javítását a magasabb szintre).

A tételhez fogalmazzunk meg néhány feltételt :

1. Legyen  $\|\tilde{P}_{\ell-1}^\ell\| \leq 1$  (ez nemcsak a bilineáris interpolációra igaz, hanem a bikvadratikus, sőt a biköbös interpolációra is);
2.  $\|\tilde{P}_{\ell-1}^\ell y_{\ell-1} - y_\ell\| \leq c_0(u)h_\ell^\kappa$ , ahol  $c_0(u) = \text{const}$ ,  $u$  a peremérték feladat megoldása (ezt úgy is lehet mondani, hogy  $\tilde{P}_{\ell-1}^\ell$  ne rontsa el a diszkretizáció rendjét – amihez kézenfekvő úgy választani  $\tilde{P}_{\ell-1}^\ell$ -t, hogy rendje legalább  $\kappa$  legyen (egydimenziós példánkban, ahol  $\kappa = 2$  volt, a szakaszonkénti lineáris interpoláció megfelel ennek a feltételnek);
3. az  $\ell$ -edik szinten a (sima) többrácsos módszer csökkentse a hibát  $\zeta_\ell$ -re, ahol  $\max_{1 \leq \ell \leq \ell_{\max}} \zeta_\ell =: \zeta < 1$  függetlenül  $h$ -tól és  $\ell_{\max}$ -tól.

Legyen  $\mu = \mu_1 = \mu_2$  azon iterációk száma, amelyeket az  $\ell$ -edik szinten (az MG eljárás hívásával) végzünk a teljes többrácsos módszer keretében.

**15.18. Tétel** (Hackbusch; teljes többrácsos módszer konvergenciája). Minden  $1 \leq \ell \leq \ell_{\max}$ -ra legyen  $h_{\ell-1} = 2h_\ell$ . Válasszuk  $\mu$ -t úgy, hogy  $\zeta^\mu \leq 1/(1+2^\kappa)$  és teljesüljenek az 1-3. feltételek. Ekkor a teljes többrácsos módszer konvergál,  $\tilde{y}_\ell$  megoldásának és az  $A_\ell y_\ell = b_\ell$  egyenletrendszer  $y_\ell$  megoldásának eltérésére igaz a következő becslés :

$$\|\tilde{y}_\ell - y_\ell\| \leq c_0 h_\ell^\kappa.$$

Itt  $c_0$  a 2. feltételben szereplő konstans.

**Bizonyítás.** Teljes indukcióval bizonyítjuk a

$$\|\tilde{y}_\ell - y_\ell\| \leq \frac{\zeta^\mu}{1 - 2^\kappa \zeta^\mu} c_0 h_\ell^\kappa \quad (15.115)$$

becslést,  $1 \leq \ell \leq \ell_{\max}$ . (15.115) teljesül a legalacsonyabb szinten, hiszen ott  $\tilde{y}_1 = y_1$ .

Az  $\ell$ -edik szinten legyen  $y_\ell^0 = \tilde{P}_{\ell-1}^\ell \tilde{y}_{\ell-1}$  a kiindulási közelítés. Ekkor a sima többrácsos módszer  $\mu$  lépése után kapjuk az  $y_\ell^\mu$  közelítést, amelyet ezen a szinten végső  $\tilde{y}_\ell$  megoldásnak tekintünk.  $\tilde{y}_\ell - y_\ell = M_\ell^\mu (y_\ell^0 - y_\ell)$  miatt

$$\|\tilde{y}_\ell - y_\ell\| \leq \|M_\ell^\mu\| \|y_\ell^0 - y_\ell\| \leq \zeta^\mu \|y_\ell^0 - y_\ell\|.$$

Tehát

$$\begin{aligned} \|y_\ell^0 - y_\ell\| &= \|\tilde{P}_{\ell-1}^\ell \tilde{y}_{\ell-1} - y_\ell\| \leq \|\tilde{P}_{\ell-1}^\ell y_{\ell-1} - y_\ell\| + \|\tilde{P}_{\ell-1}^\ell\| \|\tilde{y}_{\ell-1} - y_{\ell-1}\| \\ &\leq c_0 h_\ell^\kappa + \frac{\zeta^\mu}{1 - 2^\kappa \zeta^\mu} c_0 h_{\ell-1}^\kappa \\ &= c_0 h_\ell^\kappa \left( 1 + \frac{\zeta^\mu}{1 - 2^\kappa \zeta^\mu} 2^\kappa \right), \end{aligned}$$

hiszen  $h_{\ell-1} = 2h_\ell$ . Együtt az előbbi becsléssel :

$$\|\tilde{y}_\ell - y_\ell\| \leq c_0 h_\ell^\kappa \frac{\zeta^\mu}{1 - 2^\kappa \zeta^\mu}.$$

Ha most  $\zeta^\mu \leq \frac{1}{1+2^\kappa}$ , akkor a tétel állítása következik.  $\square$

**Megjegyzések. 1.** Gyakran  $\zeta < 0.2$ , ld. a 19. feladatot (valamint a 15.4.6. pontot). Ilyenkor, ha másodrendű a diszkretizáció, elég  $\mu = 1$ .

**2.** Tételünk bizonyítása szerint a teljes többrácsos módszer folyamán elegendő  $\mu = \ln(1 + 2^\kappa) / \ln \frac{1}{\zeta}$  iterációt végezni a mindenkor  $\ell$ -edik szinten, így az utolsó,  $\ell_{\max}$ -adik szinten is. A 15.4.4. pont eredményeivel együtt ez azt jelenti, hogy  $O(N_\ell)$  a teljes aritmetikai műveletigény, amellyel az  $\ell$ -edik szinten  $O(h_\ell^\kappa)$  hibával állíthatjuk elő az  $y_\ell$  megoldást,  $1 \leq \ell \leq \ell_{\max}$ .  $\square$

### 15.4.9 Nemlineáris egyenletek megoldása

Peremérték feladatunk, ill. diszkretizációja az utolsó szinten legyen  $L(u) = 0$ , ill.

$$A_\ell(y_\ell) = 0, \quad \ell = \ell_{\max}, \quad (15.116)$$

ahol  $A_\ell$  tehát nemlineáris  $\mathbb{R}^{N_\ell} \rightarrow \mathbb{R}^{N_\ell}$  leképezés.

A többrácsos megoldás restriktói és interpolációi lehetnek olyanok, mint a lineáris esetben. Mint simító iterációt vehetjük a nemlineáris Gauss–Seidel iterációt (ld. I. 6.6.). Ilyenkor, ha az  $\ell$ -edik szinten megoldandó  $A_\ell(y_\ell) = d_\ell$  diszkretizált egyenletrendszert egyszerűsítve az

$$F(y) = 0, \quad F : \mathbb{R}^{N_\ell} \rightarrow \mathbb{R}^{N_\ell},$$

alakban írjuk, akkor az  $i$ -edik lépésben az

$$F_i(y_1^{(i+1)}, \dots, y_{i-1}^{(i+1)}, y_i, y_{i+1}^{(i)}, \dots, y_n^{(i)}) = 0, \quad 1 \leq i \leq N_\ell \text{ (ciklikusan)},$$

$y_i$ -ben nemlineáris, egyváltozós egyenlet megoldásával foglalkozunk. Ezt egyetlen Newton-lépéssel intézzük el (a deriváltat differencia-képlettel helyettesítve).

A probléma a közelítő megoldás javítása a durvább rács segítségével. A legfelsőbb rácson adott a (15.116) egyenlet, annak jobboldala  $d_{\ell_{\max}} = 0$ . Legyen ismert az  $\ell$ -edik rács egyenlete a sima többrácsos módszer folyamán :

$$A_\ell(y_\ell) = d_\ell,$$

valamint  $y_\ell$  pontos megoldásának  $y_\ell^0$  jelenlegi közelítése. Az  $\ell - 1$ -edik rácson csak maga az  $A_{\ell-1}$  leképezés ismert.

A kérdés ekkor az olyan  $d_{\ell-1}$  kiszámítása, hogy az  $\ell-1$ -edik szint megoldásának segítségével az  $y_\ell^0$ -t javíthatjuk.

A. Brandt ötlete, hogy a durvább rácstra az  $y_\ell^0$  közelítő megoldást is vigyük át, és az  $\ell-1$ -edik szint egyenletének jobboldali vektorát a következőképpen határozzuk meg  $d_\ell$ -ből és a leszűkített  $y_\ell^0$ -ból :

$$d_{\ell-1} := A_{\ell-1}(\tilde{y}_{\ell-1}^0) - R_\ell^{\ell-1}(A_\ell(y_\ell^0) - d_\ell), \quad \text{ahol } \tilde{y}_{\ell-1}^0 := R_\ell^{\ell-1}y_\ell^0. \quad (15.117)$$

Ezután a durvább rácson megközelítjük az

$$A_{\ell-1}(y_{\ell-1}) = d_{\ell-1}$$

nemlineáris egyenlet megoldását. Ehhez a kezdeti közelítés  $\tilde{y}_{\ell-1}^0$ . Ezen szint végső közelítő megoldása legyen  $\tilde{y}_{\ell-1}$ . Amikor ezzel rendelkezünk, következhet a javítás az  $\ell$ -edik szinten :

$$y_\ell^1 := y_\ell^0 + P_{\ell-1}^\ell(\tilde{y}_{\ell-1} - \tilde{y}_{\ell-1}^0).$$

A (15.117) képletet magyarázza a következő gondolatsor. Az  $\ell$ -edik rácson kellene az  $A_\ell(y_\ell) = d_\ell$  egyenletet megoldanunk. A rendelkezésünkre álló  $y_\ell^0$  közelítéssel az egyenlet nem teljesül. Így keresünk olyan  $v_\ell$  javítást, amellyel igaz

$$A_\ell(y_\ell^0 + v_\ell) = d_\ell.$$

Ekkor, a Newton-módszer ötletét és az  $A_\ell$  nemlineáris leképezés  $A'_\ell$  Jacobi-mátrixát használva,

$$d_\ell = A_\ell(y_\ell^0 + v_\ell) \approx A_\ell(y_\ell^0) + A'_\ell v_\ell,$$

vagyis

$$A'_\ell v_\ell \approx d_\ell - A_\ell(y_\ell^0).$$

Így a  $v_\ell$  javítást definiáljuk mint az

$$A'_\ell v_\ell = d_\ell - A_\ell(y_\ell^0) \quad (15.118)$$

egyenlet megoldását. De  $v_\ell$ -t nem határozzuk meg pontosan, hanem a durva rács segítségével közelítjük : az  $R_\ell^{\ell-1}$  restrikióval a durva rácstra átvitt  $y_\ell^0$  közelítéshez, azaz  $\tilde{y}_{\ell-1}^0$ -hoz keresünk olyan  $v_{\ell-1}$  javítást, amellyel teljesül

$$A_{\ell-1}(\tilde{y}_{\ell-1}^0 + v_{\ell-1}) \approx A_{\ell-1}(\tilde{y}_{\ell-1}^0) + A'_{\ell-1}v_{\ell-1}. \quad (15.119)$$

A kézenfekvő ötlet ezután az, hogy (15.118) alapján itt az  $A'_{\ell-1}v_{\ell-1}$  tagot helyettesítjük a következőképpen :

$$A'_{\ell-1}v_{\ell-1} \approx R_\ell^{\ell-1}A'_\ell v_\ell \approx R_\ell^{\ell-1}(d_\ell - A_\ell(y_\ell^0)).$$

Ezért, figyelembe véve (15.119)-et is, a durvarácsú  $v_{\ell-1}$  javítást a következő egyenletből határozzuk meg :

$$A_{\ell-1}(w_{\ell-1}) = A_{\ell-1}(\tilde{y}_{\ell-1}^0) + R_{\ell}^{\ell-1}(d_{\ell} - A_{\ell}(y_{\ell}^0)) = d_{\ell-1},$$

ahol  $w_{\ell-1} := \tilde{y}_{\ell-1}^0 + v_{\ell-1}$  (és  $\tilde{y}_{\ell-1}^0 = R_{\ell}^{\ell-1}y_{\ell}^0$ ).

$w_{\ell-1}$  kiszámítása után a  $v_{\ell-1} = w_{\ell-1} - \tilde{y}_{\ell-1}^0$  javítást az  $\ell$ -edik szintre interpoláljuk :  $v_{\ell} := P_{\ell-1}^{\ell}(w_{\ell-1} - \tilde{y}_{\ell-1}^0)$ . Így közelítő Newton-módszert használunk anélkül, hogy deriváltakat vagy szokásos differencia-approximációjukat közvetlenül kellene kiszámítanunk!

A következőkben megadjuk az előbbi, sikeresnek bizonyult ötletnek megfelelő NMG eljárás programját, amelyet – úgy mint MG-t 15.4.3-ban – a sima, valamint a teljes többrácsos módszer keretében hasznosíthatjuk. Az eljárás feladata az  $A_{\ell}(y_{\ell}) = d_{\ell}$  nemlineáris algebrai rendszer megoldása; a legmagasabb szinten a fentieknek megfelelően  $d_{\ell} = 0$ .

procedure NMG( $\ell, y_{\ell}, d_{\ell}$ );

begin

if  $\ell=1$  then  $y_{\ell} := \Phi(y_{\ell}, d_{\ell})$  *pontos megoldás az első szinten*

else begin array  $r, w$ ;

$y_{\ell} := S_{\ell}^{\nu}(A_{\ell}, d_{\ell})y_{\ell}$ ; *elősimítás*

$r := A_{\ell}(y_{\ell}) - d_{\ell}$ ; *a maradékvektor kiszámítása*

$y_{\ell-1} := R_{\ell}^{\ell-1}y_{\ell}$ ;  $d_{\ell-1} := A_{\ell-1}(y_{\ell-1}) - R_{\ell}^{\ell-1}r$ ; *átmenet az  $\ell - 1$ -edik szintre*

$w := y_{\ell-1}$ ; *a durvarácsú megoldás kezdeti közelítése*

for  $i := 1$  step 1 until  $\gamma$  do NMG( $\ell-1, w, d_{\ell-1}$ );

$y_{\ell} := y_{\ell} + P_{\ell-1}^{\ell}(w - y_{\ell-1})$ ; *javítás az  $\ell$ -edik szinten*

$y_{\ell} := S_{\ell}^{\nu^2}(A_{\ell}, d_{\ell})y_{\ell}$ ; *utósimítás*

end

end;

A számítási tapasztalatok szerint ilyen többrácsos módszerrel a nemlineáris elliptikus peremérték feladatoknak jelentős részét lényegében olyan ráfordítással meg lehet oldani, mint a lineáris feladatokat (az  $N_{\ell}$  szorzója az  $\ell$ -edik szint  $O(N_{\ell})$  műveletigényében megnövekszik annyival, amennyi művelet/rácspontra szükséges az  $y_{\ell-1}$  és  $d_{\ell-1}$  kiszámításához).

Különösen fontos a sikerhez itt viszont az a  $\Phi(y_{\ell}, d_{\ell})$  eljárás, amely az elsőszintű egyenletek megoldását intézi: ha nem megfelelő (azaz : nem vagy alig csökkenti a közelítő megoldás maradékvektorának normáját), akkor exponenciális sebességgel divergálhat az egész algoritmus! Tehát a nemlineáris

peremérték feladatok lényeges nehézsége a legdurvább rácstra megy át. Amennyiben a legdurvább rácson több megoldást találunk, akkor már ott kell közülük válogatni (ami a kevés ismeretlen miatt könnyebb). Ezután a finomabb rácson rendszerint már csak egy megoldás vonzáskörében dolgozunk.

Az, hogy többértelmű a megoldás, a nemlineáris egyenletek megoldása során ki szokott derülni – és gyakran elsőnek éppen a nemkívánt megoldást kapjuk : nem azt, amit fizikailag lehet értelmezni, pl. negatív abszolút hőmérsékleteket vagy koncentrációkat stb. Ilyenkor nagyobb munka szükséges, a program kibővítendő *folytatásos módszerre* (v.ö. I. 6.4.8-cal, valamint pl. R.E. Bank és T.F. Chan cikkével), annak érdekében, hogy a feladatnak egy paraméterét (vagy a megoldásnak normáját vagy értékét egy rögzített pontban) variálva az összes megoldási utat bejárjuk.

#### 15.4.10 Befejező megjegyzések

Tisztázzuk először is, vajon mikor nem előnyös a többrácsos módszer?

Olyan feladatoknál, ahol a tartomány elég bonyolult (pl. apró lyukai vannak) és már az első szinten sok pont kellene, annyi, hogy további finomítás (ahol a többrácsos módszer be tudná bizonyítani az előnyeit) az adott számítógépen vagy már nem kivitelezhető, vagy már nem érdekes. Ekkor az adott rács a legfinomabb. Ebből a durvább rácsokat kell létrehozni, csökkenő szintszámmal párhuzamosan egyszerűsítve a tartományt. Ehhez 15.8.1-ben megadjuk *Galjorkin-reláció* címén az  $A_{\ell-1}$  mátrix kiszámítását  $A_{\ell}$ -ből. Ezen reláció használatával a nehézség az interpoláció megszerkesztésére tolódik át. A restriktio akkor már annak adjungáltja, ld. Adams cikkét is (véges differencia esetben ld. (15.90)). A bonyolult tartomány esetén lehet pl. a perem finom részleteit fokozatosan elhagyni a durvább rácsok létrehozásához, és azokon akkor az interpolációt definiálni. De ez nem egyszerű művelet. Ekkor a nemkonform rácsok és a quadtree algoritmus használata segíthet (ld. Gáspár Cs. dolgozatait).

Ilyen feladatoknál, ill. máskor, amikor a többrácsos módszer nem vezet elfogadható eredményekre (pl. időtől függő feladatok esetén amikor mozgó vonalak vannak a tartomány belsejében, amelyeken keresztül ugrásszerűen változnak az egyenlet együtthatói vagy megoldása), ill. amikor túl bonyolult nekünk a többrácsos módszer bevetése, akkor javasolható az prekondicionált konjugált gradiens módszer megfelelő változata (ld. I. 1.6.8., itt gyakran olyan változat szükséges, ahol nemszimmetrikus lehet a mátrix, ld. van der Vorst vagy Hegedűs munkáit). Mint prekondicionáló eljárás az inkomplett Gauss-eliminációt javasoljuk (I. 1.6.5.).

A fenti problémák mellett a többrácsos módszer programjának relatív bonyolultsága miatt van az a természetes jelenség, hogy kisméretű feladatoknál (kb. 200–500 ismeretlenig) az egyenletrendszerek megoldása direkt módszerrel vagy jobb iterációs eljárással (pl. konjugált gradiens módszer inkomplett Gauss–eliminációval kombinálva) gyorsabb.

A többrácsos módszernek viszont eddig még nem tárgyalt előnyei is vannak. Míg a simító iteráció általában a peremérték feladat egyszerű (alacsonyabbrendű, de stabil) approximációján alapszik, a módszerbe beépíteni magasabbrendű (sőt kevésbé stabil) approximációt nem nehéz : elegendő a maradékvektort magasabbrendű approximációnak megfelelően kiszámítani.

Egy lehetőség az  $A_h$  kiszámításának megválasztására az is, hogy a meglévő  $A_H$  durvarácsú approximációhoz, ugyanazon a durva rácson, magasabbrendű approximációt vesszünk. Hibabecslés céljából ugyanezt tehetünk a legmagasabb szinten. Például  $A_H$  legyen elsőrendű végeelem approximáció eredménye téglalapon ( $R_4$  bilineáris elem) ill. háromszögön ( $T_3$  lineáris elem), viszont az  $A_h$  kiszámításához másodrendű approximációt választunk, az előbbi elemekben felhasznált csomópontokon kívül a felezőpontokat is használva ( $R_8$  ill.  $T_6$ ).

Ezenkívül, mivel rendelkezünk a teljes többrácsos módszerben különböző finomságú diszkretizációkhoz tartozó megoldásokkal, kézenfekvő ezeknek olymódú kombinációja, hogy magasabbrendűbb approximációt kapjunk. Ezt szolgálja az I. 5.6-ban tárgyalt Richardson extrapoláció – amely peremérték feladatokra általánosítható, ld. Marcsuk és Sajdurov könyvét.

## 15.5 Differencia-approximációk; kiegészítések

### 15.5.1 A „diszkrét” Green-féle függvény becslése

Tekintsük a  $(0, a) \times (0, b)$  téglalapon a Poisson-egyenlet elsőfajú homogén peremérték feladatát és annak szokásos differencia-approximációját az  $\bar{\omega}_h$  rácson (ld. 15.3.1),

$$\begin{aligned} -(\Delta_h y)_{ij} &:= -(y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2}) = f(x_{ij}) =: f_{ij}, & x_{ij} \in \omega_h, \\ y(x_{ij}) &:= y_{ij} = 0, & x_{ij} \in \gamma_h. \end{aligned} \quad (15.120)$$

Az adott  $f_{ij}$  értékeket és a keresett  $y_{ij}$  ismeretleneket  $\vec{f}_h$ , ill.  $y_h$  vektorokba rendezve, pl. sorfolytonosan, (15.120)-ból kapjuk a homogén peremértékek eliminációja után az

$$A_h^0 y_h = \vec{f}_h \quad (15.121)$$

$n$ -dimenziós egyenletrendszert ( $n := (N_1 - 1)(N_2 - 1)$ ), amelynek (korábban  $A_h$ -val jelölt)  $A_h^0$  mátrixáról tudjuk, hogy az szimmetrikus és pozitív definit

M-mátrix. Így inverze létezik és nemnegatív (azaz minden eleme nemnegatív – sőt, ebben a speciális esetben : minden eleme pozitív).

A diszkrét Green-függvény

$$G_h^0 := \frac{1}{h_1 h_2} (A_h^0)^{-1}, \quad (15.122)$$

tehát (az eredeti  $i, j$ -indexelésnél maradva), a (15.120) feladat megoldása felírható mint

$$y_{ij} = \sum_{k=1}^{N_1-1} \sum_{\ell=1}^{N_2-1} G_{h,i,j,k\ell}^0 f_{k\ell} h_1 h_2.$$

Célszerűbb ezt indexmentesen felírni :

$$y(x) = \sum_{\xi \in \omega_h} G_h^0(x, \xi) f(\xi) h_1 h_2, \quad x \in \omega_h, \quad (15.123)$$

mert ez takarékosabb és jobban megmutatja az analógiát a  $\Delta u + f(x) = 0$ ,  $x \in \Omega$ ;  $u(x) = 0$ ,  $x \in \Gamma$  feladattal és annak

$$u(x) = \int_{\Omega} G^0(x, \xi) f(\xi) d\xi$$

alakú megoldásával, ahol  $G^0$  a „folytonos” Green-függvény.

Célunk most a (15.122) „diszkrét” Green-függvény becslése, mert ezzel – és (15.123) segítségével a (15.120) feladat megoldása is becsülhető. Ezenkívül az itteni eredményeknek a hasznát még a végeelem módszer, valamint parabolikus és hiperbolikus egyenletek tárgyalása során is látjuk majd.

Nem várhatjuk azt, hogy  $G_h^0$ -ra  $h_1$ -ben és  $h_2$ -ben egyenletes becslést nyerhetünk, mert a folytonos feladat Green-függvénye  $x = \xi$ -nél logaritmikus szingularitással rendelkezik.

**15.19. Lemma** (a Laplace-operátor diszkrét Green-féle függvényének becslése). A (15.122)-vel definiált Green-féle függvény eleget tesz a következő becslésnek :

$$0 \leq G_h^0(x, \xi) \leq \gamma_0 + \gamma_1 \ln(\sqrt{2} \max(N_1, N_2)), \quad x, \xi \in \bar{\omega}_h, \quad (15.124)$$

ahol

$$\gamma_0 := \frac{ab}{2(a^2 + b^2)}, \quad \gamma_1 := \frac{\pi}{2}.$$

Amennyiben  $h_1 \leq 1/(2a^2)$ ,  $h_2 \leq 1/(2b^2)$ , akkor érvényes

$$0 \leq G_h^0(x, \xi) \leq \gamma_2 \ln \frac{1}{\min(h_1, h_2)}, \quad \gamma_2 := \frac{3}{2}(\gamma_0 + \gamma_1). \quad (15.125)$$



**Bizonyítás.** Először azt a feladatot írjuk fel, amelynek megoldása a diszkrét Green-függvény, majd ezt pontosan megoldjuk (diszkrét Fourier-analízissel). Onnan levezethetjük a kívánt becslést (felhasználva a diszkrét sajátérték feladat pontos megoldását).

Mivel az  $(A_h^0)^{-1}$  mátrix  $j$ -edik oszlopát úgy kapjuk meg, hogy a (15.121) egyenletrendszer jobboldala gyanánt a  $j$ -edik koordináta egységvektort veszünk, a  $G_h^0(x, \xi)$  függvény rögzített  $\xi = x_{k\ell} \in \omega_h$  esetén megoldása a következő feladatnak :

$$-(\Delta_h G_h^0)(x, \xi) = \frac{\delta(x, \xi)}{h_1 h_2}, \quad x \in \omega_h; \quad G_h^0(x, \xi) = 0, \quad x \in \gamma_h. \quad (15.126)$$

Itt

$$\delta(x, \xi) = \delta(x_{ij}, x_{k\ell}) = \delta_{ik} \cdot \delta_{j\ell}$$

két Kronecker-szimbólum szorzata. (15.126) lineáris algebrai alakja éppen

$$A_h^0 (A_h^0)^{-1} = I_h.$$

A diszkrét  $\Delta$ -operátor sajátérték feladatának megoldását a (15.41), (15.42) képletek adják 15.3.3-ban, tehát a  $v^{(k\ell)}$  normált „diszkrét” sajátfüggvényeket és a  $\lambda_{k\ell}^h$  sajátértékeket. Az ott (15.43)-mal bevezetett  $L_2(\omega_h)$ -féle skalárszorzatot is használjuk a következőkben. Ennek segítségével a (15.120) feladat (15.123) megoldását azonnal rövidebben írhatjuk fel :

$$y(x) = (G_h^0(x, \cdot) f)_{(0,h)}, \quad x \in \omega_h.$$

Sorba fejtve a sajátfüggvények szerint az  $\vec{f}_h$  vektort (azaz az  $f$  rácsfüggvényt,  $f(x_{ij}) = f_{ij}$ ) és a (15.120) feladat  $y$  megoldását, azt kapjuk  $x \in \omega_h$ -ra, hogy

$$f(x) = \sum_{k,\ell} (f, v^{(k\ell)})_{(0,h)} v^{(k\ell)}(x), \quad y(x) = \sum_{k,\ell} c_{k\ell} v^{(k\ell)}(x), \quad (15.127)$$

ahol a  $c_{k\ell}$  együtthatókat (15.120)-ból határozzuk meg :

$$\begin{aligned} 0 &= (\Delta_h y + f, v^{(ij)})_{(0,h)} \\ &= \sum_{x \in \omega_h} \left\{ \sum_{k,\ell} [-c_{k\ell} \lambda_{k\ell}^h + (f, v^{(k\ell)})_{(0,h)}] v^{(k\ell)}(x) \right\} v^{(ij)}(x) h_1 h_2 \\ &= \sum_{k,\ell} [-c_{k\ell} \lambda_{k\ell}^h + (f, v^{(k\ell)})_{(0,h)}] \delta_{ik} \delta_{j\ell} \\ &= -c_{ij} \lambda_{ij}^h + (f, v^{(ij)})_{(0,h)}. \end{aligned}$$

Innen és (15.127)-ből következik

$$y(x) = \sum_{k,\ell} \frac{(f, v^{(k\ell)})_{(0,h)}}{\lambda_{k\ell}^h} v^{(k\ell)}(x).$$

Ezt a képletet alkalmazzuk a  $G_h^0(x, \xi)$  leírására, v.ö. (15.120)-at és (15.126)-ot :

$$G_h^0(x, \xi) = \sum_{k,\ell} \frac{g^{k\ell}(\xi)}{\lambda_{k\ell}^h} v^{(k\ell)}(x),$$

ahol

$$g^{k\ell}(\xi) := \left( \frac{\delta(\cdot, \xi)}{h_1 h_2}, v^{(k\ell)} \right)_{(0,h)} = \sum_{x \in \omega_h} \delta(x, \xi) v^{(k\ell)}(x) = v^{(k\ell)}(\xi).$$

Vagyis érvényes a következő diszkrét Hilbert–Schmidt-féle magképlet :

$$G_h^0(x, \xi) = \sum_{k,\ell} \frac{v^{(k\ell)}(x) v^{(k\ell)}(\xi)}{\lambda_{k\ell}^h}. \quad (15.128)$$

Ez a képlet is példa arra, hogy a rácsfüggvényekre támaszkodó felírásmód célszerűbb lehet a vektorokat ill. mátrixokat használó felírásmódnál (ld. az 5. feladatot is). (15.128)-ból megegyeszer látjuk, hogy a Green-függvény eleget tesz a peremfeltételeknek és szimmetrikus. Mivel (15.41) alapján  $|v^{(k\ell)}(x)| \leq 2/\sqrt{ab}$  és  $G_h^0$  nemnegatív, így (15.128)-ból azonnal következik

$$0 \leq G_h^0(x, \xi) \leq \frac{4}{ab} \sum_{k,\ell} \frac{1}{\lambda_{k\ell}^h}.$$

Feladatunk ezután a sajátértékek alsó becslése. Feltettük, hogy  $N_1, N_2 \geq 2$  (mert máskülönben  $\omega_h$  üres lenne), azaz  $0 < \frac{\pi h_1}{2a}, \frac{\pi h_2}{2b} \leq \frac{\pi}{4}$ . Így a  $\lambda_{1,1}^h$  sajátérték (15.44) becslése 15.3.3-ból alkalmazható.

Általában viszont  $0 < \frac{k\pi h_1}{2a}, \frac{\ell\pi h_2}{2b} < \frac{\pi}{2}$ , és emiatt  $\sin x \geq \frac{2}{\pi}x$  segítségével következik

$$\lambda_{k\ell}^h \geq 4 \left[ \left( \frac{k}{a} \right)^2 + \left( \frac{\ell}{b} \right)^2 \right].$$

Összegezve:

$$G_h^0(x, \xi) \leq \frac{4}{ab} \left( \frac{a^2 b^2}{8(a^2 + b^2)} + \frac{1}{4} \sum_{k,\ell} \frac{1}{\left( \frac{k}{a} \right)^2 + \left( \frac{\ell}{b} \right)^2} \right), \quad (15.129)$$

ahol a duplaösszeg az  $1 \leq k \leq N_1 - 1$ ,  $1 \leq \ell \leq N_2 - 1$ ,  $k + \ell > 2$  indexekre terjed ki. Itt legyen

$$\left(\frac{k}{a}\right)^2 + \left(\frac{\ell}{b}\right)^2 = r^2 \left[ \frac{1}{a^2} \cos^2 \varphi + \frac{1}{b^2} \sin^2 \varphi \right] =: g(r, \varphi),$$

$$r^2 := k^2 + \ell^2, \quad \cos^2 \varphi := \frac{k^2}{k^2 + \ell^2}, \quad \sin^2 \varphi := \frac{\ell^2}{k^2 + \ell^2}.$$

Ezzel az egészszámú  $k, \ell$  értékekre definiált  $\left(\frac{k}{a}\right)^2 + \left(\frac{\ell}{b}\right)^2$  kifejezést minden  $r \geq 0$ ,  $0 \leq \varphi \leq \frac{\pi}{2}$  értékre terjesztettük ki.

Tekintsük a  $g(r, \varphi)$  függvényt az egészszámú  $(k, \ell)$ -rácsnak egy  $s_{k\ell} := \{k-1 \leq x_1 \leq k; \ell-1 \leq x_2 \leq \ell\}$  négyzete felett. Mivel  $\Delta_{r, \varphi} g = \frac{2}{a^2} + \frac{2}{b^2} > 0$ , így  $g(r, \varphi)$  a maximumelv szerint maximumát a négyzet peremén veszi fel, pontosabban : annak jobb felső sarkában. Ott viszont értéke  $\left(\frac{k}{a}\right)^2 + \left(\frac{\ell}{b}\right)^2$ . Ennek alapján

$$\frac{1}{\left(\frac{k}{a}\right)^2 + \left(\frac{\ell}{b}\right)^2} \leq \int_{s_{k\ell}} \frac{dx}{g(r(x_1, x_2), \varphi(x_1, x_2))},$$

és (15.129)-ből következik

$$G_h^0(x, \xi) \leq \frac{1}{ab} \left\{ \frac{a^2 b^2}{2(a^2 + b^2)} + \int_0^{\frac{\pi}{2}} \int_1^{\sqrt{2} \max(N_1, N_2)} \frac{r \, dr \, d\varphi}{g(r, \varphi)} \right\}.$$

Egy formulamanipulációs program segítségével (valamint elemi mértani ismeretek felhasználásával) kapjuk azt, hogy

$$\int_0^{\frac{\pi}{2}} \frac{d\varphi}{\frac{1}{a^2} \cos^2 \varphi + \frac{1}{b^2} \sin^2 \varphi} = \frac{\pi ab}{2};$$

ezenkívül világos, hogy

$$\int_1^{\sqrt{2} \max(N_1, N_2)} \frac{dr}{r} = \ln(\sqrt{2} \max(N_1, N_2)).$$

Ezzel bebizonyítottuk a (15.124) becslét, amelynek (15.125) már a következménye :

Mivel  $N_1, N_2 \geq 2$ , így

$$1 < \ln 2\sqrt{2} \leq \ln(\sqrt{2} \max(N_1, N_2)) \leq \ln d,$$

ahol  $d := \sqrt{2} \max(a, b) / \min(h_1, h_2)$ , tehát

$$\gamma_0 + \gamma_1 \ln(\sqrt{2} \max(N_1, N_2)) \leq (\gamma_0 + \gamma_1) \ln d.$$

Most használjuk azt, hogy  $h_1 \leq 1/(2a^2)$ ,  $h_2 \leq 1/(2b^2)$  esetén

$$2 \max(a^2, b^2) \leq 1/\min(h_1, h_2) \text{ és így } d \leq 1/\min(h_1, h_2)^{3/2},$$

és a lemma be van bizonyítva.  $\square$

Egy, a 15.19. lemma (15.124) becslését pontosító egyenlőtlenség bizonyítása a 29. feladat.

**Megjegyzés.** Numerikus kísérletek szerint az  $a = b = 1$ ,  $h_1 = h_2 = h$  esetben igaz a lemma (15.124) becslése pl. a  $\gamma_0 = 0$ ,  $\gamma_1 = \frac{2}{\pi}$  konstansokkal.  $\square$

A Green-féle függvény becslései általánosíthatók a változó együtthatójú egyenletek esetére, amit a következő pontban mutatunk meg.

### 15.5.2 Változó együtthatójú differenciáloperátorok

Most foglalkozunk az alábbi általánosabb elliptikus egyenlet peremérték feladatával, amely tartalmazza pl. a  $\operatorname{div}(k \operatorname{grad} u) + f = 0$  alakú kétdimenziós, változó együtthatójú hővezetési egyenletet mint speciális esetet, ha  $k_1 = k_2 = k$ . Az  $\Omega$  tartományt az egyszerűség kedvéért téglalapnak feltételezzük:  $\Omega := \{0 < x_1 < a, 0 < x_2 < b\}$ ; ehelyett téglalapokból összeállított  $\Omega$ -t is lehetne vizsgálni.

$$\mathcal{L}u + f = 0, \quad x \in \Omega; \quad u(x) = g(x), \quad x \in \Gamma, \quad (15.130)$$

$$\mathcal{L}u := \frac{\partial}{\partial x_1} \left( k_1(x) \frac{\partial u}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left( k_2(x) \frac{\partial u}{\partial x_2} \right) - du, \quad (15.131)$$

feltéve, hogy teljesül

$$k_1(x), k_2(x) \geq k_0 > 0, \quad d(x) \geq 0, \quad x \in \Omega. \quad (15.132)$$

Ezen feladat diszkrét megfelelője (v.ö. II. 11.4.8-cal)

$$\Lambda y = f, \quad x = (x_{1i}, x_{2j}) \in \omega_h; \quad y = g, \quad x \in \gamma_h, \quad (15.133)$$

ahol  $\Lambda$  a következő differencia-operátor :

$$\Lambda y := -((k_1 y_{x_1})_{\tilde{x}_1} + (k_2 y_{x_2})_{\tilde{x}_2}) + dy. \quad (15.134)$$

(Korábbi egyezményünk szerint differencia-kifejezésekben használjuk az  $y$  rácsfüggvényt, amelyet vektorként  $y_h$ -nak írjuk.) A (15.133)-(15.134) séma mátrixalakja a peremfeltételek eliminációja után legyen

$$A_h y_h = \varphi_h, \quad (15.135)$$

ahol  $\varphi_h$  az  $f$  és  $g$  értékeiből összeállított oszlopvektor (a 15.3.3. pontban ezt  $k_1 = k_2 = 1$  esetén részletesen felírtuk (15.46) kapcsán).

Elsőnek foglalkozunk a (15.133)-(15.134) séma stabilitásával, azaz megmutatjuk, hogy  $A_h$  reguláris és inverzének normája  $h$ -tól függetlenül korlátos. Ehhez egyrészt a (15.43)-féle skalárszorzatot használjuk, mégpedig akár  $y, z$  rácsfüggvényekre, akár a megfelelő  $y_h, z_h$  oszlopvektorokra (a rács belső pontjait pl. sorfolytonosan végigszámozva) :

$$(y, z)_{(0,h)} = (y_h, z_h)_{(0,h)}.$$

Másrészt a következő skalárszorzatokat vezetjük be a peremre nullával folytatott  $y, z$  rácsfüggvények differenciahányadosaira :

$$\begin{aligned} (y_{x_1}, z_{x_1}]_1 &:= \sum_{(x_{1i}, x_{2j}) \in \omega_h^1} (y_{x_1} z_{x_1})_{i+1/2, j} h_1 h_2, \\ (y_{x_2}, z_{x_2}]_2 &:= \sum_{(x_{1i}, x_{2j}) \in \omega_h^2} (y_{x_2} z_{x_2})_{i, j+1/2} h_1 h_2, \\ (y, z)_{H_0^1(\omega_h)} = (y, z)_{(1,h)} &:= (y_{x_1}, z_{x_1}]_1 + (y_{x_2}, z_{x_2}]_2. \end{aligned}$$

Ezen mennyiségek kiszámítása éppen azt jelenti, hogy az összegzéseket minden olyan  $\bar{\omega}_h$ -beli rácspontra ki kell kiterjesztenünk, ahol  $y_{x_k}$  és  $z_{x_k}$  definiált,  $k = 1, 2$  :

$$\begin{aligned} \omega_h^1 &:= \{x = (x_{1i}, x_{2j}) \in \bar{\omega}_h; 0 \leq x_{1i} \leq 1 - h_1, h_2 \leq x_{2j} \leq 1 - h_2\}, \\ \omega_h^2 &:= \{x = (x_{1i}, x_{2j}) \in \bar{\omega}_h; h_1 \leq x_{1i} \leq 1 - h_1, 0 \leq x_{2j} \leq 1 - h_2\}. \end{aligned}$$

Ezekhez a definíciókhoz ld. a 30. feladatot. A fenti  $H_0^1(\omega_h)$ -féle skalárszorzat-hoz tartozó normát  $|\cdot|_{(1,h)}$ -val jelöljük, két tagját viszont  $\|y_{x_k}\|_k^2$ -val (tehát ezeknél az alsó  $k$  index az irányra mutat rá,  $k = 1, 2$ ) :

$$|y_h|_{(1,h)}^2 = |y|_{(1,h)}^2 := \|y_{x_1}\|_1^2 + \|y_{x_2}\|_2^2.$$

**15.20. Tétel** (önadjungált elliptikus egyenletet approximáló differenciaséma stabilitása). A (15.135) differenciaséma stabil az  $|\cdot|_{(1,h)}$  normában; érvényes a következő becslés :

$$|y_h|_{(1,h)} \leq \frac{ab}{2k_0 \sqrt{2(a^2 + b^2)}} \|\varphi_h\|_{(0,h)}.$$

ahol  $k_0$  a (15.132)-ben szereplő konstans.

**Bizonyítás.** A differenciaséma (15.135) egyenletét skalárisan megszorozzuk  $y_h$ -val, majd alkalmazzuk a Cauchy–egyenlőtlenséget :

$$(A_h y_h, y_h)_{(0,h)} = (\varphi_h, y_h)_{(0,h)} \leq \|\varphi_h\|_{(0,h)} \|y_h\|_{(0,h)}. \quad (15.136)$$

Az  $(A_h y_h, y_h)_{(0,h)}$  kifejezés alsó becsléséhez kényelmesebb a differencia-jelölést és a parciális összegzést alkalmazni :

$$\begin{aligned} (A_h y_h, y_h)_{(0,h)} &= (\Lambda y, y)_{(0,h)} \\ &= (k_1 y_{x_1}, y_{x_1})_1 + (k_2 y_{x_2}, y_{x_2})_2 + (dy, y)_{(0,h)} \quad (15.137) \\ &\geq k_0 \{ \|y_{x_1}\|_1^2 + \|y_{x_2}\|_2^2 \} = k_0 |y_h|_{(1,h)}^2. \end{aligned}$$

Vegyük észre, hogy (15.137)-ben a

$$(A_h y_h, y_h)_{(0,h)} \geq k_0 (A_h^0 y_h, y_h)_{(0,h)} \quad (15.138)$$

becslést kaptuk, ahol  $A_h^0$  a diszkrét Poisson–egyenlet (15.121) mátrixa, hiszen

$$\|y_{x_1}\|_1^2 + \|y_{x_2}\|_2^2 = (A_h^0 y_h, y_h)_{(0,h)}.$$

Ezután azt használjuk, hogy

$$|y_h|_{(1,h)}^2 = (A_h^0 y_h, y_h)_{(0,h)} \geq \lambda_{\min}^h(A_h^0) \|y_h\|_{(0,h)}^2,$$

Ide  $\lambda_{\min}^h$  (15.44) becslését behelyettesítjük, majd az eredmény négyzetgyökét vesszük :

$$\frac{2}{ab} \sqrt{2(a^2 + b^2)} \|y_h\|_{(0,h)} \leq |y_h|_{(1,h)}.$$

Ezt az egyenlőtlenséget  $k_0 |y_h|_{(1,h)}$ -val megszorozva, folytathatjuk a becslést (15.137) és (15.136) segítségével :

$$\frac{2k_0}{ab} \sqrt{2(a^2 + b^2)} \|y_h\|_{(0,h)} |y_h|_{(1,h)} \leq k_0 |y_h|_{(1,h)}^2 \leq \|\varphi_h\|_{(0,h)} \|y_h\|_{(0,h)},$$

és innen a tétel állítását kapjuk.  $\square$

**Megjegyzések.** **1.** Ha a parciális összegzést nem  $(A_h y_h, y_h)_{(0,h)}$ -ra hanem  $(A_h y_h, z_h)_{(0,h)}$ -ra alkalmazzuk, akkor a (15.137)-nek megfelelő úton az  $A_h$  önadjungáltságot a (15.43) skalárszorzatban kapjuk.

**2.** A (15.133)-(15.134) séma  $\psi_h$  képlethibája az adatok megfelelő simasága esetén másodrendű (ld. a 31. feladatot), így a tételből az  $y_h$  megoldás konvergenciája is következik. Ekkor (15.135)-ben  $y_h$  helyett  $y_h - \vec{u}_h$  áll,  $\varphi_h$  helyett  $\psi_h$  (és annak részeként  $\vec{f}_h$ , a peremértékek zérusok).

3. (15.138) és (15.44) azt adják, hogy

$$\|A_h^{-1}\|_{(0,h)} \leq \frac{1}{k_0 \lambda_{\min}^h(A_h^0)} \leq \frac{a^2 b^2}{8k_0(a^2 + b^2)},$$

és mivel a  $(\cdot, \cdot)_{(2)}$ -vel jelölt euklideszi skalárszorzat csak a  $h_1 h_2$  szorzóval különbözik az  $L_2(\omega_h)$ -féle skalárszorzattól,  $h_1 h_2(u, v)_{(2)} = (u, v)_{(0,h)}$ , a kapott korlát egyben  $\|A_h^{-1}\|_{(2)}$  korlátja is, azaz

$$\lambda_{\min}^h(A_h) \geq k_0 \lambda_{\min}^h(A_h^0). \quad \square$$

Most a Green-függvény vizsgálatára térünk rá. A (15.133)-(15.134) feladat Green-függvényét  $G_h(x, \xi)$ -vel jelöljük, tehát rögzített  $\xi \in \omega_h$ -ra legyen

$$(\Delta G_h)(x, \xi) = \frac{\delta(x, \xi)}{h_1 h_2}, \quad x \in \omega_h; \quad G_h(x, \xi) = 0, \quad x \in \gamma_h. \quad (15.139)$$

A Green-függvény összefüggése az  $A_h^{-1}$  mátrixszal a szokásos. Mindkét mátrixnak a  $(g_\ell$ -l, ill.  $\alpha_\ell$ -l jelölt) oszlopaira lesz szükségünk lejjebb a 15.19. lemmát általánosító becslésekben :

$$(\alpha_1, \dots, \alpha_n) = A_h^{-1} = h_1 h_2 G_h = h_1 h_2 (g_1, \dots, g_n). \quad (15.140)$$

Hasonlóan az  $(\alpha_1^0, \dots, \alpha_n^0) = (A_h^0)^{-1} = h_1 h_2 (g_1^0, \dots, g_n^0)$  oszlopvektorokat is vezetjük be és az egységmátrix oszlopaikat  $e_\ell$ -l jelöljük.

**15.21. Tétel** (önadjungált elliptikus egyenlet diszkrét Green-függvényének becslése). Ha igaz (15.132), akkor a (15.139) által definiált Green-függvény teljesíti a következő becslést :

$$0 \leq G_h(x, \xi) \leq \frac{1}{k_0} G_h^0(x, \xi) \leq \frac{\gamma_2}{k_0} \ln \frac{1}{\min(h_1, h_2)},$$

ahol  $\gamma_2$  a 15.19. lemmában definiált konstans.

**Bizonyítás.** Rögzített  $\xi = x^\ell \in \omega_h$ -ra legyen  $g_\ell(x) := G_h(x, \xi)$  és  $g_\ell^0(x) := G_h^0(x, \xi)$ ,  $x \in \omega_h$ . Ha  $G_h$ -t nem kétargumentumú rácsfüggvénynek tekintjük, hanem mátrixnak, akkor  $g_\ell$  annak fent már bevezetett oszlopa és  $g_\ell(x)$  annak egy komponense (analog  $g_{\ell\ell}^0$  értelmezése). Így pl.  $g_\ell$   $\ell$ -edik komponense

$$g_{\ell\ell} = g_\ell(x^\ell) = G_h(x^\ell, x^\ell) = \frac{1}{h_1 h_2} (A_h^{-1})_{\ell\ell}.$$

Az előző tételből tudjuk, hogy

$$(A_h g_\ell, g_\ell)_{(0,h)} \geq k_0 \|g_\ell\|_{(1,h)}.$$

Ezután használjuk, hogy

$$A_h^0 \alpha_\ell^0 = e_\ell = A_h^0 h_1 h_2 g_\ell^0,$$

tehát

$$\begin{aligned} g_{\ell\ell}^2 &= (e_\ell, g_\ell)_{(2)}^2 = (A_h^0 h_1 h_2 g_\ell^0, g_\ell)_{(2)}^2 \\ &= (A_h^0 g_\ell^0, g_\ell)_{(0,h)}^2 \leq (A_h^0 g_\ell^0, g_\ell^0)_{(0,h)} (A_h^0 g_\ell, g_\ell)_{(0,h)}, \end{aligned} \quad (15.141)$$

ahol arra való hivatkozással, hogy  $A_h^0$  szimmetrikus és pozitív definit (ld. az előző tételt), az általánosított Schwarz-egyenlőtlenséget alkalmaztuk :

$$(Ax, y)^2 \leq (Ax, x)(Ay, y).$$

Ez akkor érvényes, ha  $A$  önadjungált és pozitív definit a  $(\cdot, \cdot)$  skalárszorzatban. Ilyenkor ugyanis önadjungált és pozitív definit  $A^{1/2}$  gyöke létezik, és így a Schwarz-egyenlőtlenségből következik

$$|(Ax, y)| = |(A^{1/2}x, A^{1/2}y)| \leq \|A^{1/2}x\| \|A^{1/2}y\| = ((Ax, x)(Ay, y))^{1/2},$$

hiszen  $\|A^{1/2}x\|^2 = (A^{1/2}x, A^{1/2}x) = (Ax, x)$ .

Mivel

$$(A_h^0 g_\ell^0, g_\ell^0)_{(0,h)} = \frac{1}{h_1 h_2} (e_\ell, g_\ell^0)_{(0,h)} = (e_\ell, g_\ell^0)_{(2)}$$

és

$$(A_h^0 g_\ell, g_\ell)_{(0,h)} \leq \frac{1}{k_0} (A_h g_\ell, g_\ell)_{(0,h)},$$

így, folytatva a (15.141) egyenlőtlenséget azt kapjuk, hogy

$$\begin{aligned} g_{\ell\ell}^2 &\leq \frac{1}{k_0} (e_\ell, g_\ell^0)_{(2)} (A_h g_\ell, g_\ell)_{(0,h)} \\ &= g_{\ell\ell}^0 \frac{1}{k_0} (e_\ell, g_\ell)_{(2)} = g_{\ell\ell}^0 \frac{1}{k_0} g_{\ell\ell}. \end{aligned}$$

Itt  $g_{\ell\ell} > 0$  ismert, hiszen  $A_h$ -val együtt  $A_h^{-1}$  és  $G_h$  is szimmetrikusak és pozitív definiték. Ezért

$$0 < g_{\ell\ell} \leq \frac{1}{k_0} g_{\ell\ell}^0. \quad (15.142)$$

Valójában minden  $j$  indexre  $0 \leq g_{j\ell}$ , mert  $A_h$  M-mátrix (az előjel-feltételek (15.132) mellett azonnal ellenőrizhetőek, és  $A_h$  szimmetrikus és pozitív definit). Akkor a maximumelv is alkalmazható, amely szerint  $\alpha_\ell = h_1 h_2 g_\ell$ , az  $A_h \alpha = e_\ell \geq 0$  megoldása, nemnegatív vektor és felveszi maximumát ott, ahol  $e_\ell$ , azaz

$$0 \leq g_{j\ell} \leq g_{\ell\ell} \quad \text{minden } j\text{-re.}$$



Most térjünk vissza a Green-féle függvényhez! Ekkor ez utóbbi egyenlőtlenséget és (15.142)-t kombinálva azt kapjuk, hogy

$$0 \leq G_h(x, \xi) \leq G_h(\xi, \xi) \leq \frac{1}{k_0} G_h^0(\xi, \xi), \quad \text{minden } x, \xi \in \omega_h\text{-ra.}$$

$G_h$  szimmetriája miatt az is igaz, hogy

$$0 \leq G_h(x, \xi) \leq G_h(x, x).$$

A 15.19. lemmával együtt ez adja a tétel bizonyítását.  $\square$

Még általánosabb elliptikus operátorok differencia-approximációja a 32. feladat. Polár- és gömbkoordinátákban megadott feladatokkal a parabolikus egyenletek tárgyalásakor foglalkozunk.

### 15.5.3 „Diszkrét” beágyazási tételek, maximumnormabeli konvergencia

Továbbra is csak a kétdimenziós esetet vizsgálva, itt a 15.19. lemmából és a 15.20. tételből vonunk le néhány következtetést.

**15.22. Lemma** („tilos” beágyazási tétel). Tetszőleges, az  $\bar{\omega}_h$  rács peremén eltűnő  $y$  rácsfüggvényre teljesül

$$\|y\|_{C(\omega_h)} \leq \left( \gamma_2 \ln \frac{1}{\min(h_1, h_2)} \right)^{1/2} \|y\|_{H_0^1(\omega_h)}. \quad (15.143)$$

Itt  $\gamma_2$  a (15.125)-ben definiált konstans.

**Bizonyítás.** Tekintsük a (15.120) feladat  $y$  megoldását! Ez  $f$ -fel együtt befutja az egész  $\mathbb{R}^n$ -et, mivel (15.121)-ben  $A_h^0$  reguláris. Ennek és (15.123) alapján tetszőleges  $y$  rácsfüggvényre igaz a következő azonosság :

$$y(x) = \sum_{\xi \in \omega_h} G_h^0(x, \xi) (-\Delta_h y)(\xi) h_1 h_2 = (G_h^0(x, \cdot), -\Delta_h y)_{(0,h)}.$$

Innen az általánosított Schwarz-egyenlőtlenséggel kapjuk

$$\begin{aligned} |y(x)| &= |(G_h^0(x, \cdot), -\Delta_h y)_{(0,h)}| \\ &\leq ((-\Delta_h G_h^0)(x, \cdot), G_h^0(x, \cdot))_{(0,h)}^{1/2} (-\Delta_h y, y)_{(0,h)}^{1/2} \\ &= (G_h^0(x, x))^{1/2} \|y\|_{H_0^1(\omega_h)}, \end{aligned}$$

és most a 15.19. lemma adja az állításunkat.  $\square$

**Megjegyzés.** A (15.143) egyenlőtlenséget azért hívhatjuk beágyazási tételnek, mert azt mondja ki, hogy a  $H_0^1(\omega_h)$ -beli rácsfüggvények egyben  $C(\omega_h)$ -nak is elemei. A beágyazási tétel viszont „tilos” abban az értelemben, hogy a beágyazási konstans  $h$ -től függ, azaz a rácsfüggvények terének dimenziójától. Ez azzal van összhangban, hogy a folytonos argumentumú függvények körében (amikor  $\Omega \in \mathbb{R}^2$ ) ilyen tétel nem igaz (végtelen dimenziójú térben korántsem minden norma ekvivalens).  $\square$

A (15.143) becslés igen hasznos. Legyen ugyanis a téglalapon adott egy elliptikus peremérték feladat, amelynek  $u$  megoldására és  $y_h$  differenciaközelítésére ismerünk a  $H_0^1(\omega_h)$ -normában egy konvergencia becslést (mert azt könnyebben állíthatjuk elő), mondjuk

$$\|y_h - \vec{u}_h\|_{H_0^1(\omega_h)} \leq Mh^2$$

alakban (legyen  $h_1 = h_2 = h$ ). Ekkor ebből (15.143) segítségével rögtön kapjuk az egyenletes konvergencia becslést is:

$$\|y_h - \vec{u}_h\|_{C(\omega_h)} \leq M \left( \gamma_2 \ln \frac{1}{h} \right)^{1/2} h^2.$$

Ebben – a szokásos lépéstávolságok esetén – a  $(\gamma_2 \ln \frac{1}{h})^{1/2}$  szorzó még nagy számnak sem mondható:  $a = b = 1$  esetén  $\gamma_2 < \frac{5}{2}$ , és ha most  $h = 5 \cdot 10^{-18}$ , akkor  $(\frac{5}{2} \ln \frac{1}{h})^{1/2}$  még mindig 10 alatt van.

Alkalmazzuk ezt a gondolatot a (15.130)-(15.131), (15.132) feladatra és (15.133)-(15.134) differencia-approximációjára!

**15.23. Lemma** (változó együtthatójú feladat megoldásának maximum-normabeli stabilitása). A (15.135) feladat  $y_h$  megoldására igaz a következő becslés:

$$\|y_h\|_{C(\omega_h)} \leq \frac{1}{k_0} \left( \gamma_2 \ln \frac{1}{\min(h_1, h_2)} \right)^{1/2} \|\varphi_h\|_{H_0^{-1}(\omega_h)},$$

ahol  $\gamma_2$  a (15.125) konstans és

$$\|\varphi_h\|_{H_0^{-1}(\omega_h)} := \sup_{0 \neq v \in H_0^1(\omega_h)} \frac{|(v, \varphi_h)|}{\|v\|_{H_0^1(\omega_h)}}.$$

**Bizonyítás.** A (15.133)-(15.134) feladat megoldása  $y_h = A_h^{-1} \varphi_h = h_1 h_2 G_h \varphi_h$ , így az  $\ell$ -edik komponense

$$y_\ell = (\alpha_\ell, \varphi_h)_{(2)} = (g_\ell, \varphi_h)_{(0,h)},$$

ld. a (15.140) jelöléseket. Innen

$$|y_\ell| = |(g_\ell, \varphi_h)_{(0,h)}| \leq \|g_\ell\|_{H_0^1(\omega_h)} \|\varphi_h\|_{H_0^{-1}(\omega_h)},$$

ahol (ld. a 15.21. tételt)

$$\begin{aligned} \|g_\ell\|_{H_0^1(\omega_h)}^2 &= (A_h^0 g_\ell, g_\ell)_{(0,h)} \leq \frac{1}{k_0} (A_h g_\ell, g_\ell)_{(0,h)} = \frac{1}{k_0} (e_\ell, g_\ell)_{(2)} \\ &= \frac{1}{k_0} g_{\ell\ell} \leq \frac{1}{k_0^2} g_{\ell\ell}^0. \end{aligned}$$

Ebből és a 15.19. lemmából kapjuk az állítást.  $\square$

A lemmából következik az is, hogy a (15.133)-(15.134) séma konvergál maximumnormában, ld. a 33. feladatot.

A folytonos argumentumú függvények körében egyébként igaz a  $H_0^2 \hookrightarrow C$  beágyazás (amikor  $\Omega \in \mathbb{R}^d$ ,  $d < 4$ ). Vezessük le annak diszkrét megfelelőjét a vizsgált kétdimenziós esetben! Mint a  $H_0^2(\omega_h)$  tér normáját definiáljuk

$$\|y_h\|_{H_0^2(\omega_h)} := \|A_h^0 y_h\|_{(0,h)}.$$

**15.24. Lemma** (diszkrét beágyazási tétel). Tetszőleges, az  $\bar{\omega}_h$  rács peremén eltűnő  $y$  rácsfüggvényre teljesül

$$\|y\|_{C(\omega_h)} \leq M(a, b) \|y\|_{H_0^2(\omega_h)},$$

ahol

$$M(a, b) := \left( \frac{a^3 b^3}{16(a^2 + b^2)^2} + \frac{\pi}{32}(a^2 + b^2) \right)^{1/2}. \quad (15.144)$$

**Bizonyítás.** Legyen  $y_h$  tetszőleges vektor és  $f_h := A_h^0 y_h$ . Ekkor (15.123)-nak megfelelően  $y_\ell = (g_\ell^0, f_h)_{(0,h)}$  és így

$$\begin{aligned} |y_\ell| &\leq \|g_\ell^0\|_{(0,h)} \|f_h\|_{(0,h)} = \|g_\ell^0\|_{(0,h)} \|A_h^0 y_h\|_{(0,h)} \\ &= \|g_\ell^0\|_{(0,h)} \|y_h\|_{H_0^2(\omega_h)}. \end{aligned}$$

Ezért elegendő a következő becslést bebizonyítani :

$$\|g_\ell^0\|_{(0,h)}^2 = \sum_{\xi \in \omega_h} (G_h^0(x^\ell, \xi))^2 h_1 h_2 \leq M^2,$$

alkalmas  $M$  konstanssal. Ehhez kiindulunk (15.128)-ból. Felhasználva a sajátfüggvények ortonormáltságát azt kapjuk, hogy

$$\|G_h^0(x^\ell, \cdot)\|^2 = \sum_{k=1}^{N_1-1} \sum_{j=1}^{N_2-1} \left( \frac{v^{(kj)}(x^\ell)}{\lambda_{kj}^h} \right)^2 \leq \frac{4}{ab} \sum_{k=1}^{N_1-1} \sum_{j=1}^{N_2-1} \left( \frac{1}{\lambda_{kj}^h} \right)^2,$$

és a 15.19. lemma bizonyításának gondolatmenetét megfelelően módosítva:

$$\begin{aligned} \|G_h^0(x, \cdot)\|^2 &\leq \frac{4}{ab} \left\{ \frac{a^4 b^4}{64(a^2 + b^2)^2} + \frac{1}{16} \int_0^{\pi^2} \int_1^{\sqrt{2} \max(N_1, N_2)} \frac{r \, dr \, d\varphi}{g^2(r, \varphi)} \right\} \\ &< M^2(a, b), \end{aligned}$$

ahol  $M = M(a, b)$  a (15.144) konstans. Itt felhasználtuk, hogy

$$\int_0^{\frac{\pi}{2}} \int_1^\infty \frac{r \, dr \, d\varphi}{g^2(r, \varphi)} = \int_0^{\frac{\pi}{2}} \frac{d\varphi}{\left(\frac{1}{a^2} \cos^2 \varphi + \frac{1}{b^2} \sin^2 \varphi\right)^2} \int_1^\infty \frac{dr}{r^3} = \frac{\pi ab(a^2 + b^2)}{8}. \quad \square$$

### 15.5.4 Differenciasémák vizsgálata $L_2$ -beli jobboldal esetén

Ebben a pontban vizsgáljuk a Poisson-egyenlet elsőfajú peremérték feladat megoldását differenciaséma segítségével és megmutatjuk, hogy hasonló becslés kapható, mint a lineáris végelem módszer esetén.

Legyen  $\Omega := \{0 < x_1 < a, 0 < x_2 < b\}$  és  $f \in L_2(\Omega)$ .

Ekkor a homogén peremérték feladathoz tartozó általánosított megoldás  $H^2(\Omega)$ -ból való és jellemezhető mint az az  $u \in H^2(\Omega) \cap H_0^1(\Omega)$ , amely teljesíti az

$$\iint_{\Omega} \Delta u(x) v(x) \, dx = - \iint_{\Omega} f(x) v(x) \, dx \quad (15.145)$$

egyenletet minden  $v \in L_2(\Omega)$ -ra.

Legyen  $\omega_h$  a szokásos téglalap alakú rács  $\Omega$ -ban  $h_1 := a/N_1$  és  $h_2 := b/N_2$  rácsállandókkal,  $h^2 := h_1^2 + h_2^2$ , és teljesüljön  $\kappa > 0$  konstanssal

$$\frac{1}{\kappa} \leq \frac{h_1}{h_2} \leq \kappa. \quad (15.146)$$

Ha  $\bar{x} := (\bar{x}_1, \bar{x}_2) \in \omega_h$ , akkor  $\square(\bar{x})$ -szel jelöljük a hozzátartozó elemi téglalapot :

$$\square(\bar{x}) := \left\{ (x_1, x_2) \in \Omega, |x_1 - \bar{x}_1| < \frac{h_1}{2}, |x_2 - \bar{x}_2| \leq \frac{h_2}{2} \right\}.$$

Legyen  $\chi_{\square}$  az elemi téglalap karakterisztikus függvénye :

$$\chi_{\square}(x) := \begin{cases} 1, & x \in \square, \\ 0, & \text{máskor.} \end{cases}$$

Ez a függvény  $L_2(\Omega)$ -ba tartozik. Behelyettesítve (15.145)-be :

$$\begin{aligned}
-\iint_{\Omega} f(x)\chi_{\square}(x) dx &= -\iint_{\square(\bar{x})} \Delta u dx = \int_{\bar{x}_2 - \frac{h_2}{2}}^{\bar{x}_2 + \frac{h_2}{2}} \int_{\bar{x}_1 - \frac{h_1}{2}}^{\bar{x}_1 + \frac{h_1}{2}} \frac{\partial^2 u}{\partial x_1^2} dx \\
&\quad + \int_{\bar{x}_1 - \frac{h_1}{2}}^{\bar{x}_1 + \frac{h_1}{2}} \int_{\bar{x}_2 - \frac{h_2}{2}}^{\bar{x}_2 + \frac{h_2}{2}} \frac{\partial^2 u}{\partial x_2^2} dx_2 dx_1 \quad (15.147) \\
&= \int_{\bar{x}_2 - \frac{h_2}{2}}^{\bar{x}_2 + \frac{h_2}{2}} \left[ \frac{\partial u}{\partial x_1} \left( \bar{x}_1 + \frac{h_1}{2}, x_2 \right) - \frac{\partial u}{\partial x_1} \left( \bar{x}_1 - \frac{h_1}{2}, x_2 \right) \right] dx_2 \\
&\quad + \int_{\bar{x}_1 - \frac{h_1}{2}}^{\bar{x}_1 + \frac{h_1}{2}} \left[ \frac{\partial u}{\partial x_2} \left( x_1, \bar{x}_2 + \frac{h_2}{2} \right) - \frac{\partial u}{\partial x_2} \left( x_1, \bar{x}_2 - \frac{h_2}{2} \right) \right] dx_1.
\end{aligned}$$

A fenti műveletek kényelmes leírását adják a *Sztyeklov-féle átlagolási operátorok* :

$$\begin{aligned}
(S_1 v)(\bar{x}) &:= \frac{1}{h_1} \int_{\bar{x}_1 - \frac{h_1}{2}}^{\bar{x}_1 + \frac{h_1}{2}} v(x_1, \bar{x}_2) dx_1 = \int_{-\frac{1}{2}}^{\frac{1}{2}} v(\bar{x}_1 + sh_1, \bar{x}_2) ds, \\
(S_2 v)(\bar{x}) &:= \frac{1}{h_2} \int_{\bar{x}_2 - \frac{h_2}{2}}^{\bar{x}_2 + \frac{h_2}{2}} v(\bar{x}_1, x_2) dx_2 = \int_{-\frac{1}{2}}^{\frac{1}{2}} v(\bar{x}_1, \bar{x}_2 + th_2) dt, \\
(S_1 S_2 v)(\bar{x}) &:= \frac{1}{h_1 h_2} \iint_{\square(\bar{x})} v(x) dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} v(\bar{x}_1 + sh_1, \bar{x}_2 + th_2) ds dt.
\end{aligned}$$

Az  $S_1 S_2$  operátort alkalmazva (15.145)-re megkapjuk a  $(h_1 h_2$ -vel osztott) (15.147) azonosságot. Az ott kiszámítottak szerint  $v \in H^1(\Omega)$ -ra érvényes

$$\begin{aligned}
\left( S_1 \frac{\partial v}{\partial x_1} \right) (\bar{x}) &= \frac{1}{h_1} \left[ v \left( \bar{x}_1 + \frac{h_1}{2}, \bar{x}_2 \right) - v \left( \bar{x}_1 - \frac{h_1}{2}, \bar{x}_2 \right) \right], \\
\left( S_2 \frac{\partial v}{\partial x_2} \right) (\bar{x}) &= \frac{1}{h_2} \left[ v \left( \bar{x}_1, \bar{x}_2 + \frac{h_2}{2} \right) - v \left( \bar{x}_1, \bar{x}_2 - \frac{h_2}{2} \right) \right].
\end{aligned}$$

Ezekkel a jelölésekkel (15.147) tehát felírható mint

$$\begin{aligned}
\frac{1}{h_1} S_2 \left( \frac{\partial u}{\partial x_1} \left( \bar{x}_1 + \frac{h_1}{2}, \cdot \right) - \frac{\partial u}{\partial x_1} \left( \bar{x}_1 - \frac{h_1}{2}, \cdot \right) \right) \Big|_{x_2 = \bar{x}_2} &+ \quad (15.148) \\
\frac{1}{h_2} S_1 \left( \frac{\partial u}{\partial x_2} \left( \cdot, \bar{x}_2 + \frac{h_2}{2} \right) - \frac{\partial u}{\partial x_2} \left( \cdot, \bar{x}_2 - \frac{h_2}{2} \right) \right) \Big|_{x_1 = \bar{x}_1} &+ (S_1 S_2 f)(\bar{x}) = 0.
\end{aligned}$$

Itt a 15.6-ban tárgyalásra kerülő boxmódszernek megfelelően a  $\frac{\partial u}{\partial x_1}$  ill.  $\frac{\partial u}{\partial x_2}$  integráljait approximáljuk a középpont szabály segítségével, pl.

$$\begin{aligned} S_2 \left( \frac{\partial u}{\partial x_1} \left( \bar{x}_1 + \frac{h_1}{2}, \cdot \right) \right) \Big|_{x_2=\bar{x}_2} &= \frac{1}{h_2} \int_{\bar{x}_2 - \frac{h_2}{2}}^{\bar{x}_2 + \frac{h_2}{2}} \frac{\partial u}{\partial x_1} \left( \bar{x}_1 + \frac{h_1}{2}, x_2 \right) dx_2 \\ &\approx \frac{1}{h_1} (u(\bar{x}_1 + h_1, \bar{x}_2) - u(\bar{x}_1, \bar{x}_2)). \end{aligned}$$

Ez az approximáció  $u \in H^2(\Omega)$  feltételünk mellett jól definiált, hiszen míg  $\Omega$  dimenziója nem nagyobb 3-nál, addig a  $H^2(\Omega)$ -beli függvények folytonosak is.

Így eljutunk az

$$y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} + \varphi(\bar{x}) = 0, \quad \bar{x} \in \omega_h; \quad y(\bar{x}) = 0, \quad \bar{x} \in \gamma_h, \quad (15.149)$$

differenciasémához, ahol  $\gamma_h$  a rács pereme,

$$\varphi(\bar{x}) := (S_1 S_2 f)(\bar{x}), \quad (15.150)$$

és ahol az  $y$  rácsfüggvény közelíti az  $u$  pontos megoldás értékeit a rácspontokban. Hogy milyen jó ez a közelítés, azt a következőkben tisztázzuk az

$$\|y\|_{H_0^1(\omega_h)}^2 := \|y_{\bar{x}_1}\|_1^2 + \|y_{\bar{x}_2}\|_2^2$$

normában (amelynek definícióját a 15.20. tétel előtt adtuk meg), míg a pontos megoldást a következő félnormában vesszük :

$$\|u\|_{H^2(\Omega)}^2 := \iint_{\Omega} \left[ \left( \frac{\partial^2 u}{\partial x_1^2} \right)^2 + 2 \left( \frac{\partial^2 u}{\partial x_1 \partial x_2} \right)^2 + \left( \frac{\partial^2 u}{\partial x_2^2} \right)^2 \right] dx,$$

illetve a szokásos  $H^2(\Omega)$ -normában. Az alábbi tétel a differenciaséma  $y_h$  megoldásvektorát a pontos megoldás rácspontbeli értékeiből képzett  $\vec{u}_h$  vektorral hasonlítja össze.

**15.25. Tétel** (differenciaséma konvergenciája  $f \in L_2$  esetén). Teljesüljön (15.146). Akkor a (15.149), (15.150) séma  $y$  megoldása első rendben konvergál a (15.145) feladat  $u \in H^2(\Omega) \cap H_0^1(\Omega)$  megoldásához : Igaz az

$$\|y_h - \vec{u}_h\|_{H_0^1(\omega_h)} \leq Mh \|u\|_{H^2(\Omega)} \quad (15.151)$$

becslés, ahol az  $M \geq 0$  konstans nem függ  $h$ -tól.

**Bizonyítás.** A tételt a következő lépésekben bizonyítjuk :

1) felírjuk a hibához,  $z := y - \vec{u}$ -hoz tartozó egyenlet jobboldalát „divergens” alakban, mint  $\eta_{1x_1} + \eta_{2x_2}$ , és levezetjük  $\|z\|_{H_0^1(\omega_h)}$  becslését az  $\eta_1, \eta_2$  segédfüggvények  $L_2(\omega_h)$ -féle norma segítségével;

2) alkalmazva a Bramble–Hilbert lemmát (ld. II. 11.7.2.-3.), megkapjuk  $\eta_1(\bar{x}), \eta_2(\bar{x})$  becslését  $\|u\|_{H^2(\square(\bar{x}))}$  segítségével,  $\bar{x} \in \omega_h$ ;

3) összegezve az eredményeket eljutunk a (15.151) becsléshez.

1) A hibaegyenlet

$$z_{\bar{x}_1 x_1} + z_{\bar{x}_2 x_2} + \psi(\bar{x}) = 0, \quad \bar{x} \in \omega_h; \quad z(\bar{x}) = 0, \quad \bar{x} \in \gamma_h, \quad (15.152)$$

ahol

$$\begin{aligned} \psi &:= u_{\bar{x}_1 x_1} + u_{\bar{x}_2 x_2} + S_1 S_2 f = (\eta_1)_{x_1} + (\eta_2)_{x_2}, \\ \eta_1 = \eta_1(u) &:= u_{\bar{x}_1} - S_2 \left( \frac{\partial u}{\partial x_1} \left( \bar{x}_1 - \frac{h_1}{2}, \cdot \right) \right), \\ \eta_2 = \eta_2(u) &:= u_{\bar{x}_2} - S_1 \left( \frac{\partial u}{\partial x_2} \left( \cdot, \bar{x}_2 - \frac{h_2}{2} \right) \right), \end{aligned}$$

ld. (15.148). Ezután a (15.152) differencia-egyenleteket skalárisan megszorozzuk  $z$ -vel :

$$\begin{aligned} (\psi, z)_{(0,h)} &= ((\eta_1)_{x_1}, z)_{(0,h)} + ((\eta_2)_{x_2}, z)_{(0,h)} \\ &= -(z_{\bar{x}_1 x_1}, z)_{(0,h)} - (z_{\bar{x}_2 x_2}, z)_{(0,h)} = \|z_{\bar{x}_1}\|_1^2 + \|z_{\bar{x}_2}\|_2^2. \end{aligned}$$

A parciális összegzést alkalmazhatjuk az  $\eta_1, \eta_2$ -féle tagokra is, mivel  $z$  homogén peremértékkel rendelkezik :

$$\begin{aligned} \|z\|_{H_0^1(\omega_h)}^2 &= (\eta_{1x_1}, z)_{(0,h)} + (\eta_{2x_2}, z)_{(0,h)} = -(\eta_1, z_{\bar{x}_1}]_1 - (\eta_2, z_{\bar{x}_2}]_2 \\ &\leq \|\eta_1\|_1 \|z_{\bar{x}_1}\|_1 + \|\eta_2\|_2 \|z_{\bar{x}_2}\|_2 \leq (\|\eta_1\|_1^2 + \|\eta_2\|_2^2)^{1/2} \|z\|_{H_0^1(\omega_h)}, \end{aligned}$$

ahol a Cauchy–Schwarz egyenlőtlenséget használtuk.

Ezzel megkaptuk az első közbülső eredményünket :

$$\|z\|_{H_0^1(\omega_h)} \leq (\|\eta_1\|_1^2 + \|\eta_2\|_2^2)^{1/2}. \quad (15.153)$$

2) Vizsgáljuk pl. az  $\eta_1$  segédfüggvényt. Ehhez legyen

$$\square_1(\bar{x}) := \left\{ (x_1, x_2) \in \Omega, \quad \left| x_1 - \left( \bar{x}_1 - \frac{h_1}{2} \right) \right| \leq \frac{h_1}{2}, \quad |x_2 - \bar{x}_2| \leq \frac{h_2}{2} \right\}$$

egy további elemi téglalap, amelyet

$$x_1 = \bar{x}_1 + (s-1)h_1, \quad x_2 = \bar{x}_2 + \left( t - \frac{1}{2} \right) h_2$$

segítségével az  $E$  egységnyezetre képezzük le. Legyen

$$U(s, t) := u(x_1(s), x_2(t)), \quad (s, t) \in E.$$

Ekkor

$$\eta_1 = \frac{1}{h_1} \left\{ u(\bar{x}_1, \bar{x}_2) - u(\bar{x}_1 - h_1, \bar{x}_2) - \frac{h_1}{h_2} \int_{\bar{x}_2 - \frac{h_2}{2}}^{\bar{x}_2 + \frac{h_2}{2}} \frac{\partial u}{\partial x_1} \left( \bar{x}_1 - \frac{h_1}{2}, x_2 \right) dx_2 \right\}$$

felírható mint

$$\eta_1 = \frac{1}{h_1} l(U), \quad l(U) := U\left(1, \frac{1}{2}\right) - U\left(0, \frac{1}{2}\right) - \int_0^1 \frac{\partial U}{\partial s} \left(\frac{1}{2}, t\right) dt.$$

Azért mentünk át  $\eta_1$ -ről  $l$ -hez, mert így a diszkretizációtól független konstanssal alkalmazhatjuk a Bramble–Hilbert-lemmát. Ehhez jegyezzük meg, hogy  $u \in H^2(\Omega)$  miatt  $U \in H^2(E)$ . Ennek alapján  $l$  valóban lineáris funkcionál,  $l : H^2(E) \rightarrow \mathbb{R}$ .

Korlátossága következik a beágyazási tételekből :

$$H^2(E) \hookrightarrow C(E), \quad \text{így} \quad \left| U\left(1, \frac{1}{2}\right) \right|, \left| U\left(0, \frac{1}{2}\right) \right| \leq c_0 \|U\|_{H^2(E)},$$

és  $E_1 := \{(s, t) \in E, s = 1/2\}$ -del érvényes

$$H^2(E) \hookrightarrow H^1(E_1),$$

így

$$\left| \int_0^1 \frac{\partial U}{\partial s} \left(\frac{1}{2}, t\right) dt \right| \leq \left( \int_0^1 \left( \frac{\partial U}{\partial s} \left(\frac{1}{2}, t\right) \right)^2 dt \right)^{1/2} \leq c_1 \|U\|_{H^2(E)}.$$

Itt  $c_0$  és  $c_1$  a megfelelő beágyazási konstansok (fent (15.144)-ben szerepelt a  $H_0^2 \hookrightarrow C$  beágyazási konstansa).

Összesítve :

$$\begin{aligned} |l(U)| &\leq (2c_0 + c_1) \|U\|_{H^2(E)} \\ &:= (2c_0 + c_1) \left\{ \iint_E \left[ U^2 + \left( \frac{\partial U}{\partial s} \right)^2 + \left( \frac{\partial U}{\partial t} \right)^2 \right] ds dt + |U|_{H^2(E)} \right\}^{1/2}. \end{aligned}$$

A Bramble–Hilbert-lemma alkalmazásához hiányzik még annak ellenőrzése, hogy  $l(p)$  eltűnik az összes, legfeljebb másodfokú  $p$  polinomon. Csak a  $p = s^2$  és  $p = st$  eseteket nézzük meg. Ekkor

$$l(p) = 1 - 2 \int_0^1 \frac{1}{2} dt = 0, \quad \text{ill.} \quad l(p) = \frac{1}{2} - \int_0^1 t dt = 0.$$



Most a Bramble–Hilbert-lemma szerint van olyan  $M_0$  konstans, hogy

$$|l(U)| \leq M_0 |U|_{H^2(E)} \text{ vagyis } |\eta_1(u)| \leq \frac{M_0}{h_1} |U|_{H^2(E)}.$$

Ezután következik az  $|U|_{H^2(E)}$  visszaszámítása  $|u|_{H^2(\square_1)}$ -re, amelynek során előkerül az egész gondolatmenet lényege, mert itt nyerünk egy  $h$ -hatványt :

$$\begin{aligned} |U|_{H^2(E)}^2 &= \iint_E \left[ \left( \frac{\partial^2 U}{\partial s^2} \right)^2 + 2 \left( \frac{\partial^2 U}{\partial s \partial t} \right)^2 + \left( \frac{\partial^2 U}{\partial t^2} \right)^2 \right] ds dt \\ &= \frac{1}{h_1 h_2} \iint_{\square_1} \left[ \left( h_1^2 \frac{\partial^2 u}{\partial x_1^2} \right)^2 + 2 \left( h_1 h_2 \frac{\partial^2 u}{\partial x_1 \partial x_2} \right)^2 + \left( h_2^2 \frac{\partial^2 u}{\partial x_2^2} \right)^2 \right] dx \\ &\leq \frac{h^4}{h_1 h_2} |u|_{H^2(\square_1)}. \end{aligned}$$

(Emlékeztetünk, hogy  $h^2 = h_1^2 + h_2^2$ .) Így végülis

$$|\eta_1(u)| \leq \frac{M_0 h^2}{h_1 (h_1 h_2)^{1/2}} |u|_{H^2(\square_1)}. \quad (15.154)$$

Egészen analóg módon nyerjük  $|\eta_2(u)|$  becslését, bevezetve a

$$\square_2(\bar{x}) := \left\{ (x_1, x_2) \in \Omega, \quad |x_1 - \bar{x}_1| \leq \frac{h_1}{2}, \quad \left| x_2 - \left( \bar{x}_2 - \frac{h_2}{2} \right) \right| \leq \frac{h_2}{2} \right\}$$

elemi téglalapot stb.. A végeredmény

$$|\eta_2(u)| \leq \frac{M_0 h^2}{h_2 (h_1 h_2)^{1/2}} |u|_{H^2(\square_2)}. \quad (15.155)$$

A (15.154) és (15.155) becslésekben álló  $(h_1 h_2)^{-1/2}$  szorzó „magától” tűnik el a bizonyítás 3. lépése során, míg az  $1/h_1$ , ill.  $1/h_2$  szorzók miatt kénytelenek leszünk egy  $h$ -hatványt feláldozni.

3) A (15.154) és (15.155) becsléseket behelyettesítjük (15.153)-ba. Figyelembe véve az  $\|\eta_1\|_1^2$ ,  $\|\eta_2\|_2^2$ , valamint a  $\square_1$  és  $\square_2$  definícióit, azt kapjuk, hogy

$$\begin{aligned} \|\eta_1\|_1^2 &\leq \left( M_0 \frac{h^2}{h_1} \right)^2 \sum_{h_1 \leq \bar{x}_1 \leq 1} \sum_{h_2 \leq \bar{x}_2 \leq b-h_2} |u|_{H^2(\square_1(\bar{x}))}^2 \\ &= \left( M_0 \frac{h^2}{h_1} \right)^2 \int_0^a \int_{\frac{h_2}{2}}^{b-\frac{h_2}{2}} \left[ \left( \frac{\partial^2 u}{\partial x_1^2} \right)^2 + 2 \left( \frac{\partial^2 u}{\partial x_1 \partial x_2} \right)^2 + \left( \frac{\partial^2 u}{\partial x_2^2} \right)^2 \right] dx \\ &\leq \left( M_0 \frac{h^2}{h_1} \right)^2 |u|_{H^2(\Omega)}^2. \end{aligned}$$

Hasonlóan

$$\begin{aligned} \|\eta_2\|_2^2 &\leq \left(M_0 \frac{h^2}{h_2}\right)^2 \int_{\frac{h_1}{2}}^{a-\frac{h_1}{2}} \int_0^b \left[ \left(\frac{\partial^2 u}{\partial x_1^2}\right)^2 + 2 \left(\frac{\partial^2 u}{\partial x_1 \partial x_2}\right)^2 + \left(\frac{\partial^2 u}{\partial x_2^2}\right)^2 \right] dx_2 dx_1 \\ &\leq \left(M_0 \frac{h^2}{h_2}\right)^2 |u|_{H^2(\Omega)}^2, \end{aligned}$$

és így

$$\begin{aligned} \|z\|_{H_0^1(\omega_h)}^2 &\leq M_0^2 (h_1^2 + h_2^2)^2 \left(\frac{1}{h_1^2} + \frac{1}{h_2^2}\right) |u|_{H^2(\Omega)}^2 \\ &= M_0^2 (h_1^2 + h_2^2) \left(\frac{h_1}{h_2} + \frac{h_2}{h_1}\right)^2 |u|_{H^2(\Omega)}^2 \\ &\leq M_0^2 h^2 \left(\kappa + \frac{1}{\kappa}\right)^2 |u|_{H^2(\Omega)}^2 \leq \left(M_0 h \left(\kappa + \frac{1}{\kappa}\right) \|u\|_{H^2(\Omega)}\right)^2, \end{aligned}$$

felhasználva a (15.146) feltételt. Ezzel a tétel be van bizonyítva.  $\square$

**Megjegyzések. 1.** Feltételeink mellett folytatható a (15.151) becslés, mert van olyan  $c$  konstans, hogy  $\|u\|_{H^2(\Omega)} \leq c \|f\|_{L_2(\Omega)}$ .

**2.** Akkor is igaz a tétel, ha – elhagyva az  $u \in H_0^1$  feltételt – (15.145) mellé előírjuk az

$$u|_{\Gamma} = g \in H^{3/2}(\Gamma)$$

peremfeltételt. Ugyanis ekkor is  $u \in H^2(\Omega)$ , a  $g$  függvény folytonos  $\Gamma$ -n, így  $y$  számára előírhatjuk  $g$  értékeit a rács  $\gamma_h$  peremén. Ennek eredményeképpen továbbra is érvényes  $z = y - \vec{u} = 0$ ,  $x \in \gamma_h$ , és az egész bizonyítás változatlan marad. Jelen esetben van olyan  $c$  konstans, hogy

$$\|y - \vec{u}\|_{H_0^1(\omega_h)} \leq Mh \|u\|_{H^2(\Omega)} \leq cMh [\|f\|_{L_2(\Omega)} + \|g\|_{H^{3/2}(\Gamma)}].$$

**3.** Hasonlóan, mint a végeelem módszer esetén, itt is felmerül a kérdés, hogyan kell a (15.150) integráljait kiszámítani. Amennyiben  $f$  – egy-két szingularitástól eltekintve – sima függvény, akkor azokban az elemi téglalapokban, amelyekbe a (négyzetben integrálható) szingularitás esik, annak lényeges részének integrálját analitikusan számítjuk ki. Egyébként (a szingularitást tartalmazó elemi téglalapokban a maradékfüggvényre, az egyéb téglalapokban  $f$ -re) elegendő a középpont szabályt alkalmazni.

**4.** Az  $S_1 S_2$  átlagolási operátor akkor is használható differenciaséma levezetésére, ha a peremérték feladat

$$\begin{aligned} \frac{\partial}{\partial x_1} \left( k_1(x) \frac{\partial u}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left( k_2(x) \frac{\partial u}{\partial x_2} \right) + f(x) &= 0, & x \in \Omega, \\ u &= 0, & x \in \Gamma, \end{aligned}$$

ahol

$$k_1, k_2 \in C^1(\overline{\Omega}), \quad 0 < c_1 \leq k_1(x), k_2(x) \leq c_2, \quad x \in \overline{\Omega},$$

és  $f \in L_2(\Omega)$ ,  $u \in H^2(\Omega)$ . Ekkor is érvényes a tétel (valamivel bonyolultabb bizonyítással: az  $\eta$  hibamennyiségek ekkor az  $u$  megoldás és a  $k_1, k_2$  együtt-hatók lineáris funkcionáljai). A differenciaséma – eltekintve a jobboldaltól – a szokásos:

$$\begin{aligned} (a_1 y_{\bar{x}_1})_{x_1} + (a_2 y_{\bar{x}_2})_{x_2} + S_1 S_2 f &= 0, & x \in \omega_h, \\ y &= 0, & x \in \gamma_h, \end{aligned}$$

ahol

$$a_1(x) = k_1\left(x_1 - \frac{h_1}{2}, x_2\right), \quad a_2(x) = k_2\left(x_1, x_2 - \frac{h_2}{2}\right).$$

### 15.5.5 Dirac-féle $\delta$ -függvényt tartalmazó jobboldal

Előző eredményeink nem fedik le azt a gyakorlatban fontos esetet, hogy a Poisson-egyenlet  $f$  jobboldala tartalmaz Dirac-féle  $\delta$ -függvényt. Ilyen feladatra jutunk pl. abban a környezetvédelmi problémában, amelyet a 16.1.1. pontban vázoljuk, ha ott nincs szél és ha a forrás időtől függetlenül működik (ez pl. egy ipari kémény modellje lehet). Hővezetési értelmezés esetén a ponthegesztés vezet ilyen feladatra, és ha az egyenlet a megfeszített, megterhelt membrán modellje, akkor egy pontszerű teherről van szó. Ilyen esetekben a szingularitás kis környezetében valójában más modellt kellene alkalmazni. Ennek ellenére a  $\delta$ -függvényt tartalmazó egyenletek elfogadottak a gyakorlatban.

Míg egy vonalra koncentrált  $f$  forrás esetén – mint amilyennel modellezhető környezetvédelmi szempontból egy forgalmas országút –  $f \in H^{-1}(\Omega)$ , addig a  $\delta$ -függvényt tartalmazó  $f$ -ről ismert (ha  $\Omega \in \mathbb{R}^2$ ), hogy  $f \in H^\gamma(\Omega)$ , tetszőleges  $\gamma < -1$ -gyel, a következő értelemben :

$$\|f\|_{H^\gamma} := \sup_{g \neq 0} \frac{(f, g)_0}{\|g\|_{H^{(-\gamma)}}} \text{ véges.}$$

Ez azt jelenti, hogy a peremérték feladat megfogalmazása a szokásos variációs egyenlet segítségével sem elegendő, hiszen ott feltételezzük, hogy legalább  $f \in H^{-1}(\Omega)$ .

Amennyiben  $\gamma > -3/2$ , Nečas szerint mégis definiálható egy általánosított megoldás.

A másodrendű feladatokra korlátozva Nečas vizsgálatát, megállapítható, hogy az

$$a(u, v) := \int_{\Omega} \sum_{|\alpha|, |\beta| \leq 1} a_{\alpha\beta} \frac{\partial^{|\alpha|} u}{\partial x^\alpha} \frac{\partial^{|\beta|} v}{\partial x^\beta} dx$$

bilineáris formának, ha  $H^1$ -elliptikusnak tételezzük fel, akkor is van jól definiált értéke, ha

$$u \in H^{1-\delta}, \quad \frac{1}{2} > \delta > 0, \quad \text{de } v \in H^{1+\delta},$$

ahol multiindexes jelölést alkalmaztunk, ld. II. 11.7.2. elején (itt  $\alpha = (\alpha_1, \alpha_2)$ ,  $|\alpha| = \alpha_1 + \alpha_2$  stb.).

A peremfeltételt egy alkalmas  $u_0 \in H^{1-\delta}(\Omega)$  függvény segítségével fogalmazzuk meg : legyen

$$u - u_0 \in H_0^{1-\delta}(\Omega).$$

Ezen feltételt kielégítő  $u$ -t keressük, mint az

$$a(u, v) = (f, v)_0 \quad \text{minden } v \in H^{1+\delta}\text{-re}$$

variációs egyenlet megoldását, amikor  $f \in H^{-1-\delta}$ . Ennek van Nečas szerint pontosan egy  $u \in H^{1-\delta}$  megoldása úgy, hogy ( $u$ -tól független  $C$  konstanssal) érvényes

$$\|u\|_{H^{1-\delta}(\Omega)} \leq C (\|f\|_{H^{-1-\delta}(\Omega)} + \|u_0\|_{H^{1-\delta}(\Omega)}).$$

Ez nekünk éppen elég lesz ahhoz, hogy azt az esetet kezelhessük, amikor  $f$  a Dirac-féle  $\delta$ -függvény. Közben az  $f \in H^{-1}(\Omega)$  esetet is intézzük el.

Az  $a$  bilineáris forma a Laplace-operátornak feleljen meg ( $a_{\alpha,\beta} = 1$ , ha  $\alpha = \beta = (1, 0)$  vagy  $\alpha = \beta = (0, 1)$ , de  $a_{\alpha,\beta} = 0$ , ha  $\alpha = \beta = (1, 0)$ ).

Mint differenciasémát megint a (15.149) sémát fogjuk vizsgálni, de más jobboldallal :

$$\varphi(\bar{x}) := T_1 T_2 f, \quad \bar{x} \in \omega_h, \quad (15.156)$$

ahol

$$(T_1 f)(\bar{x}) := \frac{1}{h_1} \int_{\bar{x}_1 - h_1}^{\bar{x}_1 + h_1} \left( 1 - \left| \frac{x_1 - \bar{x}_1}{h_1} \right| \right) f(x_1, \bar{x}_2) dx_1 = (S_1^2 f)(\bar{x}),$$

$$(T_2 f)(\bar{x}) := \frac{1}{h_2} \int_{\bar{x}_2 - h_2}^{\bar{x}_2 + h_2} \left( 1 - \left| \frac{x_2 - \bar{x}_2}{h_2} \right| \right) f(\bar{x}_1, x_2) dx_2 = (S_2^2 f)(\bar{x}).$$

Ezekkel az átlagolási operátorokkal már a spline-függvények tárgyalása során találkoztunk (ld. I. 267. o., (4.62)) : a másodrendű differenciahányados segítségével kifejezhető :

$$T_i \frac{\partial^2 f}{\partial x_i^2} = f_{\bar{x}_i x_i}, \quad i = 1, 2.$$

A  $T_1 T_2$  operátort a Poisson-egyenletre alkalmazva, megkapjuk a

$$0 = T_2 u_{\bar{x}_1 x_1} + T_1 u_{\bar{x}_2 x_2} + \varphi = (T_2 u)_{\bar{x}_1 x_1} + (T_1 u)_{\bar{x}_2 x_2} + \varphi \quad (15.157)$$

azonosságot. Ennek approximációjára a (15.149), (15.156) sémát fogjuk felhasználni.

Mivel majd a konvergencia bizonyítása során meg fogjuk mutatni, hogy a különböző (15.157)-ben fellépő kifejezések jól definiáltak, itt csak a következőre hívjuk fel a figyelmet :

Ha (15.156) jól definiált, akkor a (15.149), (15.156) séma  $y$  megoldása is az. Közben viszont az  $u$  általánosított megoldás pontbeli értékei már akkor sem jól definiáltak, amikor  $f \in H^{-1}(\Omega)$  és így  $u \in H^1(\Omega)$ .

Vagyis : nem jó kérdés az, vajon  $y \rightarrow u$  amikor  $h \rightarrow 0$ ?

Azt fogjuk bebizonyítani, hogy  $y \rightarrow \bar{u}$ , ahol

$$\bar{u}(x) := \begin{cases} (S_1 S_2 u)(\bar{x}), & x \in \omega_h, \\ 0, & x \in \gamma_h. \end{cases} \quad (15.158)$$

A differenciaséma megoldása – a (15.156) jobboldal esetén – tehát a pontos megoldás *integrálját* közelíti a  $\square(x)$  elemi téglalap felett.

**15.26. Tétel** (Szamarszkij, Lazarov, Makarov; differenciaséma konvergenciája Dirac-féle  $\delta$ -függvény jelenlétében). Legyen  $f \in H^\gamma$ ,  $-\frac{3}{2} < \gamma \leq -1$ ,  $u$  a Poisson-egyenlet általánosított megoldása, homogén peremfeltételek mellett (Nečas értelmében), és  $y_h$  legyen a (15.149), (15.156) differenciaséma megoldásvektora. Ekkor érvényes

$$\|y_h - R_h u\|_{L_2(\omega_h)} \leq M h^{2+\gamma} \|f\|_{H^\gamma(\Omega)},$$

ahol  $R_h u$  jelöli a (15.158) által definiált rácsfüggvényhez tartozó vektort (a 15.3.1-beli jelöléssel  $R_h u = (\vec{u})_h$ ).

**Bizonyítás.** A tételt hasonló három olyan lépéssel fogjuk bebizonyítani, mint a 15.25. tételt.

1) Legyen  $z := y - \bar{u}$  a hiba. Ekkor (használva (15.157)-et)

$$\begin{aligned} z_{\bar{x}_1 x_1} + z_{\bar{x}_2 x_2} &= -T_1 T_2 f - \bar{u}_{\bar{x}_1 x_1} - \bar{u}_{\bar{x}_2 x_2} = (\eta_1)_{\bar{x}_1 x_1} + (\eta_2)_{\bar{x}_2 x_2}, \quad (15.159) \\ \eta_1 &:= T_2 u - \bar{u}, \quad \eta_2 := T_1 u - \bar{u}, \quad \bar{x} \in \omega_h. \end{aligned}$$

Ugyanezekkel a képletekkel definiálható  $\eta_1$  (ill.  $\eta_2$ ) akkor is, amikor  $\bar{x} \in \gamma_h$  és  $\bar{x}_1 = 0$  vagy  $\bar{x}_1 = a$  (ill.  $\bar{x}_2 = 0$  vagy  $\bar{x}_2 = b$ ), és értékük ekkor nulla. A perem egyéb oldalain előírjuk a homogén peremértékeket  $\eta_1$ , ill.  $\eta_2$  számára. Így mindenütt  $\eta_1(\bar{x}) = \eta_2(\bar{x}) = 0$ ,  $\bar{x} \in \gamma_h$ .

Legyenek  $A_h, A_{1h}, A_{2h}$  azok a mátrixok, amelyek (a homogén peremértékek eliminációja után) hozzátartoznak  $z_{\bar{x}_1 x_1} + z_{\bar{x}_2 x_2}$ -höz,  $z_{\bar{x}_1 x_1}$ -hez, ill.  $z_{\bar{x}_2 x_2}$ -höz. Már korábban megmutattuk, hogy  $A_h = A_{1h} + A_{2h}$  szimmetrikus és

pozitív definit, viszont  $A_{1h}$  és  $A_{2h}$  felcserélhetőek (mint az  $x_1$ , ill.  $x_2$  szerinti differenciahányadosak képzése is az). Így (15.159) a

$$z_h = -A_h^{-1}(A_{1h}\eta_{1h} + A_{2h}\eta_{2h}) = -A_{1h}A_h^{-1}\eta_{1h} - A_{2h}A_h^{-1}\eta_{2h} \quad (15.160)$$

vektoralakban felírható. Mivel tetszőleges,  $\gamma_h$ -n eltűnő  $w$  rácsfüggvény esetén az  $L_2(\omega_h)$ -normában igaz

$$\begin{aligned} \|A_h w_h\|^2 &= \|w_{\bar{x}_1 x_1}\|^2 + 2(w_{\bar{x}_1 x_1}, w_{\bar{x}_2 x_2}) + \|w_{\bar{x}_2 x_2}\|^2 \\ &= \|A_{1h} w_h\|^2 + 2(w_{\bar{x}_1 x_2}, w_{\bar{x}_1 x_2}) + \|A_{2h} w_h\|^2 \\ &\geq \|A_{1h} w_h\|^2, \|A_{2h} w_h\|^2, \end{aligned}$$

így

$$\|v_h\| \geq \|A_{ih} A_h^{-1} v_h\| = \|A_h^{-1} A_{ih} v_h\|, \quad i = 1, 2,$$

ahol  $v_h := A_h w_h$  is tetszőleges,  $\gamma_h$ -n eltűnő rácsfüggvény vektora. Ezért (15.160)-ból következik

$$\|z_h\|_{L_2(\omega_h)} \leq \|\eta_{1h}\|_{L_2(\omega_h)} + \|\eta_{2h}\|_{L_2(\omega_h)}. \quad (15.161)$$

2) Nečas szerint, ha  $f \in H^\gamma(\Omega)$  és  $\gamma > -\frac{3}{2}$ , akkor  $u \in H^{2+\gamma}$ . Azt fogjuk megmutatni, hogy  $\eta_1$  és  $\eta_2$  az  $u$  funkcionáljai a  $H^{2+\gamma}(\square)$  téren, ahol most  $\square = \square(\bar{x}) := \{x \in \Omega, |x_1 - \bar{x}_1| \leq h_1, |x_2 - \bar{x}_2| \leq h_2\}$  az elemi téglalap  $\bar{x} \in \omega_h$ -hoz.

Először is,  $\int_{\square} u \, dx$  jól definiált, mivel a beágyazási tételek szerint  $H^{2+\gamma}(\square) \hookrightarrow L_2(\square)$  ha  $0 < 2 + \gamma$ . Továbbá,  $H^{2+\gamma}(\square) \hookrightarrow L_2(\square_1)$ , ha  $2(2 + \gamma) \geq d = 2$ , vagy ha  $0 < 2 - 2(2 + \gamma) < 1$ , azaz : ha vagy  $\gamma \geq -1$  vagy  $-\frac{3}{2} < \gamma < -1$ . Itt  $\square_1 := \{x \in \square, x_2 = \bar{x}_2\}$ , ha  $T_1 u$ -t vizsgáljuk, ill.  $\square_1 := \{x \in \square, x_1 = \bar{x}_1\}$ , ha  $T_2 u$ -t nézzük. Ezzel megvan  $\eta_1$  és  $\eta_2$  korlátossága (additivitásuk világos),

$$|\eta_1|, |\eta_2| \leq c_0 \|u\|_{H^{2+\gamma}(\square)}.$$

A Bramble–Hilbert lemma alkalmazásához megjegyezzük, hogy  $u = \text{const}$  esetén

$$S_1 S_2 u = T_1 u = T_2 u = u.$$

Úgy mint a 15.25. tétel bizonyításában, elvégezzük a transzformációt az egységnyezetre ( $\square \rightarrow E, u \rightarrow U$ ), ahol a Bramble–Hilbert lemma szerint

$$|\eta_1|, |\eta_2| \leq M_0 |U|_{H^{2+\gamma}(E)}$$

$h_1$ -től és  $h_2$ -től független  $M_0$  konstanssal és a  $H^{2+\gamma}(E)$  félnormájával. Visszatranszformálva  $\square$ -ra kapjuk

$$|\eta_1|, |\eta_2| \leq M_0 \frac{h^{2+\gamma}}{(4h_1 h_2)^{1/2}} |u|_{H^{2+\gamma}(E)}, \quad (15.162)$$

ahol  $4h_1h_2$  a  $\square(\bar{x})$  elemi téglalap térfogata és  $h^2 = h_1^2 + h_2^2$ .

Most vizsgáljuk még

$$\varphi = \varphi(\bar{x}) = (T_1T_2f)(\bar{x}) = \int_{\square(\bar{x})} \psi(\bar{x}, x)f(x) dx$$

jó definiáltságát is, ahol  $\psi(\bar{x}, x) := (1 - |x_1 - \bar{x}_1|/h_1)(1 - |x_2 - \bar{x}_2|/h_2)$  !

Ha  $f \in H^{-1}(\Omega)$ , akkor ismeretes, hogy felírható mint

$$f = f_0 + \frac{\partial}{\partial x_1}f_1 + \frac{\partial}{\partial x_2}f_2, \quad f_i \in L_2(\Omega), \quad i = 0, 1, 2.$$

Itt a deriváltakat általánosított értelemben kell érteni. Folytatva a  $\psi$  függvényt  $\square(\bar{x})$ -on kívül nullával, írhatunk az  $L_2(\Omega)$  skalárszorzatában

$$\varphi := (f, \psi)_0 = (f_0, \psi)_0 - \left(f_1, \frac{\partial}{\partial x_1}\psi\right)_0 - \left(f_2, \frac{\partial}{\partial x_2}\psi\right)_0,$$

ahol már szokásos módon (majdnem mindenütt) értelmezhetjük a deriváltakat. Ha most

$$f = f_0 + \frac{\partial}{\partial x_1}f_1 + \frac{\partial}{\partial x_2}f_2 + \delta(x - x_*), \quad x_* \in \square(\bar{x}),$$

tehát  $f \in H^\gamma(\Omega)$ ,  $\gamma < -1$ , akkor

$$\varphi = (f_0, \psi)_0 - \left(f_1, \frac{\partial}{\partial x_1}\psi\right)_0 - \left(f_2, \frac{\partial}{\partial x_2}\psi\right)_0 + \psi(x_*).$$

Az  $x_*$ -hoz legközelebbi négy rácspontban a Dirac-féle  $\delta$ -függvény ad nemnulla értéket  $\varphi$ -hez, és ezen négy érték összege  $1/(h_1h_2)$ .

3) A (15.161) és (15.162) becslésekből következik a tétel állítása, hasonlóan mint a 15.25. tétel 3. bizonyítási lépésben. Itt csupán arra kell figyelni, hogy most az elemi téglalapok részlegesen átfedik egymást, és emiatt

$$\begin{aligned} \|\eta_{1h}\|_{L_2(\omega_h)}^2, \|\eta_{2h}\|_{L_2(\omega_h)}^2 &\leq 2 \left(M_0 \frac{h^{2+\gamma}}{2}\right)^2 \|u\|_{H^{2+\gamma}(\Omega)}^2 \\ &\leq \frac{1}{2} (M_0 h^{2+\gamma})^2 \|u\|_{H^{2+\gamma}(\Omega)}^2. \quad \square \end{aligned}$$

**Megjegyzések.** 1. Ha tehát  $f$   $\delta$ -függvényt tartalmaz, akkor  $f \in H^{-1-\varepsilon}(\Omega)$ , és ekkor  $\|y_h - R_h u\|_{L_2(\omega_h)} = O(h^{1-\varepsilon})$ . Ha  $f \in H^{-1}(\Omega)$ , tehát amikor  $u$  a szokásos variációs feladattal definiálható, akkor  $\|y_h - R_h u\|_{L_2(\omega_h)} = O(h)$ .

**2.** A bizonyításból kiderül, hogy az inhomogén peremértékek esete nehézség nélkül kezelhető (sem a bizonyításban, sem az eredményben nem lesz változás), ha  $u(x) = g(x)$ ,  $x \in \Gamma$ ,  $g \in H^{3/2+\gamma}(\Gamma)$ , és ha a diszkrét feladat peremfeltételeként előírjuk  $y = T_{3-i}g$  a perem azon oldalain, amelyek  $x_i$ -vel párhuzamosak ( $i = 1, 2$ ).

**3.** A tétel lényeges feltétele  $\gamma > -\frac{3}{2}$ , de ilyen általánosságban ( $\eta_1, \eta_2$  tulajdonságai mellett) csak

$$\|y_h - R_h u\|_{L_2(\omega_h)} \leq M h^{\min(1, 2+\gamma)} \|f\|_{H^\gamma(\Omega)}$$

bizonyítható. Viszont  $\gamma > -1$  esetén már nincs szükség  $\bar{u}$  bevezetésére, így vehetünk  $\eta_1 := T_2 u - u$ ,  $\eta_2 := T_1 u - u$ , és ekkor jobb becslés levezethető (ld. a 34. feladatot).

Ezen pont tanulságaként azt mondhatjuk, hogy *a differenciasémák rácspontokhoz kötött értékei alatt nem kell feltétlenül a megoldás értékét érteni : az egy lokális átlagérték is lehet, és így a differenciaséma képes általánosított megoldás közelítésére.*

## 15.6 Véges térfogat módszer

### 15.6.1 Bevezetés

Ahogy láttuk, a differenciasémák általánosíthatók arra az esetre, hogy a differenciáloperátor változó együtthatót tartalmaz, bár ekkor bonyolultabb a stabilitási és pontossági vizsgálat : változó együtthatók esetén nehéz az alkalmas  $w$  majoráns függvény meghatározása, amely a mátrix  $M$ -tulajdonságát mutatná, ld. II. 11.4.8-ban az egydimenziós esetet. A többdimenziós esetben ekkor általában  $L_2$ -féle technikát kell alkalmazni, ld. 15.5.2-3.

A többdimenziós feladatok másik, tipikus nehézsége viszont az, hogy a megoldási tartomány nem téglalap. Ekkor már a Shortley–Weller approximáció sem garantálja azt, hogy az önadjungált elliptikus peremérték feladat diszkrétizációja szimmetrikus pozitív definit mátrixú egyenletrendszerre vezessen. Ez nem csupán a matematikai szépség csorbítása, hanem a megoldási algoritmusok munkáját megnehezíti, és sajátérték feladatok esetén minőségileg hibás eredményeket vonhat maga után.

Továbbá, a görbevonalú peremen megadott másod- vagy harmadfajú peremfeltétel differencia-approximációja sem problémamentes (ld. a 14. feladatot), különösen akkor, amikor a peremfeltételek együtthatói is változóak.

Ezen nehézségek elkerülésének az érdekében a következőkben a *véges térfogat módszert*, más néven *mérlegegyenlet-* vagy *boxmódszert* a differenciasémák szerkesztésére írjuk le.



Tekintsük az alábbi peremérték feladatot a (15.2) egyenletre :

$$\operatorname{div}(k \operatorname{grad} u) + f(x) = 0, \quad x = (x_1, x_2) \in \Omega, \quad (15.163)$$

$$u(x) = g_1(x), \quad x \in \Gamma_1, \quad k(x) \frac{\partial u}{\partial n} + \sigma(x)u = g_{23}(x), \quad x \in \Gamma_{23}, \quad (15.164)$$

$$\Gamma = \Gamma_1 \cup \Gamma_{23}, \quad \Gamma_1 \neq \emptyset.$$

ahol  $\Omega$  véges tartomány. A pereméről,  $\Gamma$ -ról az egyszerűség kedvéért felteszszük, hogy az olyan töröttvonal (nulla belső és külső szögek nélkül), amelynek csak véges sok töréspontja van.

Az együtthatók illetőleg feltesszük, hogy

$$\sigma(x) \geq \sigma_0 > 0, \quad x \in \Gamma_3, \quad \sigma(x) \equiv 0, \quad x \in \Gamma_2; \quad \Gamma_2 \cup \Gamma_3 = \Gamma_{23}, \quad (15.165)$$

$$0 < k_0 \leq k(x) \leq k_1 > 0, \quad x \in \bar{\Omega}. \quad (15.166)$$

A következő simasági feltételek teljesülését követeljük  $j = 0$ -ra ill.  $j = 1$ -re :

$$k \in C^{1+j}(\bar{\Omega}), \quad f \in H^{1+j}(\Omega), \quad \sigma, g_{23} \in H^{1+j}(\Gamma_{23}), \quad u \in H^{2+j}(\Omega),$$

ezenkívül legyen  $g_1 \in C(\Gamma_1)$ . Megjegyezzük, hogy az együtthatók felsorolt tulajdonságai nem biztosítják az  $u$  megoldás tartalmazását  $H^{2+j}$ -ben. A peremfeltételek típusa változásának pontjaiban (valamint a perem sarokpontjaiban, ld. 15.2.) olyan szingularis komponensek jelenhetnek meg a megoldásban, amelyek miatt csupán  $u \in H^1(\Omega)$  igaz; maguk a szingularis megoldási komponensek  $C(\bar{\Omega}) \cap C^2(\Omega)$ -ből valók.

A lejjebb részletesen szemléltetett boxmódszer olyan  $y_h$  megoldásvektort nyújt, amely távolsága a pontos megoldás rácspontbeli értékeinek  $\vec{u}_h$  vektorától becsülhető mint

$$\|y_h - \vec{u}_h\|_{(1,h)} \leq M_1 h^{1+j/2}, \quad \|y_h - \vec{u}_h\|_{C(\omega_h)} \leq M_2 h^{1+j/2} \left( \ln \frac{1}{h} \right)^{1/2}, \quad j = 0, 1,$$

ahol még a  $j = 2$  eset is elérhető, ha  $\Gamma = \Gamma_1$ ,  $u \in H^3(\Omega)$  és a rács csak téglalapokból ill. egyenlőoldalú háromszögekből áll.

Az előbbi konvergencia eredményt nem fogjuk bebizonyítani. A bizonyítás Heinrich könyvében található és a következő lépésekből áll :

1. a differencia-egyenletek mátrixából kiindulva diszkrét Poincaré-féle egyenlőtlenségeknek (ilyen pl. II. 11.4.10-ben a (174) reláció) a levezetése;

2. a képlethiba elemzése és becslése a Bramble–Hilbert lemma (ld. II 11.7.) segítségével;

3. a  $z_h := y_h - \vec{u}_h$  vektor egyenletéből ezután adódik, 1-re és 2-re támaszkodva, a konvergencia becslés az  $(1, h)$ -normában;

4. az előzőből következik a becslés  $C(\omega_h)$ -ban „tilos” beágyazási tétel segítségével (ld. 15.5.3, (15.143)).

Arra is mutatunk rá, hogy e könyv három helyén szerepel a boxmódszer által létrehozott differenciasémák konvergencia vizsgálata: a Poisson-egyenletre (15.5.4-ben  $f \in L_2(\Omega)$ -ra, 15.5.5-ben  $f \in H^{-1-\varepsilon}(\Omega)$  esetére), és a fenti (15.163) egyenletre 15.7.9-ben (összehasonlítás a Courant-végeselemes módszerrel,  $f \in L_2(\Omega)$  ill.  $f \in H^1(\Omega)$  esetén).

### 15.6.2 A diszkretizáció levezetése

Célunk ebben a pontban mindenekelőtt az, hogy a differencia-egyenleteknek a gyakorlati levezetését részletesen megmutassuk. A \*\*\* 17.4. pontban ezt folytatjuk az olyan egyenlet esetén, amely (15.163)-hoz képest még az  $u$  első deriváltjait is tartalmazza  $\vec{v} \cdot \text{grad } u$  alakjában, és \*\*\* 18.10.3-ban elsőrendű hiperbolikus egyenletet közelítünk a boxmódszer segítségével.

Első lépésként a tartományt diszkretizáljuk egy téglalap alakú, nem ekvidisztáns rács segítségével, amelynek egyenesei a koordináta irányokkal párhuzamosak :  $x_1 = x_{1i}$ ,  $x_2 = x_{2j}$ . A rács tipikus pontját  $x^0 = (x_{1i}, x_{2j})$ -vel jelöljük.

Úgy választjuk meg a rácsot, hogy a perem minden töréspontja, a peremfeltétel típusa változásának ((15.164)-ben a  $\Gamma_1$  és  $\Gamma_{23}$  közötti átmenet) minden pontja, valamint  $\Gamma$  minden metszéspontja a rács egyenesével maga is rácspontra legyen. Ez azt jelenti, hogy a rács elég finom kell, hogy legyen.

A kapott rács a tartomány belsejét „elemi” téglalapokra osztja fel, míg a perem mentén téglalapok helyett derékszögű háromszögek is lehetnek. (Más elemi tartományok is lehetségesek, de éppen az említettek igen kényelmesek.) A rács egyenesei közötti távolságokat a következőképpen jelöljük :

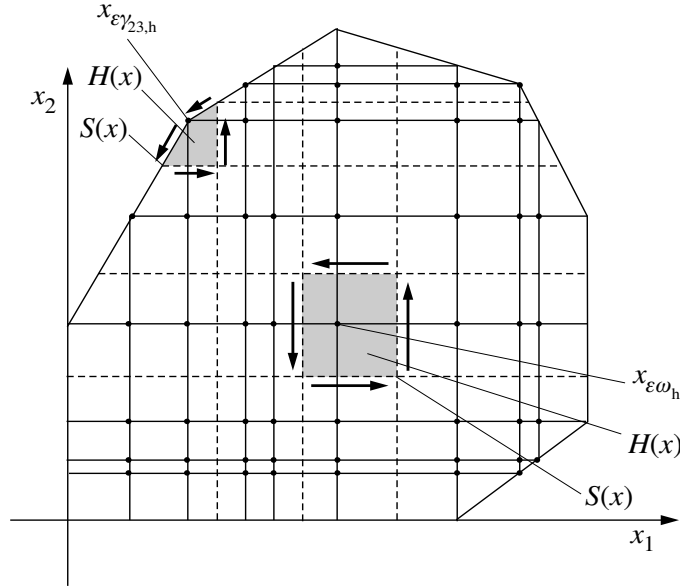
$$h_{1,i\pm 1/2} = h_{1i}^{\pm} := |x_{1,i\pm 1} - x_{1i}|, \quad h_{2,j\pm 1/2} = h_{2j}^{\pm} := |x_{2,j\pm 1} - x_{2j}|.$$

A maximális ilyen távolság legyen  $h$ . Azt követeljük meg, hogy a rács ne legyen túl irreguláris a következő értelemben. A rács tetszőleges  $h_{1i}^{\pm}$  ill.  $h_{2j}^{\pm}$  lépéstávolsága elégítse ki a

$$0 < \frac{h}{\kappa} \leq h_{1i}^{\pm}, h_{2j}^{\pm} \leq \kappa h \quad (15.167)$$

feltételt, ahol  $\kappa (\geq 1)$  alkalmas konstans. A peremmenti háromszögekben ebből a feltételből közvetlenül a

$$\frac{1}{\kappa^2} \leq \tan \vartheta \leq \kappa^2$$



ábra 15.12: Véges térfogatok

egyenlőtlenség adódik a háromszögek minden nem derékszögű  $\vartheta$  belső szögére. Ez az eredeti ( $\Gamma$ -ra vonatkozó, nulla szögeket kizáró) szögfeltétel pontosításának tekinthető. A rács minden  $\bar{\Omega}$ -ba ill.  $\Omega$ -ba eső pontjának a halmazát itt is  $\bar{\omega}_h$ -val ill.  $\omega_h$ -val jelöljük. A

$$\gamma_{23,h} := \Gamma_{23} \cap \bar{\omega}_h$$

perempontok az approximáció és az algoritmusok szempontjából ugyanolyan kezelést kívánnak, mint az  $\omega_h$ -beliek, ezért többször szerepel a  $\omega_h \cup \gamma_{23,h}$  halmaz is.

A következő módon, az elemi téglalapok oldalfelezőinek a segítségével a differenciálegyenlet diszkretizációjának előkészítésére minden  $x^0 \in \omega_h \cup \gamma_{23,h}$  pont körül kijelölünk egy  $H(x^0)$  „elemi” tartományt, a boxot, cellát avagy *véges térfogatot*: Ezekkel a  $H(x^0)$  elemi tartományokkal létrejön az  $\Omega$ -nak egy másodlagos felosztása; az új, második rács vonalait a  $H(x^0)$  tartományok  $S(x^0)$  peremei képezik.

A differenciálegyenlet diszkretizációja most a következőképpen történik. Adott  $x \in \omega_h \cup \gamma_{23,h}$ -hoz az egyenletet  $H(x^0)$  felett integráljuk, a második deriváltakat tartalmazó tagokat a Gauss–Osztrogradszkij tétel segítségével alakítjuk át:

$$\int_{H(x^0)} \operatorname{div}(k \operatorname{grad} u) \, dx = \int_{S(x^0)} k \operatorname{grad} u \cdot n \, ds = \int_{S(x^0)} k \frac{\partial u}{\partial n} \, ds,$$

ahol az  $S(x^0)$  perem az óramutató járásával ellenkezőleg irányított,  $n$  az  $S$  külső normálvektora. Így adódik a  $H(x^0)$ -hoz tartozó *mérlegegyenlet*, amely azt fejezi ki, hogy a  $H(x^0)$ -ban előállított hőmennyiség egyensúlyban van a be- és kiáramló hőmennyiséggel :

$$\int_{S(x^0)} k \frac{\partial u}{\partial n} ds + \int_{H(x^0)} f dx = 0. \quad (15.168)$$

Ezen a helyen már tudunk arra a kérdésre válaszolni, amely a pont elején is felvetődhetett : minek még egy diszkretizációs módszert tárgyalni, amikor már differencia-módszer és végeelem módszer is van és bármely ilyen módszer végülis csak egy többé-kevésbé különböző út a differenciálegyenlettől az algebrai egyenletrendszerhez?

A boxmódszernek mérnöki szempontból előnyei vannak.

A (15.168) egyenlet felírásához rögtön hozzá lehet fogni, nem kell először a differenciálegyenletet levezetni, aztán azt diszkretizálni. Ehelyett elegendő a folyamat leírásához lényeges fluxusokat ismerni (itt például  $-k \frac{\partial u}{\partial n}$  a hőfluxus) és azokkal a megmaradási tételt felírni egy elemei térfogatra. Így a módszer könnyen értelmezhető, és alkalmazási területe korántsem a (15.163) egyenlet peremérték feladataira korlátozódik, hiszen csak arra van szükség, hogy a differenciálegyenlet valamilyen fluxus divergenciáját tartalmazza és adott forrásokat. Éppen ilyen formában a fizika alapvető megmaradási tételei felírhatók. Ezért pl. a Navier-Stokes (17.4.0., 327. o.) és az elsőrendű hiperbolikus (\*\*\*) (18.10.3, 453. o.) egyenletek is kezelhetők vele (ez a végeelem módszernek nem erős oldala – az pedig a negyedrendű feladatoknál, variációs elv jelenlétében felülmúlhatatlan).

Ami viszont a feladat adatainak szükséges simaságát illeti, szeretnénk arra rámutatni, hogy 15.5.4-5-ben éppen a boxmódszer szolgálta a differenciasémát abban az esetben, amikor a forrás csak  $L_2$ -ből, ill.  $H^{-1-\varepsilon}$ -ből való volt.

Vegyük észre végül azt is, hogy a boxmódszer kiindulási pontját képező lokális megmaradási tételekből következik a globális is : összeadva két  $H_1, H_2$  szomszédos elemi tartományra a hozzátartozó (15.168) egyenletet, az  $S_1 \cap S_2$  közös, belső peremszakasz feletti peremintegrál kiesik a csak az előjelben különböző  $\frac{\partial u}{\partial n}$  értékek miatt. Közben a forrástag feletti integrál már  $H_1 \cup H_2$ -re vonatkozik. Ez a megmaradási tulajdonság továbböröklődik a diszkrét rendszerre is, ha csak az  $S$  perem szakaszain az integrálokat úgy helyettesítjük kvadratúra képlettel, hogy az független legyen attól, melyik elemi tartományhoz tartozik az illető szakasz. Közben az elemi tartományok nemcsak téglalapok lehetnek, hanem akár ferde vagy görbevonalú négyszögek, háromszögek, hatszögek stb., és háromdimenziós feladatokra közvetlenül adódik az általánosítás.

A (15.168) reláció hasznos a  $\psi$  képlethiba elemzése során is (amelyet fent a konvergencia vizsgálat második lépéséként említettünk). Ehhez (15.168)-at

$$\begin{aligned} (M(u))(x) &= (\varphi(f))(x), \quad x = x^0 \in \omega_h \cup \gamma_{23,h}; \\ (M(u))(x) &:= -\frac{1}{|H(x)|} \int_{S(x^0)} k \frac{\partial u}{\partial n} ds, \quad (\varphi(f))(x) := \frac{1}{|H(x)|} \int_{H(x)} f dx \end{aligned}$$

alakjában írjuk fel. Ha most  $A_h$  a (15.168) diszkrétizációja során (amire mindjárt ezután rátérünk)  $M(u)$ -ból létrehozott egyenletrendszer mátrixa és  $b_h$  annak  $(\varphi(f))$ -ből levezetett) jobboldali vektora, akkor

$$\begin{aligned} A_h z_h &= A_h y_h - A_h \vec{u}_h = b_h - A_h \vec{u}_h =: \psi_h \\ &= b_h - A_h \vec{u}_h + R_h(M(u) - \varphi(f)) \\ &= [R_h(M(u)) - A_h \vec{u}_h] + [b_h - R_h(\varphi(f))], \end{aligned}$$

ahol  $R_h v$  jelöli az  $\Omega$ -n definiált folytonos  $v$  függvény leszűkítését az  $(x_{1i}, x_{2j})$  rácspontokban kiszámított  $v$ -értékek vektorára, és  $R_h u =: \vec{u}_h$ .

A képlethiba itt bemutatott két része (amelynek első részében az  $u$ -nak már csak első deriváltja fordul elő, és az is integrál alatt) a Bramble–Hilbert lemma segítségével becsülhető a megfelelő Szoboljev–terekben.

A (15.168) reláció egzaktul érvényes, diszkrétizációjára több lehetőség van, ld. Versteeg és Malalasekra vagy LeVeque könyvét. Itt a relációban szereplő integrálok kvadratúra képletekkel való helyettesítésével foglalkozunk, oly módon, hogy differencia-kifejezéseket kapjunk. Az  $f$ -fel kapcsolatos integrált közelítsük egy pontos képlettel: az  $x^0 = (x_{1i}, x_{2j}) \in \omega_h \cup \gamma_{23,h}$  pontban vesszünk

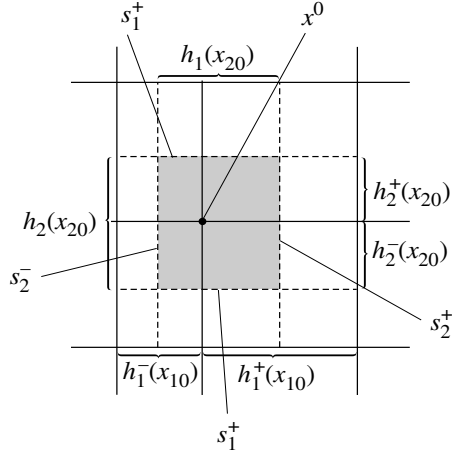
$$\int_{H(x^0)} f dx \approx |H(x^0)| f(x).$$

Amennyiben  $x^0 \in \gamma_{23,h}$ , akkor  $k \frac{\partial u}{\partial n}$  helyett a (15.164)  $\Gamma_{23}$ -ra vonatkozó peremfeltétel alapján vehetjük  $g_{23} - \sigma u$ -t, amikor az integrál  $S(x^0)$ -hoz tartozó részét approximáljuk. Ezután visszamarad az  $S(x^0) \cap \Omega$  menti integrál. Az  $S(x^0) \cap \Omega$  olyan  $s_1^\pm, s_2^\pm$  szakaszokból áll, amelyek az  $x_1, x_2$  koordinátairányokkal párhuzamosak, hosszúságuk

$$h_{1i} := \frac{1}{2}(h_{1i}^+ + h_{1i}^-), \quad \text{ill.} \quad h_{2j} := \frac{1}{2}(h_{2j}^+ + h_{2j}^-). \quad (15.169)$$

Az  $s_1^\pm$  szakasz felezi az  $x^0$  és  $(x_{1i}, x_{2,j\pm 1})$  rácspontokat összekötő szakaszt,  $s_2^\pm$  felezi az  $x^0$  és  $(x_{1,i\pm 1}, x_{2j})$  összekötő szakaszokat. Ezért kézenfekvő a következő approximációk használata:

$$\int_{s_1^\pm} k \frac{\partial u}{\partial n} ds \approx \pm h_{1i} (k v_{x_2})_{i,j\pm 1/2}, \quad \int_{s_2^\pm} k \frac{\partial u}{\partial n} ds \approx \pm h_{2j} (k v_{x_1})_{i\pm 1/2,j},$$



ábra 15.13:  $x^0 \in \omega_h$  környezete

ahol

$$k_{i\pm 1/2, j} := k \left( x_{1i} \pm \frac{h_{1i}^\pm}{2}, x_{2j} \right), \quad k_{i, j\pm 1/2} := k \left( x_{1i}, x_{2j} \pm \frac{h_{2j}^\pm}{2} \right),$$

$v$  az  $u(x)$  approximációja az  $x \in \omega_h$  rácspontokban, és a  $v$  diszkrét iránymenti deriváltjai

$$\pm v_{x_1, i\pm 1/2, j} := (v(x_1 \pm h_1^\pm, x_2) - v(x)) / h_1^\pm, \quad (15.170)$$

$$\pm v_{x_2, i, j\pm 1/2} := (v(x_1, x_2 \pm h_2^\pm) - v(x)) / h_2^\pm. \quad (15.171)$$

A  $v_{x_1, i\pm 1/2, j}$ ,  $v_{x_2, i, j\pm 1/2}$  elsőrendű differenciák helyett, valamint a  $k$  együttható “feles” pontokra vonatkozó értékei helyett rövidebben azt fogjuk írni, hogy  $v_{x_1, \pm}$ ,  $v_{x_2, \pm}$  és  $k_{i\pm, j}$ ,  $k_{i, j\pm}$ .

Írjuk ki részletesen az így kapott approximációt két speciális esetben!

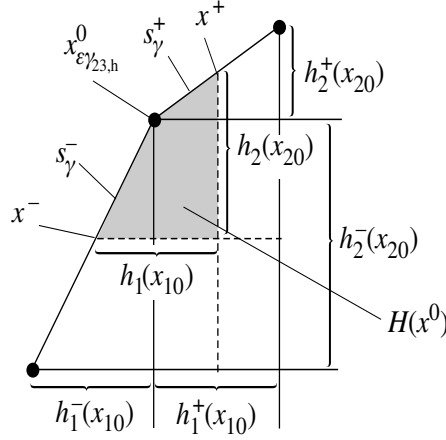
a)  $x^0 \in \omega_h$ , tehát  $x^0$  a tartomány belső pontja, és ott  $H_0 := H(x^0)$  téglalap :

$$h_{1i} [(kv_{x_2})_{i, j+1/2} - (kv_{x_2})_{i, j-1/2}] + h_{2j} [(kv_{x_1})_{i+1/2, j} - (kv_{x_1})_{i-1/2, j}] + |H_0| f_0 = 0,$$

és rövidebben

$$h_{1i} [(kv_{x_2})_{i, j+} - (kv_{x_2})_{i, j-}] + h_{2j} [(kv_{x_1})_{i+, j} - (kv_{x_1})_{i-, j}] + |H_0| f_0 = 0,$$

ahol  $|H_0| := h_{1i} h_{2j}$  és  $f_0 := f(x^0)$ . A  $|H_0|$ -ra való osztás után ez egy jól ismert differenciaképlet (ld. (15.133)-(15.134) 15.5.2-ben). A 36. feladatban annak



ábra 15.14:  $x^0 \in \gamma_{23,h}$  környezete

vizsgálata szerepel, amikor  $x^0$  egy (peremmenti) derékszögű háromszögnek is csúcsa.

b)  $x^0 \in \gamma_{23,h}$ , és  $\Gamma_{23}$  része az  $x^+$ ,  $x^0$ ,  $x^-$  pontokat összekötő töröttvonal (legyen  $x^\pm := (x_{1i} \pm \frac{h_{1i}^\pm}{2}, x_{2j} \pm \frac{h_{2j}^\pm}{2})$ ), míg  $(x_{1,i+1}, x_{2j})$ ,  $(x_{1i}, x_{2,j-1}) \in \omega_h$  :

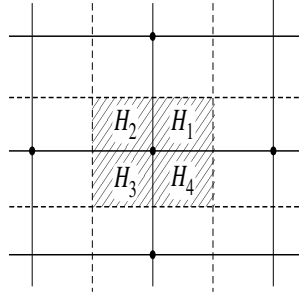
$$-h_{1i}(kv_{x_2})_{i,j-1/2} + h_{2j}(kv_{x_1})_{i+1/2,j} + (|s_\gamma^+| + |s_\gamma^-|)(g_{23} - \sigma v)(x^0) + |H_0|f_0 = 0,$$

ahol

$$|s_\gamma^\pm| := \frac{1}{2} ((h_{1i}^\pm)^2 + (h_{2j}^\pm)^2)^{1/2}, \quad \text{ill.} \quad |H_0| := \frac{h_{1i}^+}{2} \left( \frac{h_{2j}^-}{2} + \frac{h_{2j}^+}{4} \right) + \frac{h_{1i}^- h_{2j}^-}{8}$$

a peremszakaszok hossza, ill. az elemi tartomány területe. (Azon eset vizsgálata, amikor  $x^0 \in \gamma_{23,h}$ -nál a belső szög  $\frac{\pi}{2}$ , a 35. feladat tárgyát képezi, ld. a 37.a feladatot is.)

Kézenfekvő az a) esetben kapott approximációt négy tagból álló összegként felfogni és mind a négy taghoz  $H_0$ -nak egy-egy részét hozzárendelni :



ábra 15.15: Elemi tartomány felosztása

$$H_{01} : \frac{h_{2j}^+ k_{i+,j}}{2} \frac{v_1 - v_0}{h_{1i}^+} + \frac{h_{1i}^+ k_{i,j+}}{2} \frac{v_2 - v_0}{h_{2j}^+} + \frac{h_{2j}^+ h_{1i}^+}{4} f_0 =: - (A^{(1)}v)_0 + |H_{01}|f_0,$$

$$H_{02} : \frac{h_{1i}^- k_{i,j+}}{2} \frac{v_2 - v_0}{h_{2j}^+} + \frac{h_{2j}^+ k_{i-,j}}{2} \frac{v_3 - v_0}{h_{1i}^-} + \frac{h_{2j}^+ h_{1i}^-}{4} f_0 =: - (A^{(2)}v)_0 + |H_{02}|f_0,$$

$$H_{03} : \frac{h_{2j}^- k_{i-,j}}{2} \frac{v_3 - v_0}{h_{1i}^-} + \frac{h_{1i}^- k_{i,j-}}{2} \frac{v_4 - v_0}{h_{2j}^-} + \frac{h_{2j}^- h_{1i}^-}{4} f_0 =: - (A^{(3)}v)_0 + |H_{03}|f_0,$$

$$H_{04} : \frac{h_{1i}^+ k_{i,j-}}{2} \frac{v_4 - v_0}{h_{2j}^-} + \frac{h_{2j}^- k_{i+,j}}{2} \frac{v_1 - v_0}{h_{1i}^+} + \frac{h_{1i}^+ h_{2j}^-}{4} f_0 =: - (A^{(4)}v)_0 + |H_{04}|f_0.$$

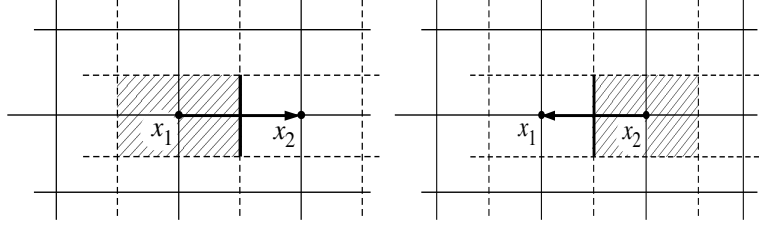
Ezeket a relációkat összefoglalva :

$$0 = \sum_{\ell=1}^4 (- (A^{(\ell)}v)_0 + |H_{0\ell}|f_0) =: - (Av)_0 + b_0, \quad b_0 = |H_0|f_0.$$

Tehát  $(A^{(\ell)}v)_0$ , ill.  $|H_{0\ell}|f_0$  értelmezhető mint a  $H_{0\ell}$  elemi résztartomány járuléka az  $Av = b$  egyenletrendszernek azon sorához, amely az  $x^0$  ponthoz tartozik. Az  $(Av)_0$  kifejezést úgy is lehetne előállítani, hogy a  $H_{0\ell}$ -kre írjuk fel a mérlegegyenletet, majd ennek a  $\ell = 1, 2, 3, 4$  értékekre kapott eredményeit összeadjuk : a  $H_{0\ell}$  „belső” ( $H(x^0)$ -beli) pereméhez tartozó mennyiségek kölcsönösen kioltják egymást, a belső peremszakaszok ellenkező átjárási iránya miatt. Eredményül  $f_0$  mellett megjelenik  $H(x^0)$  egész területe mint szorzó, de  $(A^{(\ell)}v)_0$  csak a  $H_{0\ell}$  résztartomány ( $\partial H_{0\ell} \cap S(x^0)$  külső peremével kapcsolatos. Hasonló módon értelmezhető a b) esetben kapott approximáció is.

Mindkét esetben szimmetrikus az  $A$  mátrix, hiszen két pont között csak egyidejűleg jön létre kapcsolat és a képletek csak előjelben különböznek.





ábra 15.16: A mátrix szimmetriája

Például :

$$x^0 := (x_{1i}, x_{2j}) \text{ és } x^1 := (x_{1,i+1}, x_{2j}) \text{ közötti kapcsolat : } h_{2j}k_{i+1/2,j} \frac{v_1 - v_0}{h_{1i}^+},$$

$$x^1 \text{ és } x^0 \text{ közötti kapcsolat : } h_{2j}k_{i+1/2,j} \frac{v_0 - v_1}{h_{1i}^+}.$$

Ekkor tehát  $a_{12} = -h_{2j}k_{i+1/2,j}/h_{1i}^+ = a_{21}$ .

Ha az approximációt  $|H|$ -ra osztjuk, akkor általában nonszimmetrikus lesz a mátrix, de akkor is szimmetrikus az

$$(y, v)_{(0,h)} := \sum_{x \in \omega_h \cup \gamma_{23,h}} y(x)v(x)|H(x)| \quad (15.172)$$

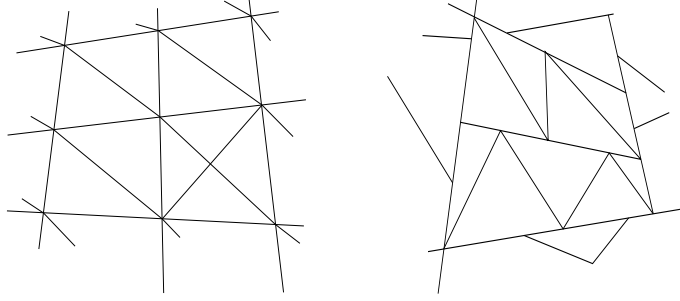
skalárszorzat értelmében (amely (15.43) kézenfekvő általánosítása).

Az itt javasolt rácserkesztés téglalapokból és peremmenti háromszögekből oda vezet, hogy az összes képlet könnyen áttekinthető lesz; az egyszerűbb struktúra révén még az eredő lineáris egyenletrendszer megoldása is könnyebb.

### 15.6.3 A véges térfogat módszer háromszögek esetén

Annak érdekében, hogy a boxmódszerben rejlő lehetőségeket jobban szemléltessük, és mivel a téglalapok esetén ismert differencia képletekre jutottunk, most azt mutatjuk meg, hogyan kell a diszkretizációt levezetni, ha a résztartományok háromszögek. Ehhez az  $\Omega$ -t véges sok háromszög összegeként állítjuk elő úgy, hogy tetszőleges háromszög csúcsa nem lehet más háromszög belsejében, és peremén csak úgy, ha ott csúccsal esik egybe. Ilyen felbontást az  $\Omega$  *konform triangulációjának* is hívjuk. Ekkor elég egyetlen,  $\Delta$ -val jelölt háromszöget vizsgálunk, amelynek csúcsai legyenek az  $x^\ell = (x_{1\ell}, x_{2\ell})$ ,  $\ell = 0, 1, 2$  pontok. Oldalait  $s_{01}$ -gyel,  $s_{12}$ -vel, ill.  $s_{20}$ -val jelöljük. Természetesen azt tesszük fel, hogy a háromszög nem elfajult :

$$|\Delta| := \frac{1}{2} \left| \det \begin{pmatrix} x_{11} - x_{10} & x_{12} - x_{10} \\ x_{21} - x_{20} & x_{22} - x_{20} \end{pmatrix} \right| \neq 0.$$



ábra 15.17: Konform és nemkonform trianguláció

Erre a  $\Delta$  háromszögre a mérlegegyenletet írjuk fel :

$$\int_D k \frac{\partial u}{\partial n} ds + \int_{\Delta} f dx = 0, \quad D := \partial\Delta, \quad (15.173)$$

és ezt approximáljuk. Ehhez mindenekelőtt az  $u(x)$  gradiensét kell közelítenünk, amire a középérték tételt és a parciális integrálást használjuk :

$$\text{grad } u \approx \frac{1}{|\Delta|} \int_{\Delta} \text{grad } u dx = \frac{1}{|\Delta|} \int_D un ds.$$

Itt  $n = n(x)$  a háromszög  $D$  peremének a külső normálvektora. A kapott integrált a trapézsabály segítségével approximáljuk :

$$\text{grad } u \approx \frac{1}{|\Delta|} \int_D un ds \approx \frac{1}{|\Delta|} \left[ \frac{v_0+v_1}{2} h_1 n^{(1)} + \frac{v_1+v_2}{2} h_3 n^{(3)} + \frac{v_2+v_0}{2} h_2 n^{(2)} \right],$$

ahol  $v_\ell \approx u(x^\ell)$ ,  $\ell = 0, 1, 2$ , és  $h_\ell n^{(\ell)}$  a háromszög  $x^0$ -t és  $x^\ell$ -t összekötő  $s_{01}$  ill.  $s_{20}$  oldalához tartozó normálvektorának a szorzata az oldal hosszával,  $\ell = 1, 2$  :

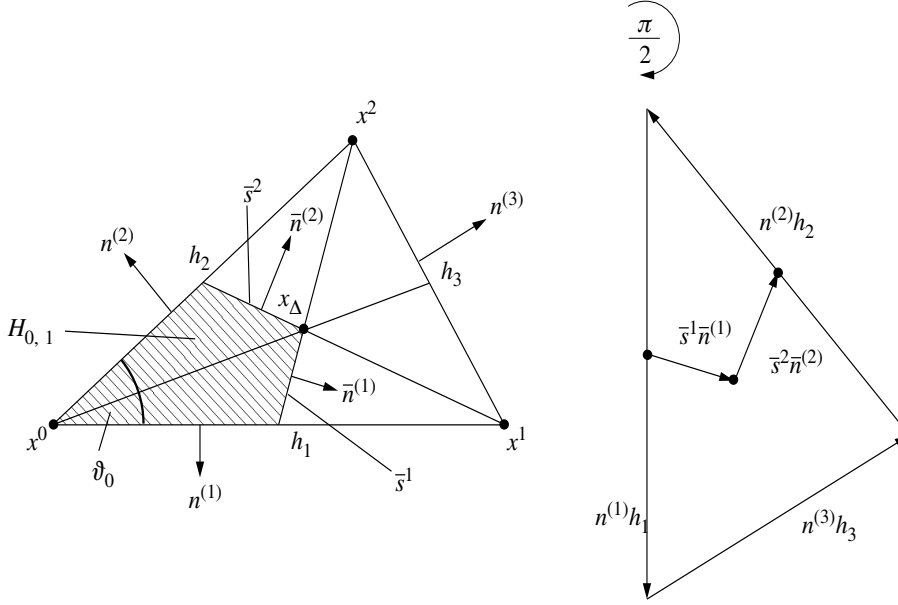
$$h_1 n^{(1)} := (x_{21} - x_{20}, -(x_{11} - x_{10}))^T, \quad h_2 n^{(2)} := (x_{20} - x_{22}, -(x_{10} - x_{12}))^T,$$

míg  $h_3 n^{(3)}$  az  $(x_{10}, x_{20})$ -val szemben lévő  $s_{12}$  oldalhoz tartozik,

$$h_3 n^{(3)} := (x_{22} - x_{21}, -(x_{12} - x_{11}))^T.$$

Felhasználva azt, hogy

$$0 = \int_D n ds = h_1 n^{(1)} + h_2 n^{(2)} + h_3 n^{(3)}, \quad (15.174)$$



ábra 15.18: Háromszöges térfogat : jölések

(arra is hivatkozhatunk, hogy az  $n^{(\ell)}$ -irányú oldalak a  $90^\circ$  fokkal elforgatott háromszöget alkotják) végülis

$$\begin{aligned} \text{grad } u &\approx \frac{1}{2|\Delta|} [(v_0 - v_2)h_1n^{(1)} + (v_0 - v_1)h_2n^{(2)}] \\ &= \frac{1}{\sin \vartheta_0} \left[ \frac{v_0 - v_2}{h_2}n^{(1)} + \frac{v_0 - v_1}{h_1}n^{(2)} \right] =: \text{grad}_h v \quad (15.175) \end{aligned}$$

adódik, ahol  $\vartheta_0 = \vartheta(x^0)$  az  $x^0$ -nál bezárt szög (tehát  $\sin \vartheta_0 = 2|\Delta|/(h_1h_2)$ ).

Az  $x^0$  pontot több háromszög veszi körül ( $x^0 \in \Omega$  esetén legtöbbször hat ilyen háromszög lesz), és mindegyiknek egy részéből az  $x^0$ -hoz hozzárendelt  $H_0$  elemi tartományt képezzük.

Ilyen részt a fenti,  $x^\ell$  csúcsú háromszögben ( $\ell = 0, 1, 2$ ) a következőképpen jelöljük ki.

Legyen  $x^\Delta$  a háromszög súlypontja és  $\bar{s}^\ell$  azok a szakaszok, amelyek  $x^\Delta$ -t és az  $s_{01}$  ill.  $s_{20}$  oldalak felezőpontjait kötik össze,  $\ell = 1, 2$ . Ekkor  $H_{0,1}$  az  $x^0$  és  $x^\Delta$  közti négyszög, ld. a 16. ábrát. Az  $x^0$ -t körülvevő  $\Delta_j$  további háromszögek megfelelő  $H_{0,j}$  részei képezik együtt a  $H_0$ -t.

További jelölések:  $\bar{n}^{(\ell)}$  az  $\bar{s}^\ell$  szakasz ( $H_{0,1}$  szempontjából) külső normálvektora,  $\ell = 1, 2$ ,  $\bar{s} := \bar{s}^1 \cup \bar{s}^2$ .

Ekkor (15.173) és (15.175) alapján a  $H_{0,1}$ -hez tartozó járulékot a követ-

kezőképpen kapjuk meg :

$$\begin{aligned} \int_{\bar{s}} k \frac{\partial u}{\partial n} ds + \int_{H_{0,1}} f dx dy &= \int_{\bar{s}} k \text{grad } u \cdot n ds + \int_{H_{0,1}} f dx dy \\ &\approx k(x^\Delta) \text{grad}_h v \cdot \int_{\bar{s}} n ds + |H_{0,1}| f_0 \\ &= \frac{k(x^\Delta)}{2|\Delta|} [(v_0 - v_2)h_1 n^{(1)} + (v_0 - v_1)h_2 n^{(2)}] \cdot \int_{\bar{s}} n ds + |H_{0,1}| f_0. \end{aligned}$$

Itt az

$$\int_{\bar{s}} n ds = \int_{\bar{s}^1 \cup \bar{s}^2} n ds = |\bar{s}^1| \bar{n}^{(1)} + |\bar{s}^2| \bar{n}^{(2)} = \frac{1}{2} h_3 \bar{n}^{(3)}$$

képletet behelyettesítve (amely összefüggés onnan következik, hogy a fent említett,  $90^\circ$ -kal elforgatott háromszögben a  $|\bar{s}^1| \bar{n}^{(1)} + |\bar{s}^2| \bar{n}^{(2)}$  vektor az  $s_{01}$  és  $s_{20}$  oldalak középpontjait köti össze), a  $\Delta$  háromszög keresett járuléka áll elő :

$$(A(\Delta)v)_0 = -\frac{k(x^\Delta)}{4|\Delta|} [(v_0 - v_2)h_1 n^{(1)} + (v_0 - v_1)h_2 n^{(2)}] \cdot h_3 n^{(3)}. \quad (15.176)$$

Ez tehát a járulék az  $x^0$  sorához az  $A$  mátrixban, míg  $|H_{0,1}| f_0 = \frac{|\Delta|}{3} f_0$  a járulék a jobboldali vektor  $x^0$ -hoz tartozó komponenséhez.

Felhasználva, hogy  $\Delta$ -ban minden mennyiség  $h_1$ -gyel,  $h_2$ -vel és  $\vartheta_0$ -val meghatározott, és pl.

$$n^{(1)} \cdot h_3 n^{(3)} = h_2 \cos \vartheta_0 - h_1, \quad n^{(2)} \cdot h_3 n^{(3)} = h_1 \cos \vartheta_0 - h_2,$$

(15.176) a következő alakba is írható át :

$$(A(\Delta)v)_0 = -\frac{k(x^\Delta)}{2 \sin \vartheta_0} \left[ \frac{v_0 - v_2}{h_2} (h_2 \cos \vartheta_0 - h_1) + \frac{v_0 - v_1}{h_1} (h_1 \cos \vartheta_0 - h_2) \right].$$

$\Delta = \Delta(x^0)$  háromszögünket most  $\Delta_{01}$ -val jelöljük, és figyelembe vesszük, hogy  $x^0$  összesen  $n_0 = n_0(x^0)$  háromszög csúcsa. Ezeket  $\Delta_{0j} = \Delta_j(x^0)$ -val jelöljük,  $1 \leq j \leq n_0$ .

Az  $x^0$ -hoz tartozó sort a lineáris rendszerben a (15.176) szerinti járulékokból adjuk össze :

$$(Av)_0 = \sum_{j=1}^{n_0} (A(\Delta_{0j})v)_0 = \sum_{j=1}^{n_0} \frac{|\Delta_{0j}|}{3} f_0 =: |H_0| f_0. \quad (15.177)$$

Ezt a sort lehetne még (a differencia módszer hagyományát követve)  $|H_0|$ -val osztani. Az akkor eredő mátrixot  $A_h$ -val fogjuk jelölni.

Számítsuk ki ezeket a fenti kifejezéseket és az adódó approximációt arra a speciális esetre, hogy az összes háromszög szabályos (a derékszögű háromszög vizsgálata a 36. feladat), az oldalak hossza  $h$  és egyik oldaluk párhuzamos az  $x_1$ -tengellyel! Ekkor

$$\begin{aligned} \vartheta &= \frac{\pi}{3}, & \sin \vartheta &= \frac{1}{2}\sqrt{3}, & |H_{0,1}| &= \frac{\sqrt{3}}{12}h^2, \\ n^{(1)} &= (0, -1)^T, & n^{(2)} &= \left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right)^T, & x^\Delta &= \left(\frac{1}{2}, \frac{\sqrt{3}}{6}\right)h + x^0, \\ \bar{n}^{(1)} &= (1, 0)^T, & \bar{n}^{(2)} &= \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)^T, & |\bar{s}^1| &= |\bar{s}^2| = \frac{\sqrt{3}}{6}h. \end{aligned}$$

Így a háromszög járuléka az approximációhoz

$$\frac{2}{\sqrt{3}}k(x^\Delta) \left[ \frac{v_0 - v_2}{h} \left(-\frac{h}{4}\right) + \frac{v_0 - v_1}{h} \left(-\frac{h}{4}\right) \right] + \frac{\sqrt{3}}{12}h^2 f_0.$$

Hasonló járulékok tartoznak a további háromszögekhez, és pl. a  $k \equiv 1$  esetben, tehát amikor (15.163) a Poisson-egyenlet, az  $x^0$ -hoz tartozó (15.177) approximáció

$$\frac{h}{\sqrt{3}} \sum_{j=1}^6 \frac{v_j - v_0}{h} + \frac{\sqrt{3}}{2}h^2 f_0 = 0,$$

vagyis a lineáris egyenletrendszernek az  $x^0$  ponthoz tartozó sora (ahol most a  $|H_0|$ -val osztottunk) :

$$(A_h v)_0 = (-\Delta_h v)_0 := -\frac{2}{3h} \sum_{j=1}^6 \frac{v_j - v_0}{h} = f_0. \quad (15.178)$$

Taylor-sorfejtéssel beláthatjuk, hogy ezen képlet hibája másodrendű. Ha viszont  $f_0$  helyett  $f_0 + \frac{h^2}{16}(\Delta f)_0$ -t vesszük, akkor az approximáció hibája negyedrendű. Ebben a képletben a  $(\Delta f)_0$  helyettesíthető a  $(\Delta_h f)_0$ -val.

A (15.178) képlet egyik általánosítása a 37.b feladat. A másik az, ha az adott rácspontot nem 6, hanem  $n$  pont veszi körül szimmetrikusan :

$$(-\Delta_h v)_0 = -\frac{4}{nh} \sum_{j=1}^n \frac{v_j - v_0}{h}. \quad (15.179)$$

A Laplace-operátornak ezen másodrendű approximációját előnyösen használhatjuk *polárkoordinátákban* megfogalmazott feladatoknál az origóban.

### 15.6.4 A véges térfogat módszer mátrixai

Térjünk vissza az általános, nem elfajult  $\Delta = \Delta(x^0)$  háromszöghöz! Ami minket elsősnek érdekel, az az  $A$  mátrix szimmetriája, amely a (15.176) képletből nem azonnal látható (de a szabályos háromszögre vonatkozó (15.178) approximáció esetén megvan). Ezért írjuk fel ugyanazon  $\Delta(x^0)$  háromszög járulékait a többi vele kapcsolatos mátrixsorokhoz! Az  $x^1$  sorába tartozó járulék :

$$(A(\Delta)v)_1 = -\frac{k(x^\Delta)}{4|\Delta|} [(v_1 - v_0)h_3n^{(3)} + (v_1 - v_2)h_1n^{(1)}] \cdot h_2n^{(2)}, \quad (15.180)$$

míg az  $x^2$  sorába tartozik

$$(A(\Delta)v)_2 = -\frac{k(x^\Delta)}{4|\Delta|} [(v_2 - v_1)h_2n^{(2)} + (v_2 - v_0)h_3n^{(3)}] \cdot h_1n^{(1)}. \quad (15.181)$$

Most közvetlenül ellenőrizhető, hogy pl.  $(A(\Delta)v)_0$ -ban a  $v_2 = v(x^2)$  szorzója

$$\frac{k(x^\Delta)}{4|\Delta|} h_1n^{(1)} \cdot h_3n^{(3)},$$

míg  $(A(\Delta)v)_2$ -ben a  $v_0 = v(x^0)$  szorzója

$$\frac{k(x^\Delta)}{4|\Delta|} h_3n^{(3)} \cdot h_1n^{(1)}.$$

A többi megfelelő szorzó is egyenlő, és így szimmetrikus az ezen szorzók segítségével definiált (15.177) mátrix. Ha még  $|H_0|$ -val osztjuk az  $x^0$  sorát, akkor keletkezik az  $A_h$  mátrix és az szimmetrikus lesz a (15.172) skalárszorzat értelmében, ha ott  $\gamma_{23,h} = \emptyset$  :

$$(A_h y, v)_{(0,h)} = (y, A_h v)_{(0,h)}, \quad \text{ahol } (y, v)_{(0,h)} := \sum_{x \in \omega_h} y(x)v(x)|H(x)|.$$

Az  $A_h$  és  $A$  közti összefüggés alapján az  $(\cdot, \cdot)$  euklideszi skalárszorzattal teljesül

$$(A_h y, v)_{(0,h)} = \sum_{x \in \omega_h} (A_h y)(x)v(x)|H(x)| = \sum_{x \in \omega_h} (A y)(x)v(x) = (A y, v).$$

Ha harmadfajú a peremfeltétel, akkor még egy peremmenti összeg is szerepel a (15.172) skalárszorzatban, v.ö. II. 189. o., (102)-vel is.

A (15.176), (15.180), (15.181) képletek alapján az  $A$  valamint  $A_h$  mátrixok pozitív definitiségének feltételét is tudjuk levezetni. Ehhez is elegendő a

$\Delta = \Delta(x^0)$  háromszöggel foglalkozni. Mint az összes belső pontra kiterjedő  $(A_h v, v)_{(0,h)}$  összeg részét tekintjük a következő kifejezést, amely tartalmazza a háromszög  $x^\ell$  csúcaiban felvett  $v^\ell$  értékeket,  $\ell = 0, 1, 2$  :

$$\begin{aligned} S(\Delta) &:= \sum_{\ell=0}^2 (A(\Delta)v)_\ell v_\ell \\ &= -\frac{k(x^\Delta)}{4|\Delta|} \left\{ [(v_0 - v_2)h_1n^{(1)} + (v_0 - v_1)h_2n^{(2)}] \cdot h_3n^{(3)}v_0 \right. \\ &\quad + [(v_1 - v_0)h_3n^{(3)} + (v_1 - v_2)h_1n^{(1)}] \cdot h_2n^{(2)}v_1 \\ &\quad \left. + [(v_2 - v_1)h_2n^{(2)} + (v_2 - v_0)h_3n^{(3)}] \cdot h_1n^{(1)}v_2 \right\}. \end{aligned}$$

Egyszerű átrendeződéssel azt kapjuk, hogy

$$\begin{aligned} S(\Delta) &= -\frac{k(x^\Delta)}{4|\Delta|} \left\{ (v_0 - v_2)^2 h_1n^{(1)} \cdot h_3n^{(3)} \right. \\ &\quad \left. + (v_0 - v_1)^2 h_2n^{(2)} \cdot h_3n^{(3)} + (v_1 - v_2)^2 h_1n^{(1)} \cdot h_2n^{(2)} \right\} \\ &= \frac{k(x^\Delta)}{2} \left\{ (v_0 - v_2)^2 \cot \vartheta_1 + (v_0 - v_1)^2 \cot \vartheta_2 \right. \\ &\quad \left. + (v_1 - v_2)^2 \cot \vartheta_0 \right\}. \end{aligned} \quad (15.182)$$

Itt használtuk, hogy érvényes

$$\begin{aligned} \cot \vartheta_0 &= -\frac{h_1n^{(1)} \cdot h_2n^{(2)}}{2|\Delta|}, \quad \cot \vartheta_1 = -\frac{h_1n^{(1)} \cdot h_3n^{(3)}}{2|\Delta|}, \\ \cot \vartheta_2 &= -\frac{h_2n^{(2)} \cdot h_3n^{(3)}}{2|\Delta|}. \end{aligned} \quad (15.183)$$

(15.182)-ből az olvasható le, hogy  $0 < \vartheta_\ell < \frac{\pi}{2}$  biztosítja a pozitív definitiséget változó  $k$  együttható mellett is. Amikor viszont  $k = \text{const}$ , akkor figyelembe vehetjük, hogy az egész rácsra kiterjedő összegzés során a másik, az  $x^0$  és  $x^2$  pontokat összekötő  $s_{20}$  oldalra illeszkedő háromszög is sorra kerül. Annak  $s_{20}$ -val szemközti szöge legyen  $\vartheta_3$ . Ekkor  $(v_0 - v_2)^2$  szorzója végülis

$$\frac{k}{2} (\cot \vartheta_1 + \cot \vartheta_3).$$

Ez a pozitív  $k$  mellett akkor pozitív, ha  $0 < \vartheta_1 + \vartheta_3 < \pi$ , mert ekkor (mivel  $(0, \pi)$ -ben a  $\cot$ -függvény monoton csökkenő)  $\cot \vartheta_1 > \cot(\pi - \vartheta_3) = -\cot \vartheta_3$ .

Az általános feltétel ezután az, hogy minden belső  $s_\ell$  oldallal szemközti két szög,  $\vartheta_\ell^+$  és  $\vartheta_\ell^-$ , teljesítse a

$$0 < \vartheta_\ell^+ + \vartheta_\ell^- < \pi \quad (15.184)$$

feltételt. A másod- vagy harmadfajú peremfeltétel esetén vannak olyan háromszögek, amelyeknek egyik oldala a peremvonal része. Ekkor a szemközti szög (mivel nincsen párja) kell, hogy  $(0, \frac{\pi}{2})$ -ből való legyen.

Annak érdekében, hogy a pozitív definitéség  $h$ -ban egyenletes legyen (azaz  $\lambda_{\min}(A_h) \geq \text{const} > 0$ , amikor  $h \rightarrow 0$ ), (15.184) helyett az olyan  $0 < \vartheta^* < \frac{\pi}{4}$  szög létezését kell követelnünk (azaz a rácssorozatot úgy kialakítanunk), hogy igaz legyen

$$0 < \vartheta^* \leq \vartheta_\ell^+ + \vartheta_\ell^- \leq \pi - \vartheta^* \quad (15.185)$$

minden belső oldallal szemközti szög esetén, ill.

$$0 < \vartheta^* \leq \vartheta_\ell \leq \frac{\pi}{2} - \vartheta^* \quad (15.186)$$

minden perementi oldallal szemközti szög esetén.

Következőnek vegyük észre, hogy (15.183) és (15.176), (15.180), (15.181) alapján az  $A$  mátrix elemeinek előjelei megfelelnek az M-mátrixok előjel-eloszlásának, és

$$Ae \geq 0 \quad (e := (1, \dots, 1)^T) \quad (15.187)$$

is érvényes, ha

$$0 \leq \vartheta_\ell \leq \frac{\pi}{2}.$$

(15.187)-ben a szigorú egyenlőtlenség a harmad- és elsőfajú perempontokra vonatkozik, míg az egyenlőség a belső pontokra, ill. másodfajú perempontokra.

Emlékeztetünk arra (ld. 15.3.2-ben a 15.2. lemmát), hogy amennyiben a diszkretizáció mátrixa M-típusú, akkor pl. az alkalmazások szempontjából fontos tulajdonság megvan, hogy a nemnegativitás megőrződik : nemnegatív jobboldal esetén nemnegatív megoldást kapunk.

Az előjeleloszlás nem elegendő ahhoz, hogy a mátrix valóban M-típusú legyen. Ezt biztosíthatja pl. a mátrix irreducibilitása. (Annak definíciója I. 3.7.3-ban található.)

Az irreducibilitás akkor megvan, ha (a boxmódszer 15.6.2.-ben leírt verziójának megfelelően) csak a perem mentén vannak (derékszögű) háromszögek. Amennyiben a tartományt tetszőleges háromszögekre bontottuk fel (15.6.3. pont), akkor (15.176)-(15.183) szerint az irreducibilitás elégséges feltétele az, hogy a háromszögek bármely belső szöge kisebb legyen  $\pi/2$ -nél.

Mutassuk meg most azt, hogy (15.187)-ből, az

$$a_{ij} \leq 0, \quad 1 \leq i \neq j \leq n \quad (15.188)$$

feltételből, valamint az  $A = (a_{ij}) \in \mathbb{R}^n$  irreducibilitásából következik M-mátrix tulajdonsága!



Ha  $A$  irreducibilis, akkor (15.187) mellett reguláris is, ld. az 1. feladatot, továbbá, ekkor  $A + \varepsilon I$  M-mátrix minden  $\varepsilon > 0$ -ra, azaz  $(A + \varepsilon I)^{-1} \geq 0$  elemenként. De ez, folytonosság révén,  $A^{-1} \geq 0$ -t vonja maga után.

Ez már, (15.188) mellett, az M-mátrix ekvivalens definíciója. Ugyanis ekkor egy majoráns vektor azonnál megadható : vegyük az  $Ag = e$  megoldását! Így  $Ag > 0$  biztosított, és  $g > 0$  komponensenként, ld. I. 52. o., 3. megjegyzés. Ezután viszont (ld. megint az 1. feladatot) a 15.13. tétel szerint még a maximumelv is érvényes.

**Megjegyzések. 1.** A-priori-becslést nem kaptunk, így az sem világos, vajon a boxmódszer stabil-e, azaz  $A_h$  mátrixa eleget tesz-e  $\|A_h^{-1}\| \leq \text{const}$  függetlenül  $h$ -tól? De a stabilitás a pont elején említett diszkrét Poincaré-féle egyenlőtlenség és a mátrix fent levezetett pozitív definitéseg segítségével belátható.

**2.** (15.185) és (15.186) mellett az  $A$  mátrix elemei egyenletesen korlátozottak  $h_1, h_2 \rightarrow 0$  esetén, és vagy  $a_{ij} = 0$  vagy  $\kappa^{-2}k_0 \leq a_{ij} \leq k_1\kappa^2$  ld. (15.165)-(15.166)-ot és (15.167)-et. Ennek megfelelően az  $A_h$  mátrix elemeinek nagyságrendje  $O(h^{-2})$ .

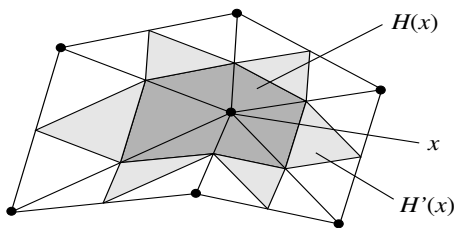
**3.** Ha a (15.173) mérlegegyenletben a  $\int_D k \frac{\partial u}{\partial n} ds$  tagot úgy approximáljuk, hogy  $\frac{\partial u}{\partial n}$  helyett  $\text{grad}_h u \cdot n$ -t vesszük (a (15.175) képlet szerint),

$$\int_D k \frac{\partial u}{\partial n} ds \approx \int_D k \text{grad}_h u \cdot n ds,$$

akkor ez annak felel meg, hogy  $u$ -t a  $\Delta$  háromszögön lineárisnak tekintjük. Ennek alapján, míg a  $H$  résztartomány  $\Delta$ -beli részének pereme az  $s_{01}$  és  $s_{20}$  oldalak felezőpontjait összekötő folytonos, de egyébként tetszőleges vonalból áll, azt meg lehet mutatni (ld. a 15.36. tételt 15.7.9-ben), hogy az eredő lineáris egyenletrendszer mátrixa megegyezik azzal a mátrixszal, amely egy speciális végeselem módszerből keletkezik. A jobboldalak különbsége viszont  $h$ -val ill.  $h^2$ -tel becsülhető. Ebből is következik az eljárás stabilitása, valamint hibabecslése.  $\square$

### 15.6.5 Speciális kérdések

A boxmódszernek van olyan változata is (ez – mint az előbb említett 15.36. tétel 15.7.9-ben – Hackbusch-ra megy vissza), amely előnyösebb a többrácsos módszer szempontjából. Az egyes háromszögeket tehát mind négy hasonló háromszögre bontjuk fel, az oldalak felezőpontjait összekötve. A  $H = H(x)$  résztartomány mind az  $x$ -csúcsú kisebb háromszögeket tartalmazza, míg a nagy háromszögek belsejében elhelyezkedő kis háromszögeket  $H' = H'(x)$ -be



ábra 15.19: Hackbusch-felosztás : másodrendű boxmódszer

soroljuk. Annak peremét  $S'$ -vel jelöljük. Vegyük észre, hogy  $S'$  tartalmazza a  $H$  tartomány  $S$  peremét is!

Legyen a mérlegegyenlet ekkor ((15.173) helyett)

$$0 = \int_S (k \operatorname{grad} u) \cdot n \, ds + \frac{1}{3} \int_{S'} (k \operatorname{grad} u) \cdot n \, ds + \int_H f \, dx + \frac{1}{3} \int_{H'} f \, dx.$$

Ehhez megjegyezzük, hogy az  $\Omega$  feletti összegzésnél minden kis belső háromszög háromszor szerepel, míg a  $H(x)$  elemi résztartományok csak egyszer.

Ha  $k$ -t helyettesítjük háromszögenkénti  $\bar{k}$  átlagával és  $\operatorname{grad} u$ -t (15.175) szerinti approximációjával, akkor az  $S'$  feletti integrál (15.174) miatt nulla lesz, így a lineáris egyenletrendszer mátrixát csak az  $S$  feletti integrárból kell kiszámítanunk. Most alkalmazható az idézett 15.36. tétel, és a kapott megoldás távolsága a speciális végeelem megoldástól másodrendű.

Az előny itt az, hogy a többrácsos módszernél legelterjedtebb rácshinómítást használjuk, és hogy a négy kisebb háromszög területe és alakja megegyezik.

Mi fent csak a (15.163) egyenlet esetén szemléltettük a boxmódszert. Amennyiben a differenciálegyenlet elsőrendű deriváltakat is tartalmaz (mint *áramlási feladatok* esetén, ld. \*\*\* 17.4.), akkor hasonlóan folyik a diszkrét egyenletek levezetése. De mivel egy a problémára jellemző irány jelent meg, az áramlási irány, ilyenkor már nem közömbös, hogy a résztartományok peremei (amelyekre merőlegesen approximáljuk a fluxusokat) hogyan fekszenek az áramlási irányhoz képest! Ugyanis kedvezőtlen felosztás esetén előfordulhat (különösen háromszöges felosztásnál), hogy csupán olyan irányú fluxusok szerepelnek, amelyek a valódi áramlás szempontjából lényegtelenek. Ekkor jobb téglalapokkal dolgozni, legalább a peremek mentén.

## 15.7 A végeelem módszer II

### 15.7.1 Bevezetés

A Ritz–Galjorkin módszer fejlődése akkor torpant meg, amikor többdimenziós, általános tartományban megfogalmazott peremérték feladatokra már nem sikerült olyan függvényrendszert megadni, amellyel egyrészt a konvergencia garantált, de másrészt a diszkrétizált feladat jól kezelhető lenne. (A javaslatok olyanok voltak : vegyük a lényeges részére leegyszerűsített differenciáloperátor sajátfüggvényeit, ill. ha  $\Omega \subset \mathbb{R}^2$  és  $\Gamma$  egy  $\omega(x)$  függvénnyel jellemezhető :  $\omega(x) = 0 \iff x = (x_1, x_2) \in \Gamma$ , akkor vegyük az  $\{x_1^i x_2^j \omega(x_1, x_2)\}_{i,j}$  rendszert.)

Ekkor a szakaszonként ill. résztartományonként definiált polinomokkal való közelítés jelentette a kiutat, ld. II. 11.6-ban az egydimenziós eset tárgyalását. Az ott bevezetett és szemléltetett végeelem módszer fő alkalmazási területe a többdimenziós, általános tartományon megfogalmazott másodrendű és negyedrendű peremérték feladatok numerikus megoldása. Nézzük meg a módszernek az ilyen feladatokra való alkalmazásának lényeges lépéseit (ld. II. 249. o.) az elméleti oldalról (a gyakorlati oldalát később 15.7.4-ben és 15.7.7-15.7.11-ben részletesen vizsgáljuk majd). Itt kiderül, hogy a peremérték feladat dimenziójától függetlenül hivatkozhatunk a II. 11.20. tételre, valamint a 11.21. lemmára - mint lényeges elméleti alapjainkra.

1) Legyen  $\Omega \subset \mathbb{R}^d$ . Parciális integrálást alkalmazva a peremérték feladatról áttérünk a következő variációs egyenletre az alkalmasan kiválasztott  $V$  Hilbert-térben.

Keressünk olyan  $u \in V$  elemet, hogy

$$a(u, v) = \varphi(v), \quad \text{minden } v \in V\text{-re.} \quad (15.189)$$

Itt pl.  $V = H^1(\Omega)$ , ha a feladat másodrendű, és ekkor a norma

$$\begin{aligned} \|u\|_V = \|u\|_1 &:= \left( |u|_1^2 + \|u\|_0^2 \right)^{1/2}, \\ |u|_1^2 &:= \int_{\Omega} \sum_{i=1}^d \left( \frac{\partial u}{\partial x_i} \right)^2 dx = \int_{\Omega} |\text{grad } u|^2 dx, \\ \|u\|_0^2 &:= \int_{\Omega} u^2 dx, \end{aligned}$$

ahol  $|\text{grad } u|$  jelöli a  $\text{grad } u$  vektor euklideszi normáját.

Negyedrendű feladatnál  $\|u\|_1$  helyett  $\|u\|_2$  szerepel, ahol

$$\|u\|_2^2 := \int_{\Omega} \sum_{i,j=1}^d \left| \frac{\partial^2 u}{\partial x_i \partial x_j} \right|^2 dx + \|u\|_1^2.$$

2) A (15.189) feladatnak van egyértelmű, stabil megoldása (II. 11.20. tétel), ha az  $a(\cdot, \cdot)$  szimmetrikus bilineáris forma, és az  $V$ -elliptikus, azaz érvényes

$$a(u, u) \geq m_a \|u\|_V \quad \text{minden } u \in V\text{-re, } m_a := \text{const} > 0,$$

és ha  $\varphi(\cdot)$  lineáris funkcionál. Az ellipticitás bizonyítása itt a döntő lépés.

3) A variációs feladat közelítő megoldásához a  $V$  végtelen dimenziójú térről áttérünk a

$$V_h := \text{span}\{w_j\}_{j=1}^N \subset V$$

végesdimenziójú térre. Ez utóbbiban is van egyértelmű, stabil megoldása a variációs feladatnak :

keressünk olyan  $u_h \in V_h$ -t, azaz egy

$$u_h = \sum_{j=1}^N y_j w_j$$

alakú függvényt, hogy

$$a(u_h, v_h) = \varphi(v_h) \quad \text{minden } v_h \in V_h\text{-re.} \quad (15.190)$$

A Céa-lemmára támaszkodva mondhatjuk (II. 11.21. lemma), hogy az  $u_h$  távolsága a pontos megoldástól – egy konstans szorzót nem számítva – ugyanakkora, mint a legjobb közelítése.

4) A  $V_h$ -beli variációs feladat ekvivalens egy lineáris egyenletrendszerrel :

$$a(u_h, w_i) = \sum_{j=1}^N y_j a(w_j, w_i) =: (Ay)_i = b_i := \varphi(w_i), \quad (15.191)$$

ahol

$$A = (a_{k\ell}), \quad a_{k\ell} = a(w_\ell, w_k), \quad b_k = \varphi(w_k),$$

és annak  $A$  mátrixa pozitív definit és szimmetrikus – ha a differenciálegyenlet operátora önadjungált (azaz a bilineáris  $a$  forma szimmetrikus) és elliptikus.

A fenti lépéseket most megvizsgáljuk egy általánosabb  $d$ -dimenziós, másodrendű differenciálegyenlet példáján (elsősorban a  $d = 2$  és  $d = 3$  esetekre gondolunk) :

Legyen  $\Gamma$  az  $\Omega \subset \mathbb{R}^d$  korlátos tartomány Lipschitz-folytonos pereme, a differenciálegyenlet legyen

$$\sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left( k_{ij}(x) \frac{\partial u}{\partial x_j} \right) - \sum_{i=1}^d v_i(x) \frac{\partial u}{\partial x_i} - c(x)u + f(x) = 0, \quad (15.192)$$

ahol  $c, v_i, k_{ij} \in C(\bar{\Omega})$  és  $f \in L_2(\Omega)$ . A  $K$  mátrix legyen szimmetrikus és pozitív definit : az euklideszi skalárszorzatban teljesüljön

$$k_0|\xi|^2 \leq (K(x)\xi, \xi) \leq K_0|\xi|^2, \text{ minden } \xi \in \mathbb{R}^d\text{-re;} \quad (15.193)$$

$$K := (k_{ij}), \quad 0 < \text{const}_0 = k_0 \leq K_0 = \text{const}_1.$$

A (15.192) egyenletre (ill. az ebből és a megfelelő peremfeltételekből levezethető variációs feladatra) nézve ez a tulajdonság szoros kapcsolatban áll az ellipticitási feltétellel, ahogyan alább kiderül.

Elsőnek azt vizsgáljuk meg, hogy a fenti 1-2. lépések által kijelölt úton milyen akadályok lehetnek? (A 3-4. lépésekkel a 15.7.4. ponttól kezdve foglalkozunk.)

a) A parciális integrálás elégséges előfeltétele az, hogy az  $\Omega$  tartomány pereme Lipschitz-folytonos legyen (v.ö. 15.2-vel).

b) Az  $a$  forma és a lineáris funkcionál folytonosságával gond lehet nem korlátos tartomány, ill. együtthatók esetén. Ezen megfelelő súlyfüggvény bevezetése segíthet.

c) Az  $a$  forma azon tulajdonságának bizonyítása, hogy folytonos, vagyis az

$$|a(u, w)| \leq M_a \|u\|_V \|w\|_V \quad \text{minden } u, w \in V\text{-re}$$

becslés belátása gyakran megoldható a Schwarz- és a Cauchy-egyenlőtlenségek, valamint a beágyazási tételek (ld. 11.4.6., valamint \*\*\* 19. fejezet) segítségével. Hasonlóan belátható

$$|\varphi(w)| \leq M_\varphi \|w\|_V \quad \text{minden } w \in V\text{-re.}$$

d) A forma pozitív definittségét, azaz az

$$a(u, u) \geq m_a \|u\|_1^2$$

becslés érvényességének belátását megfelelő Poincaré-féle egyenlőtlenségek (ld. II. 11.7.2.) és beágyazási tételek segítik.

e) A forma szimmetriája sem feltétlenül áll fenn, mert a differenciálegyenlet, az adott peremfeltételek mellett, esetleg nem önadjungált – mint ez a (15.192) egyenlet esetén általában így van (pl. az elsőrendű deriváltak jelenléte miatt).

Az elmondottakat most a (15.192) egyenlet példáján szemléltetjük. Először (15.192)-re a parciális integrálás

$$\int_{\Omega} \frac{\partial w}{\partial x_i} z \, dx = \int_{\Gamma} w z n_i \, ds - \int_{\Omega} w \frac{\partial z}{\partial x_i} \, dx \quad (15.194)$$

alakú képletét alkalmazzuk, amelyben  $n_i = \cos(\vec{n}, x_i)$  a  $\Gamma$  perem  $\vec{n}$  külső normálvektorának  $i$ -edik komponense. Ha (15.194)-ben  $z$  helyett egy adott  $\vec{y} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  vektorfüggvénynek  $i$ -edik komponensét értjük és  $i$  felett összegzünk, akkor ebből az azonosságból a következőt kapjuk :

$$\int_{\Omega} \text{grad } w \cdot \vec{y} \, dx = \int_{\Gamma} w \vec{y} \cdot \vec{n} \, ds - \int_{\Omega} w \text{div } \vec{y} \, dx. \quad (15.195)$$

Itt  $\vec{y} \cdot \vec{n}$  az euklideszi skalárszorzatot jelöli. Legyen most először  $\vec{y} = K \text{grad } u$ . Ekkor

$$\int_{\Omega} \text{grad } w \cdot (K \text{grad } u) \, dx = \int_{\Gamma} w (K \text{grad } u) \cdot \vec{n} \, ds - \int_{\Omega} w \text{div}(K \text{grad } u) \, dx. \quad (15.196)$$

$K = I$  esetén ez az ismert 1. Green-féle képlet. Ezután legyen (15.195)-ben  $w = u$  és  $\vec{y} = w\vec{v}$ , akkor következik

$$\int_{\Omega} w \text{grad } u \cdot \vec{v} \, dx = \int_{\Gamma} u w \vec{v} \cdot \vec{n} \, ds - \int_{\Omega} u \text{div}(w\vec{v}) \, dx.$$

(15.196)-tal együtt azt kapjuk, hogy a  $w$ -vel megszorozott és  $\Omega$  felett integrált (15.192) egyenlet a következő formába alakítható át :

$$\begin{aligned} a(u, w) := & \int_{\Omega} \left[ (K \text{grad } u) \cdot \text{grad } w + \frac{1}{2} w \vec{v} \cdot \text{grad } u - \frac{1}{2} u \text{div}(w\vec{v}) + cuw \right] dx \\ & + \int_{\Gamma} w \left[ -K \text{grad } u + \frac{1}{2} u \vec{v} \right] \cdot \vec{n} \, ds = \varphi(w) := \int_{\Omega} f(x) w(x) \, dx. \end{aligned} \quad (15.197)$$

A

$$\text{div } w\vec{v} = w \text{div } \vec{v} + \vec{v} \cdot \text{grad } w$$

azonosságból következik

$$a(u, w) = \int_{\Omega} \left[ (K \operatorname{grad} u) \cdot \operatorname{grad} w + c u w - \frac{1}{2} u w \operatorname{div} \vec{v} + \frac{1}{2} \vec{v} \cdot (w \operatorname{grad} u - u \operatorname{grad} w) \right] dx - \int_{\Gamma} w \left[ K \operatorname{grad} u - \frac{1}{2} u \vec{v} \right] \cdot \vec{n} ds.$$

Ahogy látjuk, az  $a$  forma  $\vec{v} \neq 0$  esetén nem szimmetrikus: a  $w \operatorname{grad} u - u \operatorname{grad} w$  előjele megfordul az  $u \leftrightarrow w$  cserénél. (Éppen az volt a fenti átalakítások célja: hogy az  $a(u, w)$  kifejezést szimmetrikus és antiszimmetrikus részre bontsuk fel. Mátrixok esetén jól ismert az  $A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$  képlet.) A peremtag kiesik, ha a peremfeltétel elsőfajú:

$$w = u = 0, \quad x \in \Gamma = \Gamma_1. \quad (15.198)$$

Ekkor a  $V$  tér  $H_0^1(\Omega)$ , tehát a végtelen sokszor deriválható, a  $\Gamma$  perem közelében eltűnő függvények lezárása  $H^1(\Omega)$ -ban; normája a  $|\cdot|_1$  norma, és az ekvivalens a  $\|\cdot\|_1$ -normával.

A (15.198) peremfeltétel mellett (másod- és harmadfajú feltétellel lent 15.7.2-ben foglalkozunk)

$$a(u, w) = \int_{\Omega} \left[ (K \operatorname{grad} u) \cdot \operatorname{grad} w + c u w - \frac{1}{2} u w \operatorname{div} \vec{v} + \frac{1}{2} \vec{v} \cdot (w \operatorname{grad} u - u \operatorname{grad} w) \right] dx. \quad (15.199)$$

Vizsgáljuk most az  $u = w$  esetet! Ekkor az  $a$  pozitív definitése következik:

$$a(u, u) = \int_{\Omega} \left[ (K \operatorname{grad} u) \cdot \operatorname{grad} u + u^2 \left( c - \frac{1}{2} \operatorname{div} \vec{v} \right) \right] dx \geq m_a |u|_1^2, \quad m_a = k_0, \quad (15.200)$$

amennyiben teljesül (15.193) és

$$c - \frac{1}{2} \operatorname{div} \vec{v} \geq 0. \quad (15.201)$$

A (15.201) feltétel egyoldali korlátot szab  $\operatorname{div} \vec{v}$ -re. A feltétel pl. akkor teljesül, ha  $\operatorname{div} \vec{v} = 0$  (összenyomhatatlan közeg áramlása, ld. 17. fejezet) és  $c \geq 0$  (nyelő).

A  $c - \frac{1}{2} \operatorname{div} \vec{v} \geq 0$  feltétel úgy teljesíthető, hogy  $u$ -t az  $\exp(a_1 x_1 + a_2 x_2 + a_3 x_3) w(x)$  alakban keressük megfelelő  $(a_1, a_2, a_3)$  konstans vektorral, de ilyen transzformáció gyakorlatilag nem javasolható az exponenciális függvény gyors növekedése miatt.

A (15.200) becsléshez ld. a 38.a-c feladatokat is.

### 15.7.2 A peremfeltételekről

Az inhomogén elsőfajú peremfeltételek kezeléséről már II. 11.6.8-ban szöveltünk. Az ott elmondottak a többdimenziós esetnek megfelelő változtatással ugyanúgy alkalmazhatók. A lényeg az, hogy a feladatot (megfelelő függvény kivonásával) olyanra vezetjük vissza, ahol a peremen homogén elsőfajú peremfeltétel adott, majd a diszkretizáció során a homogén peremfeltételt beépítjük a végeselem térbe (azáltal, hogy csak ezeket a homogén feltételeket teljesítő bázisfüggvényekkel dolgozunk). Erre nincs szükség a másodfajú és harmadfajú peremfeltételeknél.

Az általános szabály kimondásához meg kell különböztetni a feladat *lényeges* és *természetes peremfeltételeit*.

A  $2m$ -edrendű elliptikus differenciálegyenlet peremérték feladatának *lényeges* az a peremfeltétele, amely legfeljebb  $m - 1$ -edrendű deriváltakat tartalmaz. *Természetes* az a peremfeltétel, amely a variációs feladat következménye.

Ami a lényeges peremfeltételeket illeti, arra mutatunk rá, hogy az önadjungált feladat esetén a variációs fogalmazás a keresett függvénynek a  $k$ -adrendű deriváltjait  $0 \leq k \leq m - 1$ -re a peremtagjaiban nem is tartalmazza (ld. pl. (15.189), (15.197), ha  $\vec{v} \equiv 0$ ): a lényeges peremfeltételeket már felhasználtuk a variációs fogalmazás levezetéséhez és a  $V$  tér definíciójának részévé tettük.

Másodrendű differenciálegyenlet esetén a másodfajú és harmadfajú peremfeltételek természetesek, ahogyan lejjebb megmutatjuk. Az elsőfajú peremfeltétel ekkor lényeges.

Negyedrendű feladat esetén viszont a keresett függvény és első deriváltjának kombinációját tartalmazó peremfeltétel lényeges, és azzal kezelendő, hogy (alkalmas függvénynek a keresett megoldásból való kivonása után) minden bázisfüggvénytől azt követeljük: teljesítse a megfelelő homogén feltételt a peremnek azon szakaszán, ahol a vizsgált peremfeltétel adott.

A végeselem diszkretizációnál a lényeges peremfeltételeket be kell építeni a bázisfüggvényekbe, a természetes peremfeltételeket nem (az utóbbiaknak megfelelő tagok a variációs egyenletben jelennek majd meg).  $2m$ -edrendű feladat esetén, ha csak természetes peremfeltételek vannak, nem megfelelő, ha olyan bázisfüggvényekkel akarunk dolgozni, amelyeknek értéke, az első  $m - 1$  deriválttal együtt, eltűnik a peremen. Ekkor ugyanis nem a kívánt, hanem a homogén elsőfajú peremérték feladatot oldjuk meg valójában és a végeselem tér túl szűk az adott peremérték feladat approximációjához.

Fentieket szemléltetve az alábbiakban a másodfajú és harmadfajú perem-



érték feladatot vizsgáljuk a (15.192) egyenlet azon esetében, amikor

$$\vec{v}(x) = 0, \quad c(x) \geq c_0 > 0, \quad \text{minden } x \in \overline{\Omega}\text{-re,} \quad (15.202)$$

(a  $\vec{v} \neq 0$  eset a 38.d feladat tárgya) és a peremen adott a

$$((K \operatorname{grad} u) \cdot \vec{n} + \sigma u)(x) = g(x), \quad \sigma(x) \geq 0, \quad x \in \Gamma \quad (15.203)$$

peremfeltétel, amely  $\sigma \equiv 0$  esetén másodfajú (a  $c(x) \geq c_0 > 0$  feltételnek az értelme az, hogy a másodfajú feltétel esetén ne jussunk szinguláris feladatra) és  $\sigma > 0$  esetén harmadfajú.

Készítsük el ezen feladat variációs megfogalmazását! Ekkor (15.197) helyett másképpen osztjuk fel a parciális integrálás adta tagokat  $a(u, w)$ -re és  $\varphi(w)$ -re :

$$\begin{aligned} \int_{\Omega} f(x)w(x) dx &= \int_{\Omega} [(K \operatorname{grad} u) \cdot \operatorname{grad} w + cuw] dx - \int_{\Gamma} w(K \operatorname{grad} u) \cdot \vec{n} ds \\ &= \int_{\Omega} [(K \operatorname{grad} u) \cdot \operatorname{grad} w + cuw] dx + \int_{\Gamma} w(\sigma u - g) ds, \end{aligned} \quad (15.204)$$

a (15.203) feltételt is figyelembe véve. Ezután a bilineáris formát, valamint a lineáris funkcionált a következőképpen definiáljuk :

$$a(u, w) := \int_{\Omega} [(K \operatorname{grad} u) \cdot \operatorname{grad} w + cuw] dx + \int_{\Gamma} w\sigma u ds, \quad (15.205)$$

$$\varphi(w) := \int_{\Omega} f(x)w(x) dx + \int_{\Gamma} w(s)g(s) ds. \quad (15.206)$$

Ebből azt látjuk, hogy az információ a (15.203) peremfeltételről teljesen elveszne, ha a fenti megfontolások ellenére mégis  $V_h \subset H_0^1$  végeelem térrel megpróbálnánk a (15.189), (15.205)-(15.206) variációs feladat közelítő megoldását. Nyilván a  $V_h$  végeelem tér kell, hogy a peremen nem eltűnő bázisfüggvényeket is tartalmazzon.

Az új  $a(u, w)$  forma pozitív definittségét kevés módosítással úgy láthatjuk be, mint korábban. Az  $a(u, u)$  peremintegrálja a (15.203) feltételnek köszönhetően alulról nullával becsülhető, és (15.193) és (15.202) mellett a (15.200) becslés helyett az

$$\begin{aligned} a(u, u) &= \int_{\Omega} [(K \operatorname{grad} u) \cdot \operatorname{grad} u + cu^2] dx + \int_{\Gamma} \sigma u^2 ds \\ &\geq m_a \|u\|_1^2, \quad m_a := \min(k_0, c_0), \end{aligned}$$

egyenlőtlenség is igaz.

A peremmenti integrálok értelmezésével nincsen problémánk,  $u$  és  $w$  le-szűkítése  $\Gamma$ -ra jól definiált, mivel  $u$  és  $w$  (a vizsgált kétdimenziós esetben, de a háromdimenziós esetben is) mint a  $H^1(\Omega)$  függvényter elemei a beá-gyazási tételek szerint egyben  $L_2(\Gamma)$ -nak is elemei (ld. a táblázatot a \*\*\* 19. fejezetben) :

$$\|u\|_{L_2(\Gamma)} \leq C_1 \|u\|_{H^1(\Omega)}. \quad (15.207)$$

Ez a becslés érvényes mint speciális beágyazási tétel (és a  $C_1$  konstans csak  $\Omega$ -tól és  $\Gamma$ -tól függ), ha az  $\Omega$  tartomány pereme pl. Lipschitz-folytonos (ehhez ld. a 60. feladatot is). A (15.207) becslés segítségével megmutathatjuk, hogy a ((15.199)-hez képest) peremintegrálokkal kibővített  $a(u, w)$  forma bilineáris és a  $\varphi(w)$  funkcionál lineáris. Ehhez már csak ezeket a peremintegrálokat kell felülről becsülnünk. Ha pl.  $\sigma$  szakaszonként folytonos és korlátos, akkor a Schwarz–egyenlőtlenség alapján

$$\left| \int_{\Gamma} \sigma u w \, ds \right| \leq \max_{x \in \Gamma} \sigma(x) \int_{\Gamma} |u w| \, ds \leq \max_{x \in \Gamma} \sigma(x) \|u\|_{L_2(\Gamma)} \|w\|_{L_2(\Gamma)}.$$

Ezt a becslést (15.207)-cal folytatva következik a kívánt eredmény :

$$\left| \int_{\Gamma} \sigma u w \, ds \right| \leq C_2 \|u\|_{H^1(\Omega)} \|w\|_{H^1(\Omega)}, \quad C_2 := C_1^2 \max_{x \in \Gamma} \sigma(x).$$

Hasonlóan kapjuk azt, hogy alkalmas  $C_3$  konstanssal igaz

$$\left| \int_{\Gamma} w g \, ds \right| \leq C_3 \|w\|_{H^1(\Omega)} \|g\|_{L_2(\Omega)}.$$

Ezek után a II. 11.20. tétel biztosítja a (15.189), (15.205)-(15.206) variá-ció-s feladat megoldásának létezését és unicitását. Mutassuk meg befejezésül, a II-beli 11.19. lemmát általánosítva, hogy ez a megoldás, ha elég sima, nemcsak a (15.192) egyenletet, hanem a (15.203) peremfeltételt is teljesíti, másszóval, ez természetes peremfeltétel.

**15.27. Lemma** (a másodfajú és harmadfajú peremfeltétel természetes). Legyen (15.192)-ben  $c, f, v_i, k_{ij} \in C(\bar{\Omega})$  és  $g, \sigma \in C(\Gamma)$ , teljesüljön (15.193) és (15.202). A (15.189), (15.205)-(15.206) variációs feladatnak az  $u \in H^1(\Omega)$  megoldása legyen  $C^2(\bar{\Omega})$ -ból való.

Ekkor (15.192) mellett (15.203) is teljesül.

**Bizonyítás.** Ha a (15.189), (15.205)-(15.206) feladatnak az  $u$  megoldása klasszikus, akkor a parciális integrálást megfordítva, eljutunk (15.204)-től a következő azonossághoz :

$$\int_{\Omega} w [f + \operatorname{div}(K \operatorname{grad} u) - cu] \, dx = \int_{\Gamma} w [(K \operatorname{grad} u) \cdot \vec{n} + \sigma u - g] \, ds. \quad (15.208)$$

Ebből  $w \in H_0^1(\Omega) \subset H^1(\Omega)$  esetén következik, hogy  $u$  kielégíti a (15.192) egyenletet is (itt  $w$  tetszőleges, és  $H_0^1$  tartalmazza azokat a végtelenül sokszor differenciálható függvényeket is, amelyek a perem közelében eltűnnek). Ekkor viszont (15.208)-ból az maradt vissza, hogy

$$0 = \int_{\Gamma} w[(K \operatorname{grad} u) \cdot \vec{n} + \sigma u - g] ds, \quad w \in H^1(\Omega),$$

vagyis,  $H^1(\Omega)$ -t az  $\bar{\Omega}$ -ban végtelenül sokszor differenciálható függvényekre szűkítve, a (15.203) érvényessége következik.  $\square$

**Megjegyzés.** Amennyiben pl. a perem  $\Gamma_1 \neq \emptyset$  részén elsőfajú,  $\Gamma_3 \neq \emptyset$  részén harmadfajú peremfeltétel adott, akkor az ellipticitás egy speciális Poincaré-egyenlőtlenségből következik, ld. Aubin könyvét. Ilyenkor olyan bázisfüggvényekkel kell dolgoznunk, amelyek  $\Gamma_1$ -en eltűnnek,  $\Gamma_3$ -on viszont nem. A  $\Gamma_3$ -on adott peremfeltétel ekkor is a variációs feladat következménye.  $\square$

### 15.7.3 A variációs feladat megoldhatósága

Annak ellenére, hogy a (15.199) szerint definiált  $a(u, w)$  forma nem szimmetrikus, a 11.20. tétel létezési eredménye érvényes. Ehhez viszont ahhoz a (15.189) alakú variációs feladathoz másképpen állunk hozzá, amelyben  $a$ , ill.  $\varphi$  a (15.199), ill. (15.197) által definiált és teljesül a (15.201) feltétel (és ezen variációs feladat eredeti, peremérték feladat alakja (15.192), (15.193), (15.198)).

Rögzített  $u$ -ra az  $a(u, w)$  egy lineáris funkcionál a  $w$  argumentummal, amelyet  $a(u, \cdot)$ -tal fogjuk jelölni. Ez azt jelenti, hogy  $u \in V$ -hez hozzá van rendelve az  $a(u, \cdot) \in V'$ -höz, ahol  $V'$  a  $V$ -hez tartozó lineáris funkcionáloknak a tere. A leképezést  $V$ -ből  $V'$ -be  $A$ -val jelöljük. Ez a leképezés additív és homogén, hiszen pl.  $u + v$ -hez hozzátartozik  $a(u + v, \cdot) = a(u, \cdot) + a(v, \cdot)$ . A leképezés korlátos is, mert az  $a$  forma korlátos és így a leképezés eredménye,  $a(u, \cdot)$ , becslhető  $u$  segítségével:

$$|a(u, v)| \leq M_a \|u\|_V \|v\|_V \text{ minden } v \in V\text{-re,}$$

azaz  $M_a \|u\|_V$  az  $a(u, \cdot)$  funkcionál normájának felső korlátja. Ebből következik továbbá az, hogy  $\|A\| \leq M_a$ .

A (15.189) variációs feladatot ezután úgy értelmezhetjük, hogy megoldásához olyan  $u$ -t kell meghatároznunk, amellyel  $a(u, \cdot)$  éppen az adott jobboldali  $\varphi$  funkcionállal egyenlő. Mivel  $a(u, \cdot)$  és  $\varphi$  egyaránt  $V$ -n definiált lineáris funkcionálok, ennek a feladatnak lehetnek megoldásai.

A variációs feladatnak az ebben az értelemben vett megoldhatóságát, valamint az  $a$  bilineáris forma és az általa definiált  $A$  operátor kapcsolatát rendezzi a következő eredmény.

**15.28. Lemma** (Babuška; variációs feladat megoldhatósága). Legyenek  $V_1$  és  $V_2$  Hilbert-terek és  $a(v_1, v_2)$  bilineáris forma úgy, hogy pozitív  $C_1, C_2$  konstansokkal teljesül

1.  $|a(v_1, v_2)| \leq C_1 \|v_1\|_1 \|v_2\|_2$ , minden  $v_i \in V_i$ -re,  
ahol  $\|\cdot\|_i$  a  $V_i$  normája,  $i = 1, 2$ ;
2.  $\sup_{v_1 \in V_1, \|v_1\|_1=1} a(v_1, v_2) > 0$ , minden nemzérus  $v_2 \in V_2$ -re;
3.  $\sup_{v_2 \in V_2, \|v_2\|_2=1} a(v_1, v_2) \geq C_2 \|v_1\|_1$ , minden  $v_1 \in V_1$ -re.

Végülis, legyen  $f$  tetszőleges lineáris funkcionál  $V_2$ -n. Ekkor van pontosan egy  $v = v_f \in V_1$  úgy, hogy

$$a(v_f, v_2) = f(v_2) \quad \text{minden } v_2 \in V_2\text{-re,}$$

továbbá, teljesül, hogy

$$\|v_f\| \leq \frac{1}{C_2} \|f\|_{V_2'}. \quad \square$$

**15.29. Következmény** (lineáris funkcionál által definiált operátor invertálhatósága). A 15.28. lemma feltételei mellett az  $a$  formához rendelhető  $A : V_1 \rightarrow V_2'$  operátor invertálható. Az első feltétel alapján  $A$  korlátos (és így lineáris, mivel homogén és additív), a második feltétel miatt érvényes  $A(V_1) = V_2'$ , a harmadik feltétel viszont biztosítja az inverz operátor korlátosságát :  $\|A^{-1}\| \leq 1/C_2$ .  $\square$

**Megjegyzések. 1.** A lemma 2. és 3. feltétele aszimmetrikus  $a(v_1, v_2)$ -re nézve. De az eredmény, az inverz operátor létezése és korlátossága, egyben az  $A^*$  adjungált operátor,  $A^* : V_2 \rightarrow V_1'$ , inverzének létezését és korlátosságát is maga után vonja. Ez azt jelenti, hogy a következő feladatnak pontosan egy megoldása van:

Keressük  $v = v_g$ -t  $V_2$ -ben úgy, hogy

$$a(v_1, v) = g(v_1) \quad \text{minden } v_1 \in V_1\text{-re.}$$

Ezenkívül ez a megoldás stabil : érvényes

$$\|v_g\|_2 \leq \frac{1}{C_2} \|g\|_{V_1'}.$$

Valóban, meg lehet mutatni, hogy a 2. és 3. feltétel azzal ekvivalens, hogy érvényes 3. és

$$2'. \quad \sup_{v_1 \in V_1, \|v_1\|_1=1} a(v_1, v_2) \geq C_2 \|v_2\|_2 \quad \text{minden } v_2 \in V_2\text{-re.}$$

2. Amikor a Babuška–lemmában szereplő két Hilbert-tér megegyezik :  $V_1 = V_2 = V$ , és a bilineáris forma  $V$ -ben elliptikus :

$$a(v, v) \geq m_a \|v\|_V^2,$$

akkor a lemma 2. és 3. feltétele teljesül  $C_2 = m_a$ -val, és a hozzátartozó operátor linearitása és invertálhatósága következik. Ezt az eredményt ekkor (amikor  $a$  nem feltétlenül szimmetrikus) *Lax–Milgram tétel*nek hívjuk. Mivel Riesz tétele szerint a  $V$  Hilbert-téren definiált tetszőleges funkcionál mint  $(\psi, \cdot)_V$  írható fel, alkalmas  $\psi \in V$ -vel, azért az  $A$  operátort mint  $V \rightarrow V$  leképezést is értelmezhetjük. Eszerint

$$a(u, \cdot) = (Au, \cdot)_V \quad \text{és } \varphi(v) = (\psi, \cdot)_V.$$

Ekkor a fentiek szerint a  $V$ -beli  $Au = \psi$  egyenletnek van egyértelmű és stabil  $u$  megoldása.  $\square$

A 2. megjegyzéshez ld. a 39. feladatot is.

A végeelem módszer konvergencia vizsgálatánál (ld. a 15.32. következményt), valamint a végeelem alapú többrácsos módszer approximációs tulajdonságának bizonyításánál (ld. a 15.37. lemmát 15.8.2-ben) fontos az, hogy a variációs feladat megoldása nem  $H^1$ , hanem  $H^2$ -ből való, tehát regulárisabb. Ezért itt az ezt biztosító feltételeket idézzük (ld. a részletes tárgyalást Hackbusch „elliptikus” könyvében és v.ö. II. 11.6.3-ban a Nitsche–fogással is) a mi esetünkre specializált alakban.

**15.30. Tétel** (Kadlec; variációs megoldás  $H^2$ -ben). Legyen  $\Omega \in \mathbb{R}^n$  korlátos és konvex, teljesüljön (15.193),  $\Gamma = \Gamma_1$  (elsőfajú peremfeltételek) és (15.201). Ezenfelül legyenek a  $k_{ij}$  együtthatók Lipschitz-folytonosak. Ekkor a (15.189), (15.197) variációs feladat gyenge megoldása  $H^2 \cap H_0^1(\Omega)$ -ből való és van olyan (csak a tartomány átmérőjétől, valamint  $m_a$ -tól függő  $C$  konstans, hogy

$$\|u\|_{H^2(\Omega)} \leq C \|f\|_0. \quad \square \quad (15.209)$$

**Megjegyzések. 1.** A fenti eredményhez a (15.201) feltétel nem szükséges. Ekkor az unicitás ugyan nem biztosított, de (15.193)-ból következik

egy Gårding-féle egyenlőtlenség (ld. a 38.b feladatot) és, (15.209) helyett, a variációs feladat minden megoldására igaz

$$|u|_{H^2(\Omega)} \leq C[\|f\|_0 + |u|_1]. \quad (15.210)$$

Ha viszont (15.201) is érvényes, akkor (15.200) igaz és a variációs megoldás stabilitásából (ld. II-ben a 11.20 tételt és a (216) képletet, 240. o.) :

$$|u|_1 \leq \frac{c}{m_a} \|f\|_0$$

(megfelelő  $c$  konstanssal) és (15.210)-ből következik (15.209).

2. A fenti tétel hasznos a II. 11.6.3-ban az egydimenziós modellfeladat példáján bemutatott *Nitsche-fogás* általános megalapozásához is.

3. A tartomány konvexitása lényeges. Legyen ugyanis

$$v(r, \varphi) := (1 - r^2)r^{2/3} \sin((2\varphi - \pi)/3)$$

a 15.2-ben vizsgált  $u$  függvény szorzata  $1 - r^2$ -tel. Ekkor egyrészt  $v|_{\Gamma} = 0$  a 15.2. ábra szerinti  $\Omega$  tartomány peremén, másrészt az ottani (15.191) képlet segítségével közvetlenül kiszámítható, hogy

$$\Delta_{r,\varphi} v = -\frac{20}{3}u =: -f.$$

Ekkor  $f \in L_2(\Omega)$ , mert ez a függvény folytonos, de  $v \notin H_0^2(\Omega)$ , hiszen

$$\int_{\Omega} \left( \frac{\partial^2 v}{\partial r^2} \right)^2 r \, dr \, d\varphi$$

nem véges.  $\square$

#### 15.7.4 Véges elemek 2- és 3-dimenziós feladatokban

A Galjorkin-módszer következő lényeges lépése az, hogy a megfelelő  $V$  függvénytér lineáris, végesdimenziójú  $V_h$  alterét ki kell választani.

Ehhez a végeelem módszer azt javasolja, hogy a következőképpen járjunk el :

1. Az  $\bar{\Omega} \subset \mathbb{R}^d$  tartományt felosztjuk  $\bar{\Omega}_i$  Lipschitz-folytonos részterományokra,  $i = 1, \dots, M$ , kétdimenziós feladat esetén mindenekelőtt háromszögekre vagy konvex négyszögekre gondolunk) úgy, hogy

- a)  $\bar{\Omega} = \cup_{i=1}^M \bar{\Omega}_i$ ,
- b)  $\Omega_i \cap \Omega_j = \emptyset$ ,  $i \neq j$ ,

$$c) \overline{\Omega}_i \cap \overline{\Omega}_j = \left\{ \begin{array}{l} \overline{\Omega}_i \text{ és } \overline{\Omega}_j \text{ közös csúcspontja,} \\ \text{vagy } \overline{\Omega}_i \text{ és } \overline{\Omega}_j \text{ közös oldala,} \\ \text{vagy } \emptyset, \end{array} \right\} \text{ és nincs más lehetőség.}$$

Az  $\Omega$  tartomány ilyen  $\mathcal{T}_h := \{\Omega_i\}_{i=1}^M$  felosztását *konform triangulációjának* hívjuk (akkor is, ha az  $\Omega_i$ -k nem feltétlenül háromszögek). Ehhez ld. a 17. ábrát 15.6.3-ban.

Ilyen felosztás akkor létezik, ha  $\Gamma$  töröttvonal. A  $h$  betűvel a trianguláció finomságát jellemezzük – mint az egydimenziós esetben. Itt

$$h := \max_{1 \leq i \leq M} \text{diam}(\overline{\Omega}_i),$$

tehát a maximális  $\overline{\Omega}_i$ -beli távolság maximuma az egész trianguláción.

2. Minden  $\overline{\Omega}_i$ -n egy  $\{x^{i,1}, \dots, x^{i,n_i}\}$  ponthalmazt jelölünk ki, az  $x^{i,j}$ -k a *csomópontok*. A csomópontokhoz tartozik a résztartomány összes csúcspontja, de további perem- és belső pontjai is. Itt  $n_i \geq d + 1$  és főként arra gondolunk, hogy  $n_i = \text{const}$  függetlenül  $i$ -től.

A csomópontoknak az  $\{x^{i,j}\}_{j=1}^{n_i}$  a résztartományonkénti, lokális sorszámozása. Emellett a csomópontok globális sorszámozását is elkészítjük:  $\{x^\ell\}_{\ell=1}^N$ . Ez utóbbi felsorolásban minden csomópont pontosan egyszer fordul elő.

3. Minden  $x^\ell$  csomóponthoz hozzárendeljük azt a  $\psi^\ell$  lineáris funkcionált, amely megadja az  $x^\ell$ -beli függvényértéket:  $\psi^\ell(u) = u(x^\ell)$ . Ezeket a funkcionálok *szabadsági fokoknak* hívjuk. (Magát a függvényértéket is hívhatjuk szabadsági foknak, de deriváltak értékei is előfordulnak mint szabadsági fokok, ld. a pont végét. Így jobb az összefoglaló „funkcionál” elnevezés.) Amikor – mint eddig itt – minden szabadsági fok függvényérték, akkor *Lagrange-típusú elemről* beszélünk.

Az  $\overline{\Omega}_i$ -hez tartozó szabadsági fokok halmaza legyen  $\psi_i := \{\psi^\ell, x^\ell \in \overline{\Omega}_i\}$ , amelyet a lokális megszámozásban  $\psi_i = \{\psi^{i,j}\}_{j=1}^{n_i}$ -vel jelöljük.

4. Minden  $\psi^\ell$ -hez hozzárendelünk egy  $w_\ell \in V$  függvényt úgy, hogy

a)  $w_\ell$  leszűkítése  $\overline{\Omega}_i$ -re legyen polinom (=:  $p_{i,\ell}$ );

b) minden egyes  $\overline{\Omega}_i$ -n a  $\{p_{i,\ell}\}$  polinomok a  $P(\overline{\Omega}_i)$  lineáris halmazt alkotják, és ha

$$\{x^{i,1}, \dots, x^{i,n_i}\} = \{x^{\ell_1}, \dots, x^{\ell_{n_i}}\}$$

azok a csomópontok (először a lokális, majd a globális felsorolás szerint), amelyek  $\overline{\Omega}_i$ -hez tartoznak, akkor a

$$\sum_{j=1}^{n_i} p_{i,\ell_j}(x) \alpha_{ij} =: p(x), \quad x \in \overline{\Omega}_i,$$

képlettel definiált  $P(\overline{\Omega}_i)$ -beli polinom  $\alpha_{ij}$  együtthatói egyértelműen meg legyenek határozva a  $\{\psi^{\ell_1}(p), \dots, \psi^{\ell_{n_i}}(p)\}$  értékkel, pl. annak köszönhetően, hogy érvényes

$$\psi^{\ell_j}(p_{i,\ell_m}) = \delta_{jm}, \quad j, m = 1, \dots, n_i; \quad (15.211)$$

c) ha  $x^{\ell_1}, \dots, x^{\ell_{n_i(s)}}$  azok a csomópontok, amelyek  $\overline{\Omega}_i$ -nek egy tetszőleges, rögzített  $s$  oldalán vannak, akkor a

$$\sum_{j=1}^{n_i(s)} p_{i,\ell_j}(x) \alpha_{ij} =: q(x)$$

polinom egyértelműen meg legyen határozva az  $s$  oldalon a  $\{\psi^{\ell_j}(q)\}_{j=1}^{n_i(s)}$  értékek megadása után. Viszont a többi  $p_{i,\ell_j}$  polinom ( $j = n_i(s) + 1, \dots, n_i$ ) legyen azonosan nulla azon az oldalon, tehát ha  $\psi^{\ell_j}(q) = \psi^{\ell_j}(p)$ , amikor  $j = 1, \dots, n_i(s)$ , akkor érvényes  $p|_s = q$ .

Ezután, amennyiben  $\psi^m(w_\ell) = \delta_{m\ell}$ ,  $m, \ell = 1, \dots, N$ , akkor a  $V_h$  végeselem tér bázisa éppen  $\{w_\ell\}_{\ell=1}^N$ . Maga az ( $i$ -edik) „véges elem” a következő hármas :

$$\{\overline{\Omega}_i, \psi_i, P(\overline{\Omega}_i)\},$$

tehát a {résztartomány, hozzátartozó funkcionálok halmaza, polinomtér} hármasa.

A  $P(\overline{\Omega}_i)$  polinomtér alapján definiálhatjuk az elem rendjét.

**Definíció.** Legyen  $P_k(\overline{\Omega}_i)$  az  $\overline{\Omega}_i$ -n definiált, legfeljebb  $k$ -adfokú polinomok lineáris halmaza, azaz  $P_k(\overline{\Omega}_i)$  az összes  $x_1^{\gamma_1} \cdots x_d^{\gamma_d}$  alakú polinom (ahol  $0 \leq \sum_{\ell=1}^d \gamma_\ell \leq k$ ) minden lineáris kombinációját tartalmazza. Ekkor  $k$ -adrendűnek hívjuk az elemet, ha  $P_k(\overline{\Omega}_i) \subset P(\overline{\Omega}_i)$ , de  $P_{k+1}(\overline{\Omega}_i) \not\subset P(\overline{\Omega}_i)$ .

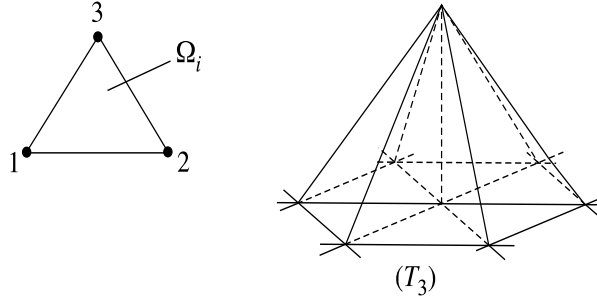
Továbbá, legyen  $Q_k(\overline{\Omega}_i)$  azon  $x_1^{\gamma_1} \cdots x_d^{\gamma_d}$  alakú tagok összes lineáris kombinációja, amelyeknél  $\gamma_1 \leq k, \dots, \gamma_d \leq k$  a feltétel.  $\square$

**Megjegyzések. 1.** A  $P_k(\overline{\Omega}_i)$  polinomterek tipikusak a háromszöges (általános dimenzióban : szimpliciális)  $\overline{\Omega}_i$ -vel rendelkező elemekre, míg a  $Q_k(\overline{\Omega}_i)$  polinomterek tipikusak a négyszöges elemekre.

**2.** A 15.7.5. pont szerint gyakran arra számíthatunk, hogy a  $k$ -adrendű elemek alkalmazása esetén az eredő végeselem megoldás konvergenciarendje  $h$ -ban éppen  $k$ .

**3.** A c) feltétel azt eredményezi, hogy a  $w_\ell$  báziselemek folytonosak két szomszédos elem közös oldalán keresztül is. Ha ott szakadás lenne, akkor  $w_\ell$  nem lenne  $H^1(\Omega)$  eleme (mert nincsen négyzetesen integrálható deriváltja) és így a  $V_h$  végeselem tér nem lenne  $H^1(\Omega)$  része. Hasonlóan a negyedrendű feladatok esetén, ahol a természetes tér a  $H^2(\Omega)$  (vagy a  $H_0^2(\Omega)$ ), arra van





ábra 15.20: A  $T_3$  elem

szükség, hogy két elem közös oldalán tetszőleges  $v \in V_h$  első deriváltja is folytonos legyen.

Az olyan végeelem módszerek, amelyek teljesítik a  $V_h \subset V$  feltételt, *konformnak* hívjuk. Ebben a könyvben – a \*\*\* 17.5. rész kivételével – csak konform módszerekkel foglalkozunk.  $\square$

Most az előbbi, eléggé absztrakt meghatározást több példával szemléltetjük.

1. *példa* :  $\bar{\Omega} = [0, 1]$  és  $\bar{\Omega}_i = [x_{i-1}, x_i]$  az  $i$ -edik intervallum a  $0 = x_0 < x_1 < \dots < x_M = 1$  felosztásból, az  $\bar{\Omega}_i$  szabadsági fokainak halmaza  $\psi_i = (\psi^{i,1}, \psi^{i,2})$ , ahol  $\psi^{i,1}(u) = u(x_{i-1})$  és  $\psi^{i,2}(u) = u(x_i)$  – tehát  $n_i = 2$ , és polinomtere  $P(\bar{\Omega}_i) = P_1(\bar{\Omega}_i)$ .

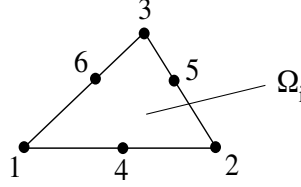
A bázist alkotják a  $w_\ell$  kalapfüggvények II. 11.6.2-ből, amelyekre teljesül  $\psi^m(w_\ell) = w_\ell(x_m) = \delta_{m\ell}$ , mégpedig elsőfajú peremérték feladat esetén a bázis  $\{w_\ell\}_{\ell=1}^{M-1}$  (így  $N = M - 1$ ), másod- és harmadfajú esetben  $\{w_\ell\}_{\ell=0}^M$  (tehát ekkor  $N = M + 1$ );

2. *példa* : Az  $\bar{\Omega} \in \mathbb{R}^2$  pereme legyen töröttvonal, a tartomány triangulációja  $\bar{\Omega}_i = \bar{\Delta}_i$  nem elfajult háromszögekre rendelkezésünkre áll. Az  $i$ -edik háromszög csomópontjai a háromszög  $x^{i,1}, x^{i,2}, x^{i,3}$  csúcsai.

$P(\bar{\Omega}_i) = P_1(\bar{\Delta}_i)$  a  $\bar{\Delta}_i$  felett definiált lineáris függvényeknek, tehát az  $a_i + b_i x_1 + c_i x_2$  alakú függvényeknek a tere. Ezeknek  $n_i = 3$  szabadsági fokuk van :  $\psi^{i,j}(u) := u(x^{i,j})$ ,  $1 \leq j \leq 3$ , amelyeknek az  $a_i, b_i, c_i$  szabad paraméterek felelnek meg. A  $\psi^{i,j}$  funkcionálok tehát a  $\bar{\Delta}_i$  háromszög csúcsaiban felvett értékeket adják meg :

$$\psi_i := \{ \psi^{i,j}; \psi^{i,j}(p) = p(x^{i,j}), p \in P_1(\bar{\Delta}_i), j = 1, 2, 3 \}.$$

Ez az úgynevezett *Courant–elem*, ld. I. 4.10.3. és lejjebb is 15.7.15.7.8-ban és 15.7.10-ben. Ezt a gyakran használt elemet röviden  $T_3$ -mal jelöljük, rendje 1;



ábra 15.21: A  $T_6$  elem

3. példa :  $\bar{\Omega}$  most is poligonális tartomány, amelyet  $\bar{\Delta}_i$  nem elfajult háromszögekre bontottuk fel. A csomópontok egyrészt a  $\bar{\Delta}_i$  háromszög  $x^{i,j}$  csúcsai,  $1 \leq j \leq 3$ , másrészt a háromszög oldalfelező pontjai :

$$x^{i,j} := \frac{1}{2} (x^{i,j-3} + x^{i,j-2}), \quad 4 \leq j \leq 6, \quad (\text{ahol } x^{i,4} := x^{i,1}).$$

$P_2(\bar{\Delta}_i)$  a  $\bar{\Delta}_i$  felett definiált másodfokú polinomok tere (tehát az  $a_i + b_i x_1 + c_i x_2 + d_i x_1^2 + e_i x_1 x_2 + f_i x_2^2$  alakú függvények. Ezeknek  $n_i = 6$  szabadsági fokuk van :  $u(x^{i,j}), j = 1, \dots, 6$ , a 6 szabad paraméternek megfelelően :  $a_i, \dots, f_i$ ). A megfeleltetés az  $u(x^{i,j})$  értékek és az  $a_i, \dots, f_i$  paraméterek között egyértelmű (ld. a 40. feladatot). A lineáris funkcionálok tehát :

$$\psi_i := \{ \psi^{i,j}; \psi^{i,j}(p) = p(x^{i,j}), p \in P_2(\bar{\Delta}_i), 1 \leq j \leq 6 \}.$$

Ezt a másodrendű háromszögű elemet  $T_6$ -tal jelöljük.

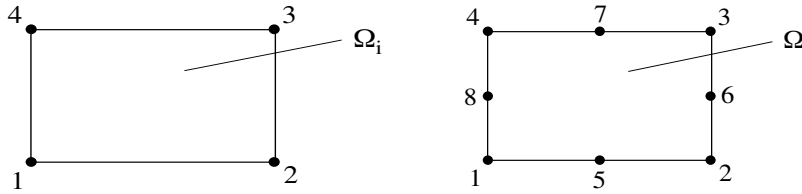
Mindkét kétdimenziós példánál a  $w_\ell$  bázisfüggvények leszűkítését  $\bar{\Delta}_i$ -ra (azaz a  $p_{i,\ell}$  polinomokat), az úgynevezett *interpolációs* vagy *formafüggvényeket* legegyszerűbben a háromszög  $\{\lambda_j(x)\}_{j=1}^3$  baricentrikus koordinátáinak a segítségével (ld. I. 4.10.3.) adhatjuk meg :

$$\begin{aligned} T_3 \text{ esetén } x^{i,j}\text{-hez tartozik } p_{i,j} &= \lambda_j, \quad 1 \leq j \leq 3; \\ T_6 \text{ esetén } x^{i,j}\text{-hez tartozik } p_{i,j} &= \lambda_j(2\lambda_j - 1), \quad 1 \leq j \leq 3, \quad (15.212) \\ x^{i,j}\text{-hez tartozik } p_{i,j} &= 4\lambda_{j-3}\lambda_{j-2}, \quad 4 \leq j \leq 6. \end{aligned}$$

Itt az interpolációs függvényeket a lokális sorszámozásban adtuk meg :  $p_{i,j}$ , ezekre teljesül (15.211) a  $\psi^{i,j}(p_{i,k}) = \delta_{jk}$ ,  $j, k = 1, \dots, n_i$  formában.

A baricentrikus koordináták függőségét a mindenkor  $\bar{\Delta}_i$  háromszögtől nem jelöltük, továbbá, a kényelmes  $\lambda_4 := \lambda_1$  jelölést használtuk.

A  $T_3, T_6$  elemek egy olyan sorozat elejét alkotják, amelynek  $k$ -adik tagja  $T_{n(k)}$ ,  $n(k) := \frac{1}{2}(k+1)(k+2)$  szabályosan elhelyezkedő csomóponttal rendelkezik, ebből a háromszög bármely oldalán  $k+1$  darab található, és amely elem  $k$ -adfokú polinomokat használ : a háromszöges standard Lagrange-elemek sorozata.



ábra 15.22:  $R_4$  elem és  $R_8$  elem

4. példa :  $\bar{\Omega}_i$  a  $\bar{\square}_i$  téglalap, csúcsai  $x^{i,j}$ ,  $1 \leq j \leq 4$ . A  $P(\bar{\Omega}_i)$  polinomtér a  $\bar{\square}_i$  felett definiált bilineáris függvények  $Q_1(\bar{\square}_i)$  tere. 4 szabadsági foka a  $\bar{\square}_i$  csúcsaiban felvett  $u$ -értékek :  $u(x^{i,j})$ ,  $j = 1, \dots, 4$ , annak megfelelően, hogy 4 szabad paramétere van az elem  $a_i + b_i x_1 + c_i x_2 + d_i x_1 x_2$  alakú polinomjainak, és láthatóan igaz  $P_1(\bar{\Omega}_i) \subset P(\bar{\Omega}_i)$ , de  $P_2(\bar{\Omega}_i) \not\subset P(\bar{\Omega}_i)$ . Ezt az elemet  $R_4$ -gyel jelöljük és *bilineáris elemnek* is hívjuk.

Polinomjainak négy szabad paraméterét az  $u(x^{i,j})$  szabadsági fokok egyértelműen határozzák meg. Az elem funkcionáljai tehát

$$\psi_i := \{ \psi^{i,j}; \psi^{i,j}(p) = p(x^{i,j}), p \in Q_1(\bar{\square}_i), 1 \leq j \leq 4 \}.$$

5. példa :  $\bar{\Omega}_i$  a  $\bar{\square}_i$  téglalap, csomópontjai az  $x^{i,j}$  csúcsai,  $1 \leq j \leq 4$ , további csomópontjai az  $x^{i,j}$ ,  $5 \leq j \leq 8$  oldalfelező pontok.

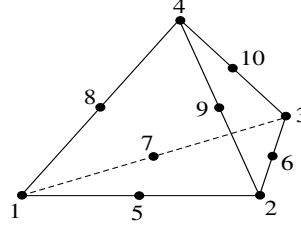
Itt  $P(\bar{\Omega}_i)$  a  $Q_2(\bar{\square}_i)$  azon polinomjaiból áll, amelyekben az  $x_1^2 x_2^2$  tag hiányzik. Ezzel  $P_2(\bar{\Omega}_i) \subset P(\bar{\Omega}_i)$ , az elemnek 8 szabadsági foka van :  $\{u(x^{i,j})\}_{j=1}^8$ , és 8 szabad paraméterrel rendelkeznek az elem polinomjai is :  $a_i + b_i x_1 + c_i x_2 + d_i x_1^2 + e_i x_1 x_2 + f_i x_2^2 + g_i x_1^2 x_2 + h_i x_1 x_2^2$ . Az elemet  $R_8$ -cal jelöljük,  $\psi^{i,j}$  funkcionáljai az  $x^{i,j}$  csomópontokbeli értékeket rendelik hozzá  $p \in P(\bar{\Omega}_i)$ -hez (amivel ez egyértelműen meghatározott, ld. a 41. feladatot).

Az alappontokhoz hozzátehetnénk mint kilencediket a téglalap középpontját, de mivel ez még nem eredményezi azt, hogy  $P_3(\bar{\Omega}_i) \subset P(\bar{\Omega}_i)$  – amivel biztosítaná a harmadrendű konvergenciát – (csak a 10. szabadsági fok tehetné, de az megbontja a szimmetriát), azért a 9. szabadsági fok (a középpont) elhagyható.

Ugyanezzel az érveléssel még további két szabadsági fok is elhagyhatónak tűnik anélkül, hogy a konvergencia sebesség csökkenne, de ez lehetetlen, ld. a 42. feladatot.

A középpont mint szabadsági fok figyelemen kívül hagyása két okból is előnyös :

1) kevesebb ismeretlennel dolgozunk, a konvergenciarend csökkentése nélkül (ez a szokásos – 15.8.2-ben tárgyalt beágyazott – rácssorozatokon igen, de általános rácssorozatban nem érvényes, ld. Arnold, Boffi, Falk dolgozatát);



ábra 15.23: A  $T_{10}^3$  tetraéderelem

2) a 9. pont behozná a  $x_1^2 x_2^2$  tagot is és ezzel negyedfokúvá válna a bázis.

Amikor az  $R_8$  formafüggvényeit szerkesztjük abból a követelményből, hogy teljesüljön  $p_{i,j}(x^{i,\ell}) = \delta_{j,\ell}$ ,  $1 \leq j, \ell \leq 8$ , akkor kiderül, hogy éppen az  $x_1^2 x_2$  és  $x_1 x_2^2$  harmadrendű tagok jönnek be. Ehhez egy tetszőleges csomópontot kijelölünk és a többi csomópontot 3 egyenessel lefedjük. Ezen egyeneseknek egyenleteit összeszorozzuk, majd a szorzatot alkalmas konstanssal megszorozzuk úgy, hogy a kiválasztott pontban a függvény értéke 1 legyen (és konstrukció szerint a formafüggvény a többi csomópontban eltűnik) :

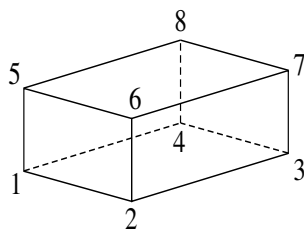
$$\begin{aligned}
 w_1(x) &:= (1-x_1)(1-x_2)(1-2x_1-2x_2), & w_2(x) &:= x_1(1-x_2)(1-2x_1+2x_2), \\
 w_3(x) &:= x_1 x_2 (3-2x_1-2x_2), & w_4(x) &:= (1-x_1)x_2(1+2x_1-2x_2), \\
 w_5(x) &:= 4x_1(1-x_1)(1-x_2), & w_6(x) &:= 4x_1 x_2 (1-x_2), \\
 w_7(x) &:= 4x_1(1-x_1)x_2, & w_8(x) &:= 4(1-x_1)x_2(1-x_2).
 \end{aligned} \tag{15.213}$$

Most néhány háromdimenziós elemet mutatunk be.

*6. példa :* A Courant–elem általánosítása három dimenzióra a lineáris tetraéder-elem ( $T_4^3$ ). Itt tehát  $\bar{\Omega}_i$  az  $x^{i,j}$  (nem egy síkban fekvő) csúcsok által meghatározott tetraéder. A négy szabadsági fok az  $x^{i,j}$  pontban felvett függvényérték,  $j = 1, 2, 3, 4$ . Az elem polinomtere az összes  $a_i + b_i x_1 + c_i x_2 + d_i x_3$  alakú lineáris függvény, vagyis  $P(\bar{\Omega}_i) = P_1(\bar{\Omega}_i)$ , és így rendje 1.

Egyébként egyáltalán nem egyszerű egy adott  $\Omega \in \mathbb{R}^3$  tartományt tetraéderekre szétbontani, ehhez ld. 15.7.11.

*7. példa :* A másodrendű tetraéder-elem ( $T_{10}^3$ ) csomópontjaihoz a csúcsokon kívül ( $x^{i,j}$ ,  $1 \leq j \leq 4$ ) a oldalfelező pontok ( $x^{i,j}$ ,  $5 \leq j \leq 10$ ) is tartoznak hozzá. A 10 csomóponttal éppen az  $x_1, x_2, x_3$ -ban legfeljebb másodfokú polinomok 10 szabad paraméterét ki is merítjük, és a felsorolt csomópontokban megadott függvényértékekből egyértelműen meghatározható egy ilyen polinom : tehát  $P(\bar{\Omega}_i) = P_2(\bar{\Omega}_i)$ . Az elem formafüggvényeit itt is a baricen-



ábra 15.24: Az  $R_8^3$  kockaelem

trikus koordináták segítségével adjuk meg :

$$\begin{array}{ll}
 x^{i,j}\text{-hez tartozik} & \lambda_j(2\lambda_j - 1), \quad 1 \leq j \leq 4, \\
 x^{i,5}\text{-höz, } x^{i,6}\text{-hoz, ill. } x^{i,7}\text{-hez} & \text{tartozik} \quad 4\lambda_1\lambda_2, 4\lambda_2\lambda_3, \text{ ill. } 4\lambda_3\lambda_1, \\
 x^{i,j}\text{-hez tartozik} & 4\lambda_{j-7}\lambda_4, \quad j = 8, 9, 10.
 \end{array}$$

8. *példa* : A trilineáris kocka-elem ( $R_8^3$ ) könnyen kezelhető és az  $R_4$ -elem általánosítása 3 dimenzióra. A kocka 8 csúcsában megadott függvényértékek képezik itt a szabadsági fokokat. Az elem polinomtere kifeszíthető az olyan  $x_1^{\gamma_1} x_2^{\gamma_2} x_3^{\gamma_3}$  alakú polinomokból, ahol minden  $j = 1, 2, 3$ -ra  $\gamma_j = 0$  vagy  $\gamma_j = 1$ . Így az elem elsőrendű, és  $P(\overline{\Omega}_i)$  polinomterének a következő bázisa van, ha  $\overline{\Omega}_i$  csomópontjai az áttekinthetőség kedvéért azok a  $(\nu_1, \nu_2, \nu_3)$  pontok, ahol minden  $j$ -re  $\nu_j = \pm 1$ , egymástól független előjelekkel :

$$w_{\nu_1, \nu_2, \nu_3}(x) := \frac{1}{8}(1 + \nu_1 x_1)(1 + \nu_2 x_2)(1 + \nu_3 x_3).$$

A kétdimenziós  $R_8$ -elem 3 dimenzióra való általánosításához, a 20 szabadsági fokú  $R_{20}^3$ -elemhez ld. a 43. feladatot.

Minden eddig tárgyalt elem Lagrange-típusú. Ha az elem valamelyik funkcionálja derivált értékét ad meg, akkor *Hermite-elem*nek nevezzük. Ezek pl. negyedrendű peremérték, valamint olyan interpolációs feladatok megoldásánál használatosak, ahol a követelmény az, hogy a függvény mindenütt differenciálható legyen ( $C^1$ -interpoláció); ld. a 44. feladatot.

Az Hermite-elemeknél a következőképpen alakulnak a pont elején adott, az elemet definiáló felsorolás részletei, ha minden csomópontban ugyanannyi derivált megadását tételezzük fel.

Az 1. és 2. pont változatlanul érvényes.

3'. Minden  $x^\ell$  csomóponthoz hozzárendelünk  $r + 1$  lineáris funkcionált :  $\{\psi_\ell^{(k)}\}_{k=0}^r$ . Ekkor  $\psi_\ell^{(0)}$  megadja az  $x^\ell$ -beli függvényértéket :  $\psi_\ell^{(0)}(u) = u(x^\ell)$ , a

$\psi_\ell^{(k)}$  funkcionál a  $k$ -adik derivált értékét adja,  $k = 1, \dots, r$ . Ezek a funkcionálok a szabadsági fokok. Az  $\overline{\Omega}_i$ -hez tartozó szabadsági fokok halmaza ezután  $\psi_i := \{\{\psi_\ell^{(k)}\}_{k=0}^r, x^\ell \in \overline{\Omega}_i\}$ . A lokális sorszámozás szerint  $\psi_{i,j}^{(k)}$ -val jelöljük a szabadsági fokokat.

4'. Minden  $\psi_\ell^{(k)}$ -hoz hozzárendelünk egy  $w_\ell^{(k)} \in V$  bázisfüggvényt úgy, hogy

a)  $w_\ell^{(k)}$  leszűkítése  $\overline{\Omega}_i$ -re legyen polinom (=  $p_{i,\ell}^{(k)}$ );

b) minden egyes  $\overline{\Omega}_i$ -n a  $\{p_{i,\ell}^{(k)}\}$  polinomok képezik a  $P(\overline{\Omega}_i)$  lineáris halmazt, és az  $\overline{\Omega}_i$ -hez tartozó

$$\{x^{i,1}, \dots, x^{i,n_i}\} = \{x^{\ell_1}, \dots, x^{\ell_{n_i}}\}$$

csomópontok  $\psi_{\ell_j}^{(k)} = \psi_{i,j}^{(k)}$  funkcionáljaival teljesül az, hogy a

$$\sum_{k=0}^r \sum_{j=1}^{n_i} p_{i,\ell_j}^{(k)}(x) \alpha_{jk}^{(i)} =: p(x)$$

polinom egyértelműen meg van határozva a

$$\{\psi_{\ell_1}^{(k)}(p), \dots, \psi_{\ell_{n_i}}^{(k)}(p)\}_{k=0}^r$$

értékekkel, mert érvényes

$$\psi_{\ell_j}^{(n)}(w_{\ell_m}^{(k)}) = \delta_{jm} \delta_{nk}, \quad 0 \leq n, k \leq r, \quad j, m = 1, \dots, n_i;$$

c) ha  $x^{\ell_1}, \dots, x^{\ell_{n_i(s)}}$  azok a csomópontok, amelyek  $\overline{\Omega}_i$ -nek egy  $s$  oldalán vannak, akkor a

$$\sum_{k=0}^r \sum_{j=1}^{n_i(s)} p_{i,\ell_j}^{(k)}(x) \alpha_{jk}^{(i)} = q(x)$$

polinom egyértelműen meg legyen határozva  $\overline{\Omega}_i$  oldalán  $\{\psi_{i,\ell_j}^{(k)}(q)\}$  megadása után ( $k = 0, \dots, r, j = 1, \dots, n_i(s)$ ). A többi  $p_{i,\ell_j}^{(k)}$  polinom ( $j = n_i(s) + 1, \dots, n_i, k = 0, \dots, r$ ) legyen azonosan nulla az  $s$  oldalon.

Ha  $\psi_m^{(n)}(w_\ell^{(k)}) = \delta_{m\ell} \delta_{nk}$ , akkor a  $V_h$  végeselem tér bázisa  $\{w_\ell^{(k)}\}_{k=0, \ell=1}^{r, N}$ .

A fenti jelölésekkel most az Hermite-féle véges elemről is elmondható, hogy az a résztartomány, a hozzá tartozó funkcionálok halmaza és a polinomtér által képzett  $\{\overline{\Omega}_i, \psi_i, P(\overline{\Omega}_i)\}$  hármas.

9. példa : (egy dimenzióban) az Hermite-bázis II. 11.6.7-ből;

10. példa : (két dimenzióban) a Felippa–Bell elem ( $FB_{18}$ ), amely folytonosan deriválható függvényekre vezet és így negyedrendű differenciálegyenletek megoldására alkalmas. 18 szabadsági fokát a háromszög minden csúcsában a függvényérték, a két elsőrendű parciális derivált és a három másodrendű parciális derivált alkotja. Ehhez azt a mellékfeltételt kell hozzátenni, hogy az oldalak mentén a normálderivált harmadfokú legyen. Ez a 21 feltétel egyértelműen határoz meg egy ötödfokú polinomot. Az elem polinomterére igaz  $P_4(\bar{\Omega}_i) \subset P(\bar{\Omega}_i) \subset P_5(\bar{\Omega}_i)$ . Bázisának explicit képleteit ld. Mitchell és Wait könyvében, 74–75. o.

Még magasabb rendű elemeket is lehet konstruálni, ld. V.G. Kornyejev, ill. Szabó és Babuška könyvét. Ilyen elemekre van szükség a *végelem módszer  $p$ -verziójában*, amikor a konvergenciát nem  $h \rightarrow 0$ -val érjük el, hanem a polinomterek rendjének növelésével. Ez a verzió kevésbé ritka mátrixokra vezet (szélesebb a sávjuk, amit már a „Velem” program magasabbrendű elemeinél is meg lehet figyelni), gyakran lényegesen gyorsabb konvergenciát ad, mint az itt tárgyalt, úgynevezett  $h$ -verzió – de jobb felkészültséget is kíván.

### 15.7.5 A végelem módszer pontossága

A Céa–lemma akkor alkalmazható, ha az  $a$  forma bilineáris és pozitív definit (ld. II. 241. o.), és azt biztosítja, hogy a végelem módszer  $u - u_h$  hibája lényegében ugyanakkora, mint a legjobb közelítése :

$$\|u - u_h\|_V \leq C \inf_{v_h \in V_h} \|u - v_h\|_V, \quad C := \frac{M_a}{m_a}. \quad (15.214)$$

Mivel a legjobb közelítés nehezen szerezhető meg, helyette vesszük az  $u$  pontos megoldás  $\Pi_h u$  interpolációját a végelem bázis segítségével :

$$\|u - u_h\|_V \leq C \|u - \Pi_h u\|_V. \quad (15.215)$$

A  $\Pi_h u$  interpoláció egy-egy  $\bar{\Omega}_i$  végelemre azonnal definiálható a következő (Lagrange–elemek esetére felírt) képlettel :

$$\begin{aligned} \Pi_h u(x)|_{\Omega_i} &:= \sum_{j=1}^{n_i} u(x^{i,j}) p_{i,\ell_j}(x) \\ &= \sum_{j=1}^{n_i} u(x^{\ell_j}) w_{\ell_j}(x) = \Pi_{h,i} u(x), \quad x \in \bar{\Omega}_i, \end{aligned} \quad (15.216)$$

ahol először a csomópontok lokális, majd globális megszámozását alkalmaztuk és ahol  $\Pi_{h,i}$  egy lokális ( $k$ -adrendű) interpolációs operátor. Figyelembe

véve, hogy az elemek peremén a  $w_\ell$  bázisfüggvények folytonosak (ld. 15.7.4-ben a 4.c követelményt), innen kapjuk a  $\Pi_h u$  definícióját az egész (töröttvonalal határolt)  $\bar{\Omega}$  tartományra.

(15.215)-tel a konvergencia kérdését arra a – peremérték feladattól független – approximáció-elméleti kérdésre vezettük vissza, vajon hogyan viselkedik az interpoláció hibája  $h \rightarrow 0$  esetén.

Az alábbiakban ezt a kérdést – és vele a konvergenciát – tisztázzuk Lagrange-elemek esetén a következő úton, feltéve, hogy  $\Omega$  kétdimenziós poligonális és triangulációját szűkebb értelemben értve, háromszöges felosztásként :

1. Leírjuk a tartomány felosztását szolgáló triangulációkból álló,  $h \rightarrow 0$ -val kapcsolatos sorozatoknak egy bizonyos (egyenletesen regulárisnak nevezett) osztályát;
2. Vizsgálunk a síkon egy speciális interpolációs problémát;
3. Bebizonyítunk egy az  $I_k(\Omega)$  interpolációs operátorra vonatkozó becslést;
4. Megmutatjuk a konvergenciát akkor, amikor a trianguláció kongruens háromszögekből áll;
5. Belátjuk a konvergenciát egyenletesen reguláris trianguláción.

Most nézzük a részleteket.

1. A kétdimenziós esetben a tartomány triangulációinak nem minden sorozata elfogadható : azt kell biztosítani, hogy a trianguláció  $h \rightarrow 0$ -val ne fájuljon el. Ezt a következő definícióval fogalmazzuk meg pontosabban.

**Definíció.** Legyen  $\{\mathcal{T}_{h_\ell}\}_{\ell=1}^\infty$  az  $\bar{\Omega}$  tartomány  $\{\bar{\Omega}_i^\ell\}_{i=1}^{M_\ell}$  triangulációinak egy sorozata, amelynek  $h_\ell := \max_i \text{diam}(\bar{\Omega}_i^\ell)$  paraméterei nullsorozatot alkotnak (úgy hogy  $M_\ell \rightarrow \infty$ , amikor  $h_\ell \rightarrow 0$ ). Ezen  $\{\mathcal{T}_{h_\ell}\}$  sorozatnak minden triangulációja legyen konform. Ekkor a sorozatot *egyenletesen regulárisnak* hívjuk, ha van olyan  $h_\ell$ -től független  $\kappa > 0$  konstans, amellyel teljesül

$$h_\ell \leq \kappa \rho(\bar{\Omega}_i^\ell) \text{ minden } 1 \leq i \leq M_\ell\text{-re és minden } \ell\text{-re.} \quad (15.217)$$

Itt  $\rho(\bar{\Omega}_i^\ell)$  a legnagyobb  $\bar{\Omega}_i^\ell$ -be beírható kör sugara.  $\square$

Egyszerűsített módon azt is fogjuk mondani, hogy maga a trianguláció egyenletesen reguláris (elképzelve, hogy ő egy egész sorozatnak a tagja), az  $\ell$  indexet is elhagyva.

**Megjegyzés.** (15.217) helyett azt is követelhetjük, hogy az  $\bar{\Omega}_i^\ell$  elemek oldalai által bezárt  $\vartheta_i^{(\ell)}$  legkisebb belső szög teljesítse a következő feltételt :

$$\sin \vartheta_i^{(\ell)} \geq \frac{1}{\kappa} \text{ minden } i\text{-re,} \quad (15.218)$$



$h$ -tól független  $\kappa > 0$  konstanssal, vagy pedig a  $0 < \text{const} \leq \vartheta_i^{(\ell)} \leq \pi - \text{const}$  feltételt.  $\square$

Ehhez ld. a 45. és 46. feladatot. Ahhoz a témához, hogyan lehet a háromszögek “minőségét” értékelni, ld. Pébay és Baker cikkét is.

**2.** Abból indulhatunk ki, hogy a  $k$ -adfokú kétváltozós

$$p(x, y) = \sum_{i+j \leq k} a_{ij} x^i y^j$$

polinomnak  $n = n(k) = \binom{k+2}{2}$  együtthatója van, és ezeknek meghatározása  $n$  (páronként különböző) pontban adott érték alapján elképzelhető, de nem mindig lehetséges.

Speciális interpolációs problémánk a következő : Dobjunk a síkra  $k + 1$  egyenest :  $\ell_0, \dots, \ell_k$ , mégpedig úgy, hogy bármely tetszőleges két egyenes legfeljebb egy pontban metszi egymást. Az  $\ell_0$  egyenesen egy pontot jelöljünk ki (legyen ez  $(x_{0,0}, y_{0,0})$ ),  $\ell_1$ -en kettőt (legyen ez  $(x_{1,0}, y_{1,0})$  és  $(x_{1,1}, y_{1,1})$ ) stb., végül  $\ell_k$ -n  $k + 1$  pontot  $((x_{k,0}, y_{k,0}), \dots, (x_{k,k}, y_{k,k}))$ , és egyik pont se legyen az egyenesek metszéspontja. Akkor pontosan egy (legfeljebb)  $k$ -adfokú kétváltozós polinom létezik, amely az  $(x_{i,j}, y_{i,j})$  pontban az adott  $f_{i,j}$  függvényértéket interpolálja,  $i = 0, \dots, k$ ,  $j = 0, \dots, i$ .

Ezt indukcióval látjuk be : Az első pontban adott érték konstans polinommal interpolálható. Tegyük fel, hogy a  $k-1$ -edfokú polinom meghatározhatósága az  $\ell_0, \dots, \ell_{k-1}$  egyenesen fekvő  $n(k-1) = \frac{1}{2}(k+1)k$  pontból ismert. Vegyük az  $\ell_k$  egyenest, amelynek képlete legyen  $\alpha_k x + \beta_k y + \gamma_k = 0$ , és annak mentén,  $n(k) - n(k-1) = k + 1$  pontja alapján a (legfeljebb  $k$ -adfokú, egydimenziós)  $L_k(x, y)$  Lagrange-féle polinomot számítjuk ki. Ekkor az interpolációs feladat megoldása (az I (4.14) Newton-rekurzió ötletét használva)

$$p_k(x, y) = L_k(x, y) + (\alpha_k x + \beta_k y + \gamma_k) p_{k-1}(x, y),$$

ahol a  $p_{k-1}$  polinomot indukció szerint ki tudjuk számítani abból, hogy

$$p_{k-1}(x_{i,j}, y_{i,j}) = \frac{f_{i,j} - L_k((x_{i,j}, y_{i,j}))}{\alpha_k x_{i,j} + \beta_k y_{i,j} + \gamma_k}, \quad i = 0, \dots, k-1, \quad j = 0, \dots, i.$$

**Megjegyzés.** A legfontosabb helyzet az, amikor a  $k+1$  egyenes párhuzamos (standard Lagrange-elem), de előfordulnak ettől eltérő esetek is.  $\square$

**3.** Legyen  $\Omega \in \mathbb{R}^2$  Lipschitz-folytonos tartomány (ez lesz majd a mintaelem tartománya),  $k \geq 1$ ,  $\overline{\Omega}$ -ban adottak a  $z_1, \dots, z_{n(k)}$  pontok, amelyek a fenti leírásnak megfelelően  $k + 1$  egyenesen fekszenek és amelyekben az  $u \in$

$H^{k+1}(\Omega)$  Szoboljev-térbeli függvény értékeit írjuk elő. Ekkor az interpoláció eredményét, a  $k$ -adfokú polinomot  $I_k(\Omega)u$ -val jelölve, igaz az alábbi becslés :

$$\|u - I_k(\Omega)u\|_{k+1} \leq c_0(\Omega, k)|u|_{k+1}, \quad (15.219)$$

ahol  $\|\cdot\|_{k+1}$ , ill.  $|\cdot|_{k+1}$  a Szoboljev-tér normája, ill. félnormája.

A bizonyítás (ld. Braess könyvét) azzal indul, hogy belátjuk :

$$\|v\|_{k+1} := |v|_{k+1} + \sum_{i=1}^{n(k)} |v(z_i)|, \quad v \in H^{k+1}(\Omega),$$

normát definiál  $H^{k+1}(\Omega)$ -ban, és  $\|v\|_{k+1}$  ekvivalens  $\|\cdot\|_{k+1}$ -gyel. Ehhez megjegyezzük elsőnek, hogy  $k+1 \geq 2$  és a beágyazási tételek miatt a  $H^{k+1}(\Omega)$ -beli függvények folytonosak, így a képlet jól definiált, továbbá  $|v(z_i)| \leq c_1(\Omega)\|v\|_{k+1}$  minden  $i$ -re, tehát

$$\|v\|_{k+1} \leq (1 + c_1 n(k))\|v\|_{k+1},$$

és ekkor  $\|v\|_{k+1}$  már félnorma  $H^{k+1}(\Omega)$ -ban. Ha most  $\|v\|_{k+1} \leq c_0\|v\|_{k+1}$  hamis lenne minden  $c_0$  és  $v \in H^{k+1}(\Omega)$  esetén, akkor létezne olyan  $\{v_m\} \subset H^{k+1}(\Omega)$  sorozat, hogy 1)  $\|v_m\|_{k+1} = 1$  és 2)  $\|v_m\|_{k+1} \leq \frac{1}{m}$ ,  $m = 1, 2, \dots$

Vizont ekkor 1) miatt a  $\{v_m\}$  sorozatnak egy részsorozata  $H^k(\Omega)$ -ban konvergens lenne (kompakt beágyazás) – amelyet újra  $\{v_m\}$ -nek jelölhetünk. Mivel  $H^k(\Omega)$  teljes, ott  $\{v_m\}$  Cauchy-sorozat. De 2)-ből és a definíció szerinti  $|v_m|_{k+1} \leq \|v_m\|_{k+1}$ -ből következik  $|v_m|_{k+1} \rightarrow 0$ , így  $\|v_m - v_\ell\|_{k+1}^2 \leq \|v_m - v_\ell\|_k^2 + 2(|v_m|_{k+1}^2 + |v_\ell|_{k+1}^2) \rightarrow 0$ , vagyis :  $\{v_m\}$  Cauchy-sorozat  $H^{k+1}(\Omega)$ -ban is. Ezután már  $H^{k+1}(\Omega)$  teljessége alapján  $\{v_m\}$ -nek van  $v_*$  limesze  $H^{k+1}(\Omega)$ -ban, és 2) miatt teljesül :  $0 = \|v_*\|_{k+1} \geq |v_*|_{k+1}$ . De ez ellentmondásra vezet, ugyanis ekkor  $v_*$   $k$ -adfokú polinom kellene, hogy legyen, és  $\|v_*\|_{k+1}$  definíciója alapján  $v_*(z_i) = 0$  is következne minden  $i$ -re, ami  $I_k(\Omega)$  egyértelműsége miatt  $v_* = 0$ -t jelentene – ellentmondással 1)-hez, amely szerint  $\|v_*\|_{k+1} = 1$ . Így kell léteznie egy  $c_0$  konstansnak úgy, hogy

$$\|v\|_{k+1} \leq c_0\|v\|_{k+1} \text{ minden } v \in H^{k+1}(\Omega)\text{-ra,}$$

ami egy *Poincaré-egyenlőtlenség*. Innen  $v = u - I_k(\Omega)u$ -val kapjuk

$$\begin{aligned} \|u - I_k(\Omega)u\|_{k+1} &\leq c_0\|u - I_k(\Omega)u\|_{k+1} \\ &= c_0\left(|u - I_k(\Omega)u|_{k+1} + \sum_{i=1}^{n(k)} |(u - I_k(\Omega)u)(z_i)|\right) = c_0|u|_{k+1}, \end{aligned}$$

hiszen  $(u - I_k(\Omega)u)(z_i) = 0$  és  $|I_k(\Omega)u|_{k+1} = 0$ .

**Megjegyzés.** A (15.219) becslés a Bramble–Hilbert lemmának (v.ö. II 11.28. lemma) egy formája, hiszen az  $L_k := I - I_k(\Omega) : H^{k+1} \rightarrow H^{k+1}$  operátor magtere a  $P_k(\bar{\Omega})$  polinomtér. (Itt  $I$  az identitást jelöli.)  $\square$

4. Legyen a mintaelemünk az  $S$  egység-szimplex :  $0 < \xi_1 < 1 - \xi_2 < 1$ , és  $S_h := hS$  a  $0 < h \leq 1$ -méretű hasonmása. Olyan  $\Omega$  tartományt és olyan triangulációját vizsgálunk, amelynek minden eleme  $S_h$ , csak esetleg el van forgatva (ami nem fog szerepet játszani, mert háromszögenként dolgozunk). Legyen  $k \geq 1$ , a mintaelemen  $I_k(S)$  a fenti interpolációs operátor, amelynek  $S_h$ -n az  $I_k(S_h) = \Pi_{h,i}$  lokális operátor felel meg, és  $\Pi_h$  a globális (15.216) interpolációs operátor. Ekkor igaz a következő becslés :

$$\|u - \Pi_h u\|_m \leq c_0 h^{k+1-m} |u|_{k+1}, \quad 0 \leq m \leq k+1, \quad c_0 = c_0(\Omega, k). \quad (15.220)$$

**Bizonyítás.** Az alábbiakban érdemes a normákhoz tartozó mindenkori tartományt is említeni.

Ha  $u \in H^{k+1}(S_h)$ , akkor  $v(\xi) := u(h\xi) = u(x)$  definiálja a  $v \in H^{k+1}(S)$  függvényt. Itt  $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = h \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = h\xi$ . Továbbá, a deriváltakra kapjuk multiindex-jelöléssel  $\partial_\xi^\alpha = h^{|\alpha|} \partial_x^\alpha$  ( $|\alpha| = \alpha_1 + \alpha_2$ ), és így

$$|v|_{m,S}^2 = \sum_{|\alpha|=m} \int_S (\partial_\xi^\alpha v)^2 d\xi \stackrel{x=h\xi}{=} \sum_{|\alpha|=m} \int_{S_h} h^{2m} (\partial_x^\alpha u)^2 h^{-2} dx = h^{2(m-1)} |u|_{m,S_h}^2. \quad (15.221)$$

Innen

$$\|u\|_{m,S_h}^2 := \sum_{j \leq m} |u|_{j,S_h}^2 = \sum_{j \leq m} h^{2(1-j)} |v|_{j,S}^2 \leq h^{2(1-m)} \|v\|_{m,S}^2,$$

mert  $h \leq 1$ . Ezt a becslést  $u - I_k(S_h)u$ -ra is alkalmazhatjuk :

$$\|u - I_k(S_h)u\|_{m,S_h} \leq h^{1-m} \|v - I_k(S)v\|_{m,S} \leq h^{1-m} \|v - I_k(S)v\|_{k+1,S},$$

hiszen  $\|\cdot\|_{m,S}^2$  egyszerűen a  $\|\cdot\|_{k+1,S}^2$  része. Most emlékezzünk a (15.219) egyenlőtlenségünkre :

$$\|v - I_k(S)v\|_{k+1,S} \leq c_0 |v|_{k+1,S},$$

ahol a  $c_0$  konstans nem függ  $h$ -tól. Itt a  $|v|_{k+1,S}$  félnormát (15.221) segítségével  $|u|_{k+1,S_h}$ -ra átszámítva következik

$$\|v - I_k(S)v\|_{k+1,S} \leq c_0 h^k |u|_{k+1,S_h},$$

és összefoglalva

$$\|u - I_k(S_h)u\|_{m,S_h} \leq c_0 h^{k+1-m} |u|_{k+1,S_h}.$$

Most már csupán ezt az egyenlőtlenséget kell négyzetre emelni és  $\bar{\Omega} = \cup_i \bar{S}_{h,i}$  esetén összegezni, hogy (15.220) következzen.  $\square$

**Megjegyzés.** Ahogyan látjuk, a (15.220)-ban az  $m < k + 1$  esetben pozitív  $h$ -hatvány onnan ered, hogy a bal oldalon az  $\|\cdot\|_{m,\Omega}$  normát, a jobb oldalon  $\|\cdot\|_{k+1,\Omega}$ -t használjuk. Példák mutatják, hogy ezt a  $h$ -hatványt nem lehet javítani.  $\square$

5. Most foglalkozunk az általános ( $\Delta$ -val jelölt, nem elfajult) háromszöggel. Erre az  $S$  mintaelemet affin leképezéssel transzformálhatjuk :

$$S \rightarrow \Delta : \quad \xi \in S \rightarrow x = x(\xi) = x_0 + B\xi \in \Delta, \quad (15.222)$$

ahol  $x_0$  a  $\Delta$  háromszög egy csúcsa és  $B$  reguláris  $2 \times 2$ -es mátrix. Ekkor a  $\xi = \xi(x)$  inverz leképezés is affin, és az  $S$ -beli és  $\Delta$ -beli polinomok megfelelnek egymásnak.

Ha  $u \in H^m(\Delta)$ , akkor  $v(\xi) := u(x(\xi)) \in H^m(S)$ , és (15.221) általánosításaként érvényes ( $\|B\|_F := (\sum_{i,k=1,2} b_{ik}^2)^{1/2}$  a Frobenius-normát) jelenti)

$$|v|_{m,S} \leq \|B\|_F^m |\det B|^{-1/2} |u|_{m,\Delta}. \quad (15.223)$$

Megfordítva, ha  $v \in H^m(S)$ , akkor  $u(x) := v(\xi(x)) \in H^m(\Delta)$ , és

$$|u|_{m,\Delta} \leq \|B^{-1}\|_F^m |\det B|^{1/2} |v|_{m,S}. \quad (15.224)$$

A két relációból elegendő pl. az elsőt bizonyítani, mert  $B^{-1}$  közvetíti az inverz leképezést. Az  $m = 0$  eset világos : a (15.222) transzformáció Jacobi-mátrixa  $B$ . Tekintsük az  $m = 1$  esetet, a láncszabályt alkalmazva :

$$(\nabla_\xi v)_i = \frac{\partial v}{\partial \xi_i} = \sum_{j=1,2} \frac{\partial u}{\partial x_j} \frac{\partial x_j}{\partial \xi_i} = \sum_{j=1,2} b_{ij}^T \frac{\partial u}{\partial x_j} = (B^T \nabla_x u)_i, \quad i = 1, 2, \quad (15.225)$$

avagy

$$\nabla_\xi v = B^T \nabla_x u.$$

Ennek négyzetét integrálva, a (15.222) transzformációt és a  $\|B^T\| = \|B\| \leq \|B\|_F$  egyenlőtlenséget használva (ahol  $\|\cdot\|$  az euklideszi norma), kapjuk

$$|v|_{1,S}^2 = \int_S \|\nabla_\xi v\|^2 d\xi = \int_\Delta \|B^T \nabla_x u\|^2 \frac{1}{|\det B|} dx \leq \|B\|_F^2 |\det B|^{-1} |u|_{1,\Delta}^2.$$

Vizsgáljuk az  $m = 2$  esetet. (15.225) alapján :

$$\frac{\partial^2 v}{\partial \xi_i \partial \xi_j} = \sum_{k=1,2} b_{ik}^T \sum_{\ell=1,2} \frac{\partial^2 u}{\partial x_k \partial x_\ell} \frac{\partial x_\ell}{\partial \xi_j} = \sum_{k=1,2} b_{ik}^T \sum_{\ell=1,2} b_{j\ell}^T \frac{\partial^2 u}{\partial x_k \partial x_\ell}, \quad i, j = 1, 2.$$

Innen kapjuk a  $D := \sum_{i=1,2} \sum_{j=1,2} \left( \frac{\partial^2 v}{\partial \xi_i \partial \xi_j} \right)^2 =: \sum_{j=1,2} D_j$  összeg

$$D_j = \sum_{i=1,2} \left( \sum_{k=1,2} b_{ik}^T a_k^{(j)} \right)^2 \quad \text{és} \quad a_k^{(j)} := \sum_{\ell=1,2} b_{j\ell}^T \frac{\partial^2 u}{\partial x_k \partial x_\ell}$$

tagjainak becslése Cauchy-egyenlőtlenségek segítségével :

$$D_j \leq \|B\|_F^2 \sum_{k=1,2} (a_k^{(j)})^2 \quad \text{és} \quad \sum_{j=1,2} (a_k^{(j)})^2 \leq \|B\|_F^2 \sum_{\ell=1,2} \left( \frac{\partial^2 u}{\partial x_k \partial x_\ell} \right)^2,$$

ahonnan  $D \leq \|B\|_F^4 \sum_{k=1,2} \sum_{\ell=1,2} \left( \frac{\partial^2 u}{\partial x_k \partial x_\ell} \right)^2$  következik. Ennek  $S$  feletti integrációja és a (15.222) transzformáció figyelembe vétele adja a (15.223) reláció  $m = 2$  esetét. A fentieknek megfelelő gondolatmenettel kapjuk az egyenlőtlenséget akkor is, amikor  $m > 2$ .

Most minden készen áll a végeelem módszer interpolációs hibájának becsléséhez :

az  $u - I_k(\Delta)u$  hibát mérjük a  $H^m(\Delta)$  félnormájában;

ezt transzformáljuk az  $S$  egység-szimplexre;

ott alkalmazzuk a (15.219) becslést (a Bramble–Hilbert lemmát) a  $h$ -tól független  $c_0(S, k) = c_0(k)$  konstanssal;

a jobboldalon megjelenő  $|v|_{k+1,S}$  félnormát visszatranszformáljuk  $\Delta$ -re :

$$\begin{aligned} |u - I_k(\Delta)u|_{m,\Delta} &\leq \|B^{-1}\|_F^m |\det B|^{1/2} |v - I_k(S)v|_{m,S} \\ &\leq \|B^{-1}\|_F^m |\det B|^{1/2} c_0 |v|_{k+1,S} \\ &\leq \|B^{-1}\|_F^m |\det B|^{1/2} c_0 \|B\|_F^{k+1} |\det B|^{-1/2} |u|_{k+1,\Delta} \\ &= c_0(k) (\|B^{-1}\|_F \|B\|_F)^m \|B\|_F^{k+1-m} |u|_{k+1,\Delta}. \end{aligned} \quad (15.226)$$

Hátra maradt a  $\|B\|_F$  és a  $\|B^{-1}\|_F$  becslése – amihez a (15.217)-beli  $\rho(\overline{M})$  jelölést használjuk  $M = S, \Delta$ -ra. Először azt fogjuk belátni, hogy az euklideszi normában igaz

$$\|B(\xi - \eta)\| \leq \frac{\text{diam}(\overline{\Delta})}{2\rho(\overline{S})} \|\xi - \eta\|, \quad \|B^{-1}(x - y)\| \leq \frac{\text{diam}(\overline{S})}{2\rho(\overline{\Delta})} \|x - y\|, \quad (15.227)$$

ahol  $\xi, \eta \in \overline{S}$  tetszőlegesen és  $x = x(\xi), y = x(\eta) \in \overline{\Delta}$  a (15.222) transzformáció alatt. Ekkor ugyanis elegendő a  $\|\xi - \eta\| = 2\rho(\overline{S})$  esetet nézni, amikor tudjuk, hogy  $\|x - y\| = \|B(\xi - \eta)\| \leq \text{diam}(\overline{\Delta})$  – és az első reláció következik. Hasonlóan a második is igaz. Felsorolva a (15.227)-beli mennyiségeket, ill. a (15.217) egyenletes regularitás szerinti becsléseit :

$$\text{diam}(\overline{S}) = \sqrt{2}, \quad \rho(\overline{S}) = \frac{1}{2 + \sqrt{2}}, \quad \text{diam}(\overline{\Delta}) \leq h_\ell, \quad \rho(\overline{\Delta}) \geq \frac{h_\ell}{\kappa}, \quad (15.228)$$

az következik  $\|B\|_F \leq \sqrt{2}\|B\|$  figyelembe vételével ( $2 = d$  a  $B$  mátrix dimenziója), hogy

$$\|B\|_F \leq h_\ell(1+\sqrt{2}), \quad \|B^{-1}\|_F \leq \frac{\kappa}{h_\ell}, \quad \text{és így } \text{cond}_F(B) \leq (1+\sqrt{2})\kappa. \quad (15.229)$$

Összesítve a (15.226)-(15.229) egyenlőtlenségeket, látjuk, hogy igaz

$$|u - I_k(\Delta)u|_{m,\Delta} \leq c_0(k)\kappa^m(1+\sqrt{2})^{k+1}h_\ell^{k+1-m}|u|_{k+1,\Delta},$$

amiből négyzetre emeléssel és  $\Delta = \Omega_i^\ell$  behelyettesítéssel, majd  $i$  feletti összegzéssel (ehhez ld. a 15.31. tétel utáni első megjegyzést) adódik az alábbi eredmény.

**15.31. Tétel** ( $k$ -adfokú Lagrange-féle interpoláció hibája egyenletesen reguláris háromszöges trianguláción). Teljesüljenek a következő feltételek :

1.  $\Omega \in \mathbb{R}^2$  korlátos tartomány,  $\Gamma$  pereme poligonális;
2. Az interpolálandó  $u$  függvény elég reguláris :  $u \in H^{k+1}(\Omega)$ , ahol  $k \geq 1$ ;
3. Az  $\bar{\Omega} = \cup_{i=1}^M \bar{\Omega}_i$  tartomány  $\mathcal{T}_h$  triangulációja egyenletesen reguláris  $\kappa$  konstanssal, ld. (15.217), és  $h = \max_i \text{diam}(\bar{\Omega}_i) \leq 1$ ;
4. Az  $m \leq k+1$  egészszám olyan, hogy  $\Pi_h u \in H^m(\Omega)$ .

Akkor  $0 \leq \ell \leq m$ -re igazak a következő becslések :

$$|u - \Pi_h u|_{H^\ell(\Omega)} \leq C_0 h^{k+1-\ell} |u|_{H^{k+1}(\Omega)}, \quad C_0 := C_0(k, m, \kappa), \quad (15.230)$$

$$\|u - \Pi_h u\|_{H^m(\Omega)} \leq C_1 h^{k+1-m} |u|_{H^{k+1}(\Omega)}, \quad C_1 := C_0 \sqrt{1+m}. \quad \square \quad (15.231)$$

A (15.231) eredmény  $h \leq 1$  miatt az előzőből úgy kapható, hogy azt négyzetre emeljük, majd  $0 \leq \ell \leq m$ -re összegzük.

**Megjegyzések.** **1.** Abból, hogy  $\Pi_{h,i}u$  polinom  $\Omega_i$ -n és így bizonyára  $\Pi_{h,i}u \in H^m(\Omega_i)$ , még nem következik  $\Pi_h u \in H^m(\Omega)$ , ld. a 3. megjegyzést (15.7.4)-ben. A 4. feltétel viszont  $m = 0$ -ra és konform elemek esetén legalább  $m = 1$ -re teljesül.

**2.** A tétel általánosítása három dimenzióra problémamentes, ha tetraédes elemekről van szó.

**3.** Kettő dimenzióban a négyszöges elemekre való általánosítás akkor nem gond, ha paralelogrammákra korlátozzunk magunkat : ekkor még működik a fent bemutatott, Ciarlet-re visszamenő "affin végeelem családok elmélete".  $S$  helyett a mintaelem ekkor az  $E$  egységnégyzet, amely affin transzformációval a paralelogrammára képezhető le. Az  $\bar{E}$ -n definiált  $R_4, R_8$  stb. véges elemhez tartozó  $I_k^E$  interpolációs operátor viszont nem az  $\Omega = E$  esetre fentiek szerint konstruált és (15.219)-ben szereplő  $I_k(E)$  operátor : utóbbinál az  $I - I_k(E)$  magtere éppen  $P_k(\bar{E})$ , míg  $I - I_k^E$  magtere ennél nagyobb. A négyszöges

elemeknél ugyanis nincs olyan  $k$  szám, hogy  $P(\overline{E}) = P_k(\overline{E})$ , de van olyan  $k$ , hogy  $P_k(\overline{E}) \subset P(\overline{E})$  és  $P_{k+1}(\overline{E}) \not\subset P(\overline{E})$  ( $k = 1 : R_4$ ,  $k = 2 : R_8$ , ld. részletesen a 15.7.4. pontot).

Emiatt arra hivatkozunk, hogy van olyan lineáris  $L_k : H^{k+1}(E) \rightarrow H^{k+1}(E)$  operátor, hogy  $L_k$  magtere *tartalmazza* a  $P_k(\overline{E})$  polinomokat (mégpedig  $L_k = I - I_k^E$ ), tehát  $L_k I_k(E) = 0$ . Ekkor (15.219) alapján  $\|L_k v\|_{k+1} = \|L_k(I - I_k(E))v\|_{k+1} \leq \|L_k\| \|v - I_k(E)v\|_{k+1} \leq \|L_k\| c_0 |v|_{k+1}$ .

Egyébként a (15.222) transzformációval kapcsolatos eredményeink érvényben maradnak és a (15.228)-beli konstansok  $S \rightarrow E$ -nek megfelelően változnak (ld. a 47. feladatot is).

4. Általános (konvex) négyszöges elemeknél (és a magasabb rendű háromszöges izoparametrikus elemek esetén) többféle bonyodalom adódik : az  $S \rightarrow \Omega_i$ -féle,  $F_i$ -vel jelölt leképezés nem lineáris, az  $\Omega_i$ -n adott  $u(x)$  polinomnak a mintaelemen megfelelő  $v(\xi) = (\Phi_i u)(\xi) := u(F_i(\xi))$  függvény általában nem polinom, és (15.226)-(15.227) helyett az  $F_i, F_i^{-1}, \Phi_i, \Phi_i^{-1}$  leképezések Lipschitz-folytonosságával kell foglalkozni. De ekkor is teljesülnek végül a (15.230)-(15.231) becslések, ld. pl. Girault és Raviart könyvét.

5. Amennyiben Hermite-elemekről van szó és a legnagyobb (szabadsági fokként előforduló) derivált rendje  $s$ , akkor azt kell követelni  $k \geq 1$  helyett, hogy igaz legyen  $H^{k+1}(\Omega_i) \hookrightarrow C^s(\overline{\Omega}_i)$ , folytonos beágyazással.  $\square$

A 15.31. tételből a Céa-lemma segítségével adódik a konvergencia.

**15.32. Következmény** (Lagrange-féle konform végeselem módszer konvergenciája egyenletesen reguláris trianguláción). Teljesüljenek a 15.31. tétel valamint a II. 11.21. lemma feltételei a (15.189) variációs feladat esetén. Annak  $u$  megoldása legyen  $H^{k+1}(\Omega)$  eleme.

Ekkor a módszer konvergens, mégpedig érvényes a következő becslés :

$$\|u - u_h\|_{H^1(\Omega)} \leq \frac{M_a}{m_a} C(k, \kappa) h^k \|u\|_{H^{k+1}(\Omega)}. \quad (15.232)$$

**Bizonyítás.** A tétel következik a (15.214), (15.215) becslésekből, valamint az előző tételből,  $\ell = 1$  behelyettesítéssel a (15.231) becslésbe.  $\square$

**Megjegyzések. 1.** Emlékeztetünk arra, hogy a 15.7.4-ben megtárgyalt elemek rendje a következő :

elem	$T_3$	$R_4$	$T_6$	$R_8$	$T_{10}$	$FB_{18}$	$T_4^3$	$R_8^3$	$T_{10}^3$	$R_{20}^3$
rendje	1	1	2	2	3	4	1	1	2	2

Egyenletesen reguláris trianguláció esetén, valamint a megoldás megfelelő simasága mellett ( $k = 1$  esetén ld. a 15.30. tételt) ez tehát egyben a konvergenciarend is.

2. A táblázat szerint egyértelműen előnyösebbnek tűnnek a háromszöges és tetraéder elemek a téglalap elemeknél, de erre a kérdésre 15.7.11-ben térünk vissza.

3. A (15.232) becslés javíthatatlan (ami a  $h$  hatványát illeti), ld. Ciarlet vagy Oganieszjan és Ruhovec könyvét.

4. A *Nitsche-fogás* segítségével (ld. II. 11.6.3.) megmutatható, hogy a másodrendű feladat megfelelő regularitása mellett (ezt a 15.30. tétel biztosítja, ha  $k = 1$ ) érvényes az

$$\|u - u_h\|_{L_2(\Omega)} \leq c(k)h^{k+1}\|u\|_{H^{k+1}(\Omega)}$$

becslés, és ez is javíthatatlan.

5. Legyen  $k = 1$ . Ekkor elméleti szempontból a 15.32. következmény szépséghibájának tűnik, hogy a variációs feladat  $u$  megoldása  $H^1$ -ben van, de konvergenciát csak  $H^2$ -ben kapunk. Viszont a  $H^1$ -beli konvergencia ténye innen következik: Mivel  $H^2$  sűrű  $H^1$ -ben, létezik olyan  $u_\varepsilon \in H^2$ , hogy  $\|u - u_\varepsilon\|_{H^1} \leq \frac{\varepsilon m_a}{2M_a}$ . Ez az  $u_\varepsilon$  azon variációs feladat  $w$  megoldása, hogy  $a(w, v) = a(u_\varepsilon, v)$  minden  $v \in H^1$ -re, és  $u_{\varepsilon, h}$  végelem közelítésére igaz (15.232) alapján, hogy  $\|u_\varepsilon - u_{\varepsilon, h}\|_{H^1} \leq \text{const } h\|u_\varepsilon\|_{H^2} \leq \frac{\varepsilon m_a}{2M_a}$ , ha  $h = h(\varepsilon)$  elég kicsi. Emiatt a Céa-lemmából kapjuk  $\|u - u_h\|_{H^1} \leq \frac{M_a}{m_a}\|u - u_{\varepsilon, h}\|_{H^1} \leq \frac{M_a}{m_a}(\|u - u_\varepsilon\|_{H^1} + \|u_\varepsilon - u_{\varepsilon, h}\|_{H^1}) \leq \varepsilon$ , vagyis a konvergenciát - de ez tetszőlegesen lassú lehet és  $h$ -val nem jellemezhető.  $\square$

## 15.7.6 Numerikus integrálás

Eddig abból indultunk ki, hogy az összes, a végelem módszer használatával kapcsolatos integrált pontosan számítjuk ki, így az előző pont becslései is ezt tételezik fel.

Az integrálokat legtöbbször csak közelítőleg tudjuk kiszámítani (és erre megy el a végelem program lényeges futási ideje). Egyrészt az  $A = (a_{kl})$  végelem mátrix  $a_{kl} = a(w_l, w_k)$  elemeiről, másrészt a jobboldallal kapcsolatos  $b_k = \varphi(w_k)$  integrálokról van szó, ld. pl.  $a$  és  $\varphi$  (15.197) definícióját. Ha ezeket az integrálokat kvadratúra képletek segítségével közelítjük (konkrét ilyen lehetőségekre lejjebb térünk ki, alapvető kvadratúra képletekkel már I. 5-ben foglalkoztunk), akkor ezt röviden úgy tudjuk jellemezni, hogy az

$$a(w_l, w_k) \approx a_h(w_l, w_k), \quad \varphi(w_k) \approx \varphi_h(w_k)$$

közelítésekkel dolgozunk, tehát egy közelítő  $a_h$  bilineáris formával és egy közelítő  $\varphi_h$  lineáris funkcionállal.

Itt azonnal két lényeges kérdés merül fel:



1. Az új bilineáris forma és lineáris funkcionál rendelkezik-e minden szükséges tulajdonsággal ahhoz, hogy a keletkező variációs feladatnak, keressünk olyan  $u_h \in V_h$ -t, hogy

$$a_h(u_h, w) = \varphi_h(w) \quad \text{minden } w \in V_h\text{-ra} \quad (15.233)$$

– úgy mint a (15.189) eredetinek – van-e egyértelmű és stabil megoldása  $V_h$ -ban?

2. Milyen jó közelítésre van szükség annak érdekében, hogy adott végeselem esetén annak konvergenciarendjét ne csökkentsük?

Az első kérdéssel kapcsolatban rögtön felmerül az a probléma, hogy az  $a$ -ból és  $\varphi$ -ből kvadratúra képletek segítségével előállított  $a_h$  bilineáris forma és  $\varphi_h$  lineáris funkcionál – eltekintve az egydimenziós esettől – általában  $V$ -n nem lehetnek definiáltak, mivel a  $V = H^1$  esetben a függvények pontbeli értéke nem definiált.

Legyen ezért az  $a_h$  bilineáris forma és a  $\varphi_h$  lineáris funkcionál definiáltak  $V_h \times V_h$ -n, ill.  $V_h$ -n. Ezenkívül legyen az  $a_h$  forma  $h$ -ra nézve *egyenletesen elliptikus*  $V_h$ -ban, azaz létezzen olyan  $h$ -tól független  $\alpha$  szám, hogy a  $V$  tér normájával érvényes

$$a_h(v_h, v_h) \geq \alpha \|v_h\|_V^2 \quad \text{minden } v_h \in V_h\text{-ra.} \quad (15.234)$$

Ekkor a 11.20. tétel alapján a fenti közelítő variációs feladatnak létezik egyértelmű és stabil megoldása  $V_h$ -ban. Ezzel a válasz a fenti első kérdésre megvan.

Az  $a_h$  forma egyenletes ellipticitásának elégséges feltételeit lent adjuk meg, megkülönböztetve a  $P_k(\overline{\Omega}_i)$  és a  $Q_k(\overline{\Omega}_i)$  polinomtereket (ld. a 15.7.4-beli definíciót).

**Definíció.** A háromszöges  $k$ -adfokú végeselemre vonatkozó kvadratúra képlet akkor teljesíti a  $P(k)$ -feltételt, ha vagy egzakt  $P_{2k-2}(\Omega_i)$ -ben, vagy az alappontjaihoz rendelhető szabadsági fokok (vagy részhalmazuk) az összes  $P_{k-1}(\Omega_i)$ -beli polinomot egyértelműen határozzák meg.

A négyszöges tartományon megadott kvadratúra képlet akkor teljesíti a  $Q(k)$ -feltételt, ha az alappontjaihoz rendelhető szabadsági fokok (vagy részhalmazuk) az összes olyan  $P_{d,k-1}(\Omega_i)$ -beli polinomot, amely egyben a  $Q_k(\Omega_i)$ -nak is eleme, egyértelműen határozzák meg.  $\square$

Mint első példát említjük, hogy az I. 5.11-ben található (5.39) képlet (a középpont szabályt általánosítja háromszöges tartományra és amelynek egyetlen alappontja a háromszög súlypontja), egzakt  $P_1(\Delta_i)$ -ben, így  $k = 1$  (azaz  $T_3$ ) esetén teljesíti a  $P(1)$ -feltétel mindkét alternatíváját. A  $P(2)$  feltétel viszont nem érvényes.

Mint második példát vesszük a középpont szabály általánosítását négyszöges tartományra (a képlet egyetlen alappontja a négyszög súlypontja). Ez egzakt  $Q_1(\square_i)$ -ben, de nem teljesíti a  $Q(1)$ -feltételt, mert ahhoz legalább 3 alappontjával kellene rendelkeznie (annak megfelelően, hogy a  $P_1(\square_i)$ -beli polinomoknak 3 szabad paramétere van).

További példaként tekintjük a 2-pontos tenzorszorzat Gauss-képletet (arra az esetre felírva, hogy a  $\square$  téglalap középpontja  $(x_{1i}, x_{2j})$ , oldalhossza  $h_1$ , ill.  $h_2$ ) :

$$\int_{\square} f(x_1, x_2) dx \approx \frac{1}{4} |\square| \sum_{\ell, m=1}^2 f(\xi_{\ell}, \eta_m), \quad (15.235)$$

$$\xi_{1,2} := x_{1i} \pm \frac{\sqrt{3}}{6} h_1, \quad \eta_{1,2} := x_{2j} \pm \frac{\sqrt{3}}{6} h_2.$$

Ez a képlet teljesíti a  $Q(1)$ -, de nem a  $Q(2)$ -feltételt.

Ezek a definíciók segítik a következő tétel rövidebb megfogalmazását, amelyet bizonyítás nélkül adjuk meg (ld. Ciarlet könyvében a 4.1.2. tételt).

**15.33. Tétel** (Ciarlet; közelítő bilineáris forma egyenletes ellipticitása). Legyen  $a$  a (15.197) bilineáris forma és ott  $\vec{v} = 0$ , teljesüljön (15.193), legyen  $a_h$  az  $a$ -ból kvadratúra képlet segítségével létrehozott közelítő bilineáris forma. Akkor  $a_h$  egyenletesen elliptikus, ha érvényesek a következő feltételek :

1. Minden  $\Omega_i$  tartomány affin leképezéssel leképezhető vagy  $d$ -dimenziós egységssimplexre, vagy  $d$ -dimenziós egységkockára.
2. A trianguláció egyenletesen reguláris.
3. A végeelem polinomtere háromszöges elem esetén  $P(\Omega_i) = P_k(\Omega_i)$ , négyszöges elemnél teljesüljön  $P_k(\Omega_i) \subset P(\Omega_i) \subset Q_k(\Omega_i)$ , és mindkét esetben legyen  $k \geq 1$ .
4. Az  $a_h$ -t létrehozó kvadratúra képlet teljesíti a  $P(k)$ - ill. a  $Q(k)$ -feltételt.
5. A képlet összes súlya legyen pozitív.  $\square$

A következő eredménnyel készítjük elő a választ a fenti 2. kérdésre, ezzel egyben a Céa-lemmát általánosítva.

**15.34. Lemma** (Strang; közelítő variációs feladat hibája). Legyen az  $a_h$  bilineáris forma egyenletesen elliptikus  $V_h \times V_h$ -n, legyen  $\varphi_h$  lineáris funkcionál  $V_h$ -n és  $u$ , ill.  $u_h$  a (15.189), ill. (15.233) variációs feladat megoldása.

Akkor érvényes a következő becslés :

$$\|u - u_h\|_V \leq C \left( \inf_{v_h \in V_h} \left[ \|u - v_h\|_V + \sup_{w_h \in V_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_V} \right] + \sup_{w_h \in V_h} \frac{|\varphi(w_h) - \varphi_h(w_h)|}{\|w_h\|_V} \right), \quad C := 1 + \frac{\max(M_a, 1)}{\alpha}.$$

**Bizonyítás.** Legyen tehát  $u_h \in V_h$  a (15.233) feladat, illetve  $u \in V$  a (15.189) megoldása. Tetszőleges  $v_h \in V_h$  függvénnyel  $u_h - v_h \in V_h \subset V$ , és ezért

$$a_h(u_h, u_h - v_h) = \varphi_h(u_h - v_h) \text{ és } a(u, u_h - v_h) = \varphi(u_h - v_h).$$

Így (15.234) alapján

$$\begin{aligned} \alpha \|u_h - v_h\|_V^2 &\leq a_h(u_h - v_h, u_h - v_h) = \varphi_h(u_h - v_h) - a_h(v_h, u_h - v_h) \\ &= \varphi_h(u_h - v_h) - a_h(v_h, u_h - v_h) \\ &\quad + [a(u - v_h, u_h - v_h) + a(v_h, u_h - v_h) - \varphi(u_h - v_h)] \\ &= a(u - v_h, u_h - v_h) + a(v_h, u_h - v_h) - a_h(v_h, u_h - v_h) + \varphi_h(u_h - v_h) - \varphi(u_h - v_h). \end{aligned}$$

Az  $a$  forma folytonosságának köszönhetően innen azt kapjuk, hogy  $u_h \neq v_h$  esetén

$$\alpha \|u_h - v_h\|_V \leq M_a \|u - v_h\|_V + \frac{|a(v_h, u_h - v_h) - a_h(v_h, u_h - v_h)|}{\|u_h - v_h\|_V} + \frac{|\varphi_h(u_h - v_h) - \varphi(u_h - v_h)|}{\|u_h - v_h\|_V},$$

ahonnan az  $\alpha$ -val való osztás után,  $\|u - u_h\|_V \leq \|u - v_h\|_V + \|u_h - v_h\|_V$  segítségével (és mivel  $v_h$  tetszőleges volt) a lemma állítása következik.  $\square$

**Megjegyzések.** **1.** Ahogyan a lemma bizonyításából is kiderül, a (15.234) feltétel éppen a közelítés stabilitását biztosítja.

**2.** Amennyiben most  $V = H^1$  vagy  $V = H_0^1$  és igaz

$$\begin{aligned} |a(v_h, w_h) - a_h(v_h, w_h)| &\leq c_1 h^k \|w_h\|_V \|v_h\|_V, \\ |\varphi(w_h) - \varphi_h(w_h)| &\leq c_2 h^k \|f\|_k \|w_h\|_V, \end{aligned}$$

minden  $v_h, w_h \in V_h$ -ra, akkor a lemmából,  $\|u - \Pi_h u\|_1 \leq c_3 h^k |u|_{k+1}$ -ből és  $\|\Pi_h u\|_1 \leq c_4 \|u\|_{k+1}$ -ből az következik, hogy

$$\|u - u_h\|_V \leq C h^k [(c_3 + c_1 c_4) \|u\|_{k+1} + c_2 \|f\|_k]. \quad \square$$

Ezzel a lemmával a 2. kérdés vizsgálatát függetlenítettük a peremérték feladattól ill. a variációs feladattól – kivéve azt a körülményt, hogy a variációs feladat Szoboljev–térbeli megoldása miatt a kvadratúra képleteknek az ilyen függvényeken való konvergenciáját kell vizsgálni. Ez megtehető a Bramble–Hilbert lemma segítségével, ld. II. 11.7.3., ahol a trapézszabály példáján a másodrendű konvergencia becslést bebizonyítottuk abban az esetben, amikor az  $f$  integrandus  $H^2$ -ből való. A következő eredmény bizonyítása mégis elég hosszú (emiatt bemutatásától eltekintünk), mert éppen nem az  $\int_{\Omega} f(x) dx$  integrál kvadratúra képletek segítségével való kiszámításának hibáját kell becsülni, hanem kettő ( $\varphi(w_h)$  közelítések), ill. három ( $a(v_h, w_h)$  közelítések) függvény szorzata szerepel az integrál alatt, és a becslést  $H^k$ -ban ill.  $H^{k+1}$ -ben kell elkészíteni meghatározott alakban. Az alábbi tételben ilyen vizsgálat végeredményét fogalmazzuk meg. A benne előforduló  $k > d/2$  feltétel a  $H^k \hookrightarrow C$  beágyazást biztosítja, amely a  $\varphi$ -integrálok becslése során fontos.

**15.35. Tétel** (Ciarlet; végeelem konvergencia kvadratúra képletek használatákor). Teljesüljenek a 15.33. tétel feltételei, a végeelemek legyenek Lagrange-típusúak és folytonosak  $\bar{\Omega} \subset \mathbb{R}^d$ -ben. Ha érvényes a  $P(k)$ -feltétel (ill.  $Q(k)$ -feltétel), akkor a kvadratúra képlet legyen egzakt  $P_{2k-2}(\Omega_i)$ -ben (ill.  $Q_{2k-1}(\Omega_i)$ -ben). Ezenkívül legyen  $k > d/2$ , továbbá teljesüljön  $k_{ij}, v_i, c \in C^k(\bar{\Omega})$  és  $f \in H^k(\Omega)$ . Az (15.189) variációs feladat megoldása legyen a  $H^{k+1}(\Omega)$ -térből való.

Ekkor érvényes a következő becslés :

$$\|u - u_h\|_1 \leq Ch^k (\|u\|_{k+1} + \|f\|_k).$$

Itt a  $C$  konstans nem függ  $h$ -től (hanem csak  $\Omega$ -tól,  $k$ -től, valamint a (15.192) egyenlet együtthatóitól).  $\square$

**Megjegyzés.** A  $2k - 2$ -fokszámú polinomokra vonatkozó feltétel azt jelenti a választott kvadratúra képletre nézve, hogy a (15.192) differenciálegyenlet főegyütthatóinak megfelelő  $(K \text{ grad } u_h) \cdot \text{grad } w_h$  integrandust a bilineáris forma (15.197) képletéből akkor kell egzaktul integrálnia, amikor a  $K$  mátrix  $k_{ij}$  elemei konstansok (hiszen a  $\text{grad } u_h$  és  $\text{grad } w_h$  vektorok komponensei  $k - 1$ -edfokú polinomok).  $\square$

A kvadratúra képletek gyakorlati kiválasztása a fenti eredmények ismeretében a következőképpen történik. A végeelem polinomtér a  $k$  számot határozza meg. A 15.35. tételből tudjuk, hogy a  $\varphi$ -integrálokat olyan képlettel kell közelíteni az  $\Omega_i$  végeelemen, amely  $P_{2k-2}(\Omega_i)$ -beli (ill.  $Q_{2k-1}(\Omega_i)$ -beli) polinomokat pontosan tud integrálni. Az  $a$  formával kapcsolatos integrálok

kiszámításánál ez lehet, hogy nem elég. Ekkor az egyenletes ellipticitása biztosítása érdekében még azt kell ellenőriznünk (ld. a 15.33. tétel 4. feltételét), vajon a kvadratúra képlet teljesíti a  $P(k)$ - ill. a  $Q(k)$ -feltételt, tehát alappontjai (mint szabadsági fokok) az összes  $\Omega_i$ -n definiált  $\ell$ -edfokú polinomot egyértelműen határozzák meg ( $\ell = k - 1$  ill.  $\ell = k \cdot d - 1$ ). Ez csak akkor lehetséges (mint szükséges feltétel), ha legalább annyi alappontja van a képletnek, mint az ilyen polinom szabad együtthatóinak száma:  $\binom{d+\ell}{\ell}$ , ha  $P_\ell(\Omega_i)$ -ről van szó és  $(\ell + 1)^d$  a  $Q_\ell(\Omega_i)$  esetén.

Egydimenziós feladatok esetén a  $k$ -pontú Gauss-képletek elegendőek, mert ekkor a képlet még a  $2k - 1$ -edfokú polinomokat is egzaktul integrálja, ld. I. 5.8., az összes súlya pozitív, és a  $k - 1$ -edfokú polinomokat biztosan meghatározza a  $k$  alappont (hiszen mind különbözőek).

A többdimenziós esetben is szóba jönnek (és elterjedtek a végeelem programokban) a Gauss-képletek, de azért gyakran helyettesíthetők hatékonyabb képletekkel, mert alappontjai mind a végeelem belsejében vannak, azaz költséges  $f(x)$ - és  $k_{ij}(x)$ -kiszámítás esetén az egyszer kiszámított függvényértéket nem lehet a szomszédos elemek kezelése során újra hasznosítani – mint amikor a képlet alappontjai a végeelem peremén vannak. Így pl. a  $T_3$  esetén használható, fent megemlített középpont szabály  $N$  háromszög esetén  $N$  függvénykiértékelést kíván. Ha ahelyett az I. 5.11-beli (5.40) képletet vesszünk (a trapézszabály általánosítását), akkor aszimptotikusan csak fele annyi kiértékelés szükséges, mert bár 3 érték kell ekkor egy elemhez, de átlagban 6 háromszög tartozik egy csúcshoz.

A fenti  $T_3, T_6, R_4, R_8$  példáknál a következő kvadratúra képletek az előző elmélet szerint megfelelőek (v.ö. a 48. feladattal, valamint az elemek 15.7.4-beli leírásával. A  $g$  integrandus lehet pl.  $fw_i$  vagy  $k_{\ell m} \frac{\partial w_j}{\partial x_m} \frac{\partial w_i}{\partial x_\ell}$  :

$$T_3 : \int_{\Delta_i} g(x) dx \approx \frac{|\Delta_i|}{3} \sum_{j=1}^3 g(x^{i,j}); \quad (15.236)$$

$$T_6 : \int_{\Delta_i} g(x) dx \approx \frac{|\Delta_i|}{3} \sum_{j=4}^6 g(x^{i,j}); \quad (15.237)$$

$$R_4 : \int_{\square_i} g(x) dx \approx \frac{|\square_i|}{4} \sum_{j=1}^4 g(x^{i,j}); \quad (15.238)$$

$$R_8 : \int_{\square_i} g(x) dx \approx \frac{|\square_i|}{36} \left( 16g(x^{i,9}) + \sum_{j=1}^4 g(x^{i,j}) + 4 \sum_{j=5}^8 g(x^{i,j}) \right). \quad (15.239)$$

Ezekben a képletekben  $|\Delta_i|$  és  $|\square_i|$  jelöli a háromszöges, ill. a téglalapos elem területét. Az  $R_8$  elemhez ajánlott képlet alappontjait az  $R_9$ -elem (ld. 41. feladat)  $x^{i,1}, \dots, x^{i,9}$  csomópontjai képezik. Az  $R_8$ -elemhez más kvadratúra képletet javasol a 49. feladat. Annak viszont negatív súlyai vannak, ami miatt azon képlet használatákor  $h$  kell, hogy elég kicsi legyen. (További, a numerikus integrálással kapcsolatos, fontos szempontokat 16.4.8-ban tárgyaljuk.)

Ezen pont elmélete ugyanúgy a háromdimenziós esetre is vonatkozik. Mint használható kvadratúra képletet a  $T_4^3$  elem esetén a középpont szabályt említjük meg (egyetlen alappontja a tetraéder súlypontja), a  $T_{10}^3$  esetén a tetraéder súlypontját ( $2/15$  súllyal) és négy csúcsát ( $1/120$  súllyal) tartalmazó, másodfokú polinomokat pontosan integráló képletet (ld. Sajdurov könyvét), és az  $R_8^3$ -elemhez a kocka mind 8 csúcsát egyenlően súlyozó képletet.

Az elmélettől megkövetelt  $2k - 2$  szám (mint a pontosan integrálandó polinomok fokszáma) még kevés is lehet: Ha az egyenlet

$$-\varepsilon \operatorname{div}(\operatorname{grad} u) + \vec{v} \cdot \operatorname{grad} u + cu = f$$

és  $\varepsilon$  kicsi  $|\vec{v}|$ -hez képest (tehát  $K = \varepsilon I$  (15.192)-ban, amely *konvekció-diffúzió egyenlet* egydimenziós esetével már II. 11.6.6-ban foglalkoztunk), akkor elméletileg elegendő és numerikus kísérletek alapján szükséges a következő.

A  $k$ -adik hibarend (egyenletesen  $\varepsilon$ -ban) akkor érhető el, ha a (három- vagy négyszöges) elem rendje  $k$ , ha a  $c, f$  függvényeket elemenként  $k - 1$ -edfokú polinomokkal interpoláljuk, a  $\vec{v}$  vektorfüggvényt pedig  $k$ -adfokú polinommal, majd az összes integrált pontosan számítjuk ki. Ez tehát  $2(k - 1)$ -edfokú polinomok pontos integrálását jelenti a (15.199) bilineáris  $a$  formának az  $\varepsilon \operatorname{grad} u_h \operatorname{grad} w_\ell$  része esetén,  $3k - 1$ -edfokúét a  $\vec{v}$ -vel ill.  $c$ -vel kapcsolatos integrálok esetén, ld. F. Schieweck dolgozatát.

### 15.7.7 Algoritmusok

Következőnek a végeelem módszer egyenletrendszerének az összeállítását mutatjuk meg. Az alapunk a (15.190)-ből levezetett  $Ay = b$  alakú (15.191) egyenletrendszer, ahol

$$A = (a_{k\ell}), \quad a_{k\ell} = a(w_\ell, w_k), \quad b_k = \varphi(w_k),$$

amelynek megoldása után, az  $y$  vektor birtokában

$$u_h = \sum_{\ell=1}^N y_\ell w_\ell$$

adja az  $u$  megoldás keresett  $u_h$  közelítését.

Legyen itt például, a (15.199) bilineáris formával és a  $\vec{v} = 0$  esetnek megfelelően,

$$\begin{aligned} a(u, w) &= \int_{\Omega} [(K \operatorname{grad} u) \cdot \operatorname{grad} w + cuw] dx, \\ \varphi(w) &= \int_{\Omega} fw dx. \end{aligned}$$

Legyen  $M$  az összes elemnek és  $N$  az összes szabadsági foknak a száma. Az elemeket és a szabadsági fokokat a trianguláció során megszámozzuk:  $1 \leq i \leq M$ ,  $1 \leq \ell \leq N$ . Ez a *globális megszámozás*. Az  $i$ -edik elemnek a szabadsági fokok  $n_i$  száma változhat  $i$ -vel, ha pl. a perem mentén a  $T_6$ -elemet használjuk, belül pedig az  $R_8$ -at. A szabadsági fokok elemen belüli számozása ( $1 \leq j \leq n_i$ ) adja a *lokális megszámozást*. A triangulációval egyidejűleg egy kétdimenziós egészértékű tömb formájában rögzítjük a rács összefüggéseit:

$$\operatorname{ind}(i, j) = \ell.$$

Tehát az **ind** tömb  $(i, j)$  eleme megadja az  $i$ -edik elem (a lokális számozás szerinti)  $j$ -edik szabadsági fokához a globális megszámozás szerinti  $\ell$  sorszámát.

Az elején az egész  $A$  mátrixot és a  $b$  vektort lenullázzuk. (A mechanikából származó feladatokban  $A$  a *merevségi mátrix* és  $b$  *terhelési vektor*.)

Ezután a következő algoritmus állítja elő a mátrixot és a jobboldali vektort:

1.  $i := 1(1)M$
2.  $[m := 1(1)n_i [k := \operatorname{ind}(i, m), b_k := b_k + \beta_m$
3.  $n := 1(1)n_i [\ell := \operatorname{ind}(i, n), a_{k\ell} = a_{k\ell} + \alpha_{mn}]_n]_m]_i$

Itt

$$\alpha_{mn} := \int_{\Omega_i} [K \operatorname{grad} p_n \cdot \operatorname{grad} p_m + cp_n p_m] dx, \quad \beta_m := \int_{\Omega_i} f p_m dx,$$

és  $p_m$  az  $\operatorname{ind}(i, m) = k$ -edik bázisfüggvény leszűkítése  $\Omega_i$ -re, azaz az elem  $m$ -edik formafüggvénye (amelyet korábban  $p_{i,m}$ -vel is jelöltük):

$$p_m(x) = \begin{cases} w_k(x), & \text{ha } \operatorname{ind}(i, m) = k \text{ és } x \in \Omega_i, \\ 0, & \text{egyébként.} \end{cases} \quad (15.240)$$

Hasonlóan  $p_n$  az  $\operatorname{ind}(i, n) = \ell$ -edik bázisfüggvény leszűkítése  $\Omega_i$ -re. Az  $\alpha_{mn}$  számok (rögzített  $i$ -re) alkotják a *lokális* (vagyis *elemi*)  $A_{loc}$  mátrixot

(míg  $A$  a *globális mátrix*). Az algoritmus természetesen azon alapszik, hogy az integrálok elemenként kiszámíthatóak :

$$a_{k\ell} = a(w_\ell, w_k) = \sum_{i=1}^M \int_{\Omega_i} [K \operatorname{grad} w_\ell \operatorname{grad} w_k + c w_\ell w_k] dx,$$

$$b_k = \varphi(w_k) = \sum_{i=1}^M \int_{\Omega_i} f w_k dx,$$

de kihagyja a sok nulla járulékot, amely keletkezik, ha  $w_\ell(x)w_k(x) \equiv 0$ ,  $x \in \Omega_i$ .

Ahogy látjuk, a végeelem program kell, hogy tartalmazza a  $p_j$  formafüggvényeket és deriváltjait (negyedrendű feladat esetén még a második deriváltakat is), de nem a  $w_\ell$  bázisfüggvényeket.

Ha nincs olyan elem, amelyben jelen volna az  $\operatorname{ind}(i, m) = k$ -adik és az  $\operatorname{ind}(i, n) = \ell$ -edik szabadsági fok, akkor  $a_{k\ell}$  érintetlenül marad nulla. Az

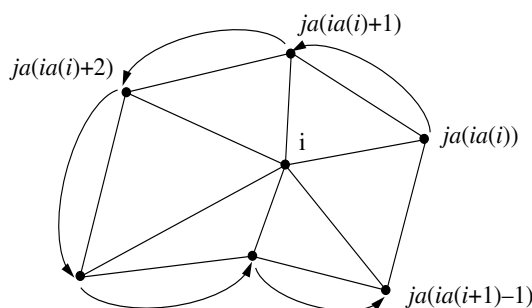
$$s := \max_{\substack{1 \leq i \leq M \\ 1 \leq m, n \leq n_i}} |\operatorname{ind}(i, m) - \operatorname{ind}(i, n)|$$

szám ezért adja az  $A$  mátrix (fél)sávszélességét. Az algoritmust lehet úgy megváltoztatni, hogy a sávon kívüli nulla elemeket nem tároljuk, vagy pedig úgy, hogy a kompakt tárolású mátrixot készíti – ami a célszerűbb. A kompakt tároláshoz (v.ö. I. 1.3.9-cel) szükséges  $\mathbf{ia}$  (sorelejét adó) és  $\mathbf{ja}$  (a globális sorszámozás szerinti oszlopszámot tartalmazó) vektorokat egy, a  $b_k$  és  $a_{k\ell}$  értékeit kiszámító, fenti ciklusra hasonló algoritmussal állítjuk elő az  $\operatorname{ind}$  tömbből (ld. az 50. feladatot).

De a rácyszerkesztés során az  $\mathbf{ia}$ ,  $\mathbf{ja}$  vektorokat közvetlenül is ki lehet számítani, az  $\operatorname{ind}$  tömb használata nélkül. Ekkor az  $A$  mátrixot és a  $b$  vektort kiszámító ciklus nem az elemeken fut végig ( $1 \leq i \leq M$ ), hanem az ismeretleneken (a szabadsági fokokon :  $1 \leq \ell \leq N$ ). Az  $\ell$ -edik szabadsági fok (csomópont) szomszédjai  $\mathbf{ja}$ -ban találhatóak az  $\mathbf{ia}(\ell)$ -edik komponenstől az  $\mathbf{ia}(\ell + 1) - 1$ -edikig. Tehát az  $\ell$ -edik csomópontnak  $m_\ell := \mathbf{ia}(\ell + 1) - \mathbf{ia}(\ell)$  darab szomszédja van, és ezeknek globális sorszámai  $j_1 := \mathbf{ja}(\mathbf{ia}(\ell))$ ,  $\mathbf{ja}(\mathbf{ia}(\ell) + 1), \dots, j_{m_\ell} := \mathbf{ja}(\mathbf{ia}(\ell + 1) - 1)$ . Ekkor, betartva egy körüljárási irányt, egyértelműen meghatározhatjuk az elem sorszámát. Pl. a  $T_3$ -elem esetén ez az  $(\ell, j_k, j_{k+1})$  csúcsok alapján megtehető,  $k = 1, \dots, m_\ell$  (ahol  $j_{m_\ell+1} \equiv j_1$ ).

A  $T_6$ -elemnél ugyanazt lehet elérni, ha a háromszögek csúcsait és oldalfelező pontjait megfelelő sorrendben végigjárjuk  $\mathbf{ja}$  elkészítése során, és  $R_4$ ,  $R_8$  esetén analóg módon dolgozhatunk.





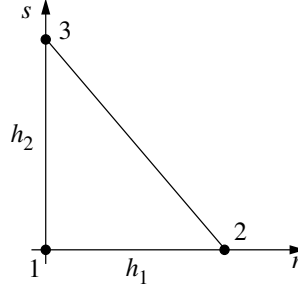
ábra 15.25: Kompakt tárolás elkészítése  $T_3$  esetén

Amennyiben a peremfeltételek elsőfajúak, akkor a teljes  $A$  mátrix és  $b$  jobboldali vektor összeállítása után a II. 11.6.8-ban részletezett műveletek következnek, tehát az elsőfajú feltételeknek megfelelő szabadsági fokok sorait a mátrixban (ezekre az indexekre használtuk II-ben az  $\mathcal{I}$  jelölést) felváltjuk az egységmátrix soraival, a  $b$  vektorba ott beírjuk a peremértéket. Ezután azonnal következhet a lineáris rendszert megoldó algoritmus, amennyiben ez nonszimmetrikus mátrixra is működik. Ha nem, akkor még a 4'. és 5. lépés következik II. 11.6.8-ból.

Már itt arra mutatunk rá, hogy a sima Gauss-elimináció alkalmazása a végeelem módszer által előállított lineáris rendszerekre műhiba, a sávos Gauss-elimináció még használatban van, de helyette javasolható a többrácsos módszer vagy a prekondicionált gradiens módszer. Ehhez a témához ld. a 15.8. és 15.4.10. pontokat.

A (15.203) harmadfajú peremfeltétel esetén mérvadó a (15.205)-(15.206) bilineáris forma és lineáris funkcionál. Ekkor a fenti algoritmusban szükséges egy ellenőrzés, vajon peremen vagyunk vagy sem – az  $\alpha_{jn}$  és  $\beta_j$  járulékok kiszámításakor (az összes index változatlan marad!). Ha igen, akkor még peremmenti integrálokat is kell közelítenünk.

Különösen egyszerű a  $\sigma = 0$ ,  $g = 0$  eset (fizikai interpretációja : szigetelt perem, ld. II. 11.2.), ekkor a peremet úgyszólván nem is vesszük észre : a peremen fekvő csomópontoknak megfelelő mátrixsorokhoz ugyanúgy számítjuk ki a járulékokat mint a belső csomópontokban, csak hogy a peremen kevesebb résztartomány találkozik a csomópontokban (ez viszont a többrácsos módszer keretében a finom és a durva rács közti, 15.8-ban megtárgyalt (15.258) összefüggést is befolyásolja).



ábra 15.26: Háromszög lokális  $(r, s)$  koordinátarendszerben

### 15.7.8 Modellfeladat végeelem megoldása

Ebben a pontban az eddigieket a legegyszerűbb modellfeladaton szemléltetjük (a pont végén egy általánosabb feladat alkalmából rátérünk a végeelem módszer és a boxmódszer közti kapcsolatra is).

Tekintsük a Poisson-egyenletet téglalapalakú tartományon homogén elsőfajú peremfeltételekkel, ld. (15.13) 15.3.1-ben!

Számítsuk ki először az elemi mátrixot a  $T_3$ -elemhez (ld. a 15.7.4-beli leírását)! Ehhez tekintsük a következő  $T$  háromszöget lokális  $(r, s)$  koordinátarendszerben : Itt 1, 2, 3 a szabadsági fokok lokális megszámozását adják, a hozzá tartozó formafüggvények (baricentrikus koordinátáknak megfelelően)

$$p_1 = 1 - \frac{r}{h_1} - \frac{s}{h_2} = \lambda_1(r, s), \quad p_2 = \frac{r}{h_1} = \lambda_2(r, s), \quad p_3 = \frac{s}{h_2} = \lambda_3(r, s).$$

A lokális mátrixot, ill. a lokális jobboldali vektort az

$$\alpha_{ij} := a_T(p_j, p_i), \quad \beta_i := \varphi_T(p_i),$$

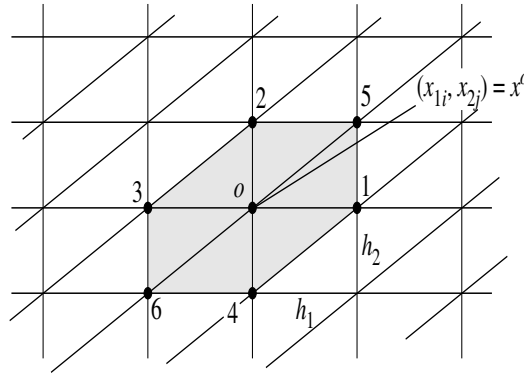
$$a_T(u, v) = \int_T \left[ \frac{\partial v}{\partial r} \frac{\partial u}{\partial r} + \frac{\partial v}{\partial s} \frac{\partial u}{\partial s} \right] dr ds, \quad \varphi_T(v) = \int_T f(r, s)v(r, s) dr ds$$

képletek határozzák meg. Innen, mivel a  $p_i, p_j$  deriváltjai konstansok, kapjuk pl., hogy

$$\alpha_{11} = \frac{h_1 h_2}{2} \left[ \left( -\frac{1}{h_1} \right)^2 + \left( -\frac{1}{h_2} \right)^2 \right],$$

$$\alpha_{12} = \frac{h_1 h_2}{2} \left[ \left( -\frac{1}{h_1} \right) \left( \frac{1}{h_1} \right) + \left( -\frac{1}{h_2} \right) \cdot 0 \right],$$

$$\alpha_{23} = \frac{h_1 h_2}{2} \left[ \frac{1}{h_1} \cdot 0 + 0 \cdot \frac{1}{h_2} \right],$$



ábra 15.27: Courant–elemek téglalapos alaprácson

és összesítve

$$A_{loc}(T) := (\alpha_{ij}) = \frac{h_1 h_2}{2} \begin{pmatrix} \frac{1}{h_1^2} + \frac{1}{h_2^2} & -\frac{1}{h_1^2} & -\frac{1}{h_2^2} \\ -\frac{1}{h_1^2} & \frac{1}{h_1^2} & 0 \\ -\frac{1}{h_2^2} & 0 & \frac{1}{h_2^2} \end{pmatrix}.$$

Ez a lokális mátrix szinguláris, ami nem véletlen : a mátrix a

$$\Delta u + f = 0, \quad (r, s) \in T; \quad \frac{\partial u}{\partial n} = 0, \quad (r, s) \in \partial T$$

feladattal áll kapcsolatban, amelynek megoldása csak egy tetszőleges konstans erejéig van meghatározva. A hozzátartozó variációs feladat ugyanis :

$$\int_T \left[ \frac{\partial v}{\partial r} \frac{\partial u}{\partial r} + \frac{\partial v}{\partial s} \frac{\partial u}{\partial s} \right] dr ds = \int_T f(r, s) v(r, s) dr ds \quad \text{minden } v \in H^1\text{-re,}$$

és a feladat következményeként kell, hogy teljesüljön

$$0 = \int_T f(r, s) dr ds,$$

egyébként nincs megoldás.

Hasznos észrevenni, hogy az egységes  $\frac{h_1 h_2}{2}$  szorzón kívül a mátrix elemeinek értékeit kölcsönhatásoknak tekinthetjük. Ez a kölcsönhatás a koordinátairányokban  $1/(\text{távolság})^2$  alakú. Az átlók irányában nincsen kölcsönhatás. Ennek alapján rögtön felírható a keletkező globális diszkretizáció, ha az egész  $\Omega$  téglapot a 15.27. ábra szerint Courant–elemekkel felosztjuk : Ekkor az előzőekből kiindulva az  $x^{ij} := (x_{1i}, x_{2j})$  pontban (amelyhez tartozik

az  $(u_h)_{ij} =: y_{ij}$  ismeretlen), azt kapjuk, hogy

$$(Ay)_{ij} = -h_1 h_2 \left( \frac{y_{i+1,j} - 2y_{ij} + y_{i-1,j}}{h_1^2} + \frac{y_{i,j+1} - 2y_{ij} + y_{i,j-1}}{h_2^2} \right). \quad (15.241)$$

Itt azért szerepel a  $\frac{h_1 h_2}{2}$  szorzó kétszerese, mert mindig két háromszögből jön egy-egy járulék két pont közötti kölcsönhatáshoz. Továbbá, itt a rácsnak megfelelő szemléletesebb kétindexű  $(Ay)_{ij}$  jelölést használtuk a mátrix leírására ahelyett, hogy

$$u_h = \sum_{\ell=1}^N y_\ell w_\ell, \quad (Ay)_k = a(u_h, w_k)_0 = \sum_{\ell=1}^N a(w_\ell, w_k) y_\ell, \quad k = 1, \dots, N,$$

alapon a globális (pl. sorfolytonos) sorszámozást alkalmaztuk volna, amely a valódi számításnál használandó.

Ahogy látjuk, a  $T_3$ - vagyis Courant–elem esetén annak globális  $A$  mátrixa pontosan megfelel a Laplace–operátorra alkalmazott véges differencia módszer ötpontos standard sémájának : utóbbinak  $A_h$  mátrixát csak még  $h_1 h_2$ -vel kell megszoroznunk,  $A = h_1 h_2 A_h$ . Innen azonnal következik, hogy hivatkozhatunk a differencia módszer tárgyalásakor 15.3.2-ben levezetett különféle becslésekre, pl. (15.35)-ből, ill. (15.45)-ből kapjuk

$$\|A^{-1}\|_{C(\omega_h)} \leq \frac{a^2 + b^2}{16h_1 h_2}, \quad \|A^{-1}\|_{(0,h)} \leq \frac{a^2 b^2}{8(a^2 + b^2)h_1 h_2}.$$

Továbbá, mivel  $A_h^{-1} = h_1 h_2 G_h$  adja az összefüggést a differencia-approximáció mátrixa és a  $G_h$  diszkrét Green-féle függvény között, most kiderül, hogy  $A^{-1} = G_h$ . Ezért használhatjuk a Green-féle függvényre levezetett becsléseket közvetlenül a végelem megoldás jellemzésére is, ld. 15.5.2. Pl. a 15.19. lemma szerint

$$0 \leq (A^{-1})_{ij} \leq \gamma_0 + \gamma_1 \ln(\sqrt{2} \min(h_1, h_2)).$$

Ezenkívül

$$\begin{aligned} \lambda_{\max}(A) &= h_1 h_2 \lambda_{\max}(A_h) < h_1 h_2 \left( \frac{4}{h_1^2} + \frac{4}{h_2^2} \right) = 4 \left( \frac{h_2}{h_1} + \frac{h_1}{h_2} \right) \\ &\leq 4 \left( \kappa + \frac{1}{\kappa} \right), \\ \lambda_{\min}(A) &\geq h_1 h_2 \left( \frac{8}{a^2} + \frac{8}{b^2} \right), \end{aligned} \quad (15.242)$$

ha érvényes  $\kappa^{-1} \leq h_2/h_1 \leq \kappa$ .

Numerikus példaként tűzzük ki a Poisson-egyenlet számára ugyanazt az  $u(x_1, x_2) = x_1^6 + x_2^6$  pontos megoldással rendelkező, elsőfajú peremérték feladatot, mint már 15.3.2-ben! A következő táblázat tartalmazza itt is  $\delta(h) := \|\vec{u}_h - y_h\|_{C(\omega_h)}$  értékeit különböző  $h$ -kra. A jobboldallal kapcsolatos integrálokat a (15.236) kvadratura képlet segítségével számítottuk ki.

$h$	1/2	1/4	1/8	1/16	1/32
$\delta(h)$	0.2969	0.1647	0.03279	$6.826 \cdot 10^{-3}$	$1.557 \cdot 10^{-3}$

Amennyiben az  $(f, w_i)_0$  integrálokat pontosan számítjuk ki, akkor esetünkben a végeelem megoldás a rácspontokban megegyezik a pontos megoldással (51. feladat).

Következőnek, a

$$-\Delta u = \lambda u, \quad (x_1, x_2) \in \Omega = (0, 1)^2; \quad u = 0, \quad (x_1, x_2) \in \Gamma = \partial\Omega,$$

sajátérték feladat miatt (v.ö. II. 11.6.4-gyel és 15.9.3-mal), vizsgáljuk még azt az esetet, hogy az  $f$  jobboldal helyett  $\lambda u$  áll. Ekkor

$$u_h = \sum_{\ell=1}^N y_\ell w_\ell, \quad (By)_k = (u_h, w_k)_0 = \sum_{\ell=1}^N (w_\ell, w_k)_0 y_\ell, \quad k = 1, \dots, N,$$

alapján képződik a Courant-elemekhez tartozó  $B$  úgynevezett *tömegmátrix*. Az  $A, B$  mátrixokkal a sajátérték feladat  $Ay = \lambda By$  alakú.

Kezdjük a számítást az  $x^0, x^1, x^5$  alappontú  $\Delta_{015}$  háromszögében! Itt  $u_h$  a  $w_0, w_1, w_5$  bázisfüggvények segítségével felírva az

$$u_h = y_0 w_0 + y_1 w_1 + y_5 w_5$$

lineáris függvény. Ezzel keletkezik egy első rész a tömegmátrixból :

$$\begin{aligned} (B^{(015)}y)_0 &:= \int_{\Delta_{015}} u_h(x) w_0(x) dx \\ &= h_1 h_2 \int_0^1 \int_0^\xi (1-\xi) [y_0(1-\xi) + y_1(\xi-\eta) + y_5\eta] d\eta d\xi \\ &= h_1 h_2 \int_0^1 \left[ y_0 \xi(1-\xi)^2 + y_1(1-\xi) \frac{\xi^2}{2} + y_5(1-\xi) \frac{\xi^2}{2} \right] d\eta d\xi \\ &= \frac{h_1 h_2}{24} \{2y_0 + y_1 + y_5\}. \end{aligned}$$

A  $\Delta_{052}$  háromszögben a képletek az előzőekből következnek az  $x_1 \leftrightarrow x_2$ ,  $h_1 \leftrightarrow h_2$  cserével :

$$(B^{(052)}y)_0 := \frac{h_1 h_2}{24} \{2y_0 + y_2 + y_5\}.$$

A  $\Delta_{023}$  háromszögben a képletek némileg eltérnek : ott  $u_h = y_0 w_0 + y_2 w_2 + y_3 w_3$ , és  $w_0(x) = 1 + \xi - \eta$ ,  $w_2(x) = \eta$ ,  $w_3(x) = -\xi$ , ahol  $\xi := x_1/h_1$ ,  $\eta := x_2/h_2$ . Ekkor

$$\begin{aligned} (B^{(023)}y)_0 &:= \int_{\Delta_{023}} u_h(x) w_0(x) dx \\ &= h_1 h_2 \int_{-1}^0 \int_0^{1+\xi} (1 + \xi - \eta) [y_0(1 + \xi - \eta) + y_2 \eta - y_3 \xi] d\eta d\xi \\ &= h_1 h_2 \int_{-1}^0 \left[ y_0 \frac{1}{3} (1 + \xi)^3 + y_2 \frac{1}{6} (1 + \xi)^3 - y_3 \frac{\xi}{2} (1 + \xi)^2 \right] d\eta d\xi \\ &= \frac{h_1 h_2}{24} \{2y_0 + y_2 + y_3\}. \end{aligned}$$

Ezután az összes többi háromszög járuléka már indexcserékkel kaphatók. Összesítve írhatunk

$$(By)_0 := \frac{h_1 h_2}{12} \left\{ 6y_0 + \sum_{k=1}^6 y_k \right\},$$

azaz

$$(By)_{ij} := \frac{h_1 h_2}{12} \{6y_{ij} + y_{i+1,j} + y_{i+1,j+1} + y_{i,j+1} + y_{i-1,j} + y_{i-1,j-1} + y_{i,j-1}\}. \quad (15.243)$$

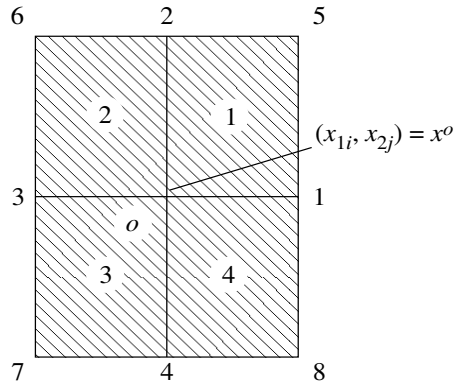
Ennek és (15.241)-nek a segítségével – figyelembe véve azt is, hogy  $y_{ij} = 0$ , amikor  $(x_{1i}, x_{2j}) \in \Gamma$  – állíthatjuk össze az  $Ay = \lambda By$  egyenleteket.

Második példaként most foglalkozzunk az  $R_4$ -elemmel. Formafüggvényei a  $\xi, \eta$ -sík egységnégyzetén

$$w_0(x) = (1 - \xi)(1 - \eta), \quad w_1(x) = \xi(1 - \eta), \quad w_2(x) = (1 - \xi)\eta, \quad w_5(x) = \xi\eta,$$

ahol  $\xi := x_1/h_1$ ,  $\eta := x_2/h_2$  adja az összefüggést az  $(x_1, x_2)$  sík elemi téglalapjával. Ennek megfelelően az ábra 1-es téglalapjában

$$u_h = y_0 w_0 + y_1 w_1 + y_2 w_2 + y_5 w_5,$$

ábra 15.28: Az  $R_8$  elem

és deriváltjaira azt kapjuk, hogy

$$\begin{aligned}\frac{\partial u_h}{\partial x_1} &= \frac{1}{h_1} \{-y_0(1-\eta) + y_1(1-\eta) - y_2\eta + y_5\eta\}, \\ \frac{\partial u_h}{\partial x_2} &= \frac{1}{h_2} \{-y_0(1-\xi) - y_1\xi + y_2(1-\xi) + y_5\xi\}.\end{aligned}$$

Ezen relációk birtokában kiszámíthatjuk az 1-es téglalap járulékát az  $A_{loc}$  lokális mátrixhoz :

$$\begin{aligned}(A_{loc}^{(1)}y)_0 &:= \int_{x_{1,0}}^{x_{1,0}+h_1} \int_{x_{2,0}}^{x_{2,0}+h_2} \left[ \frac{\partial w_0}{\partial x_1} \frac{\partial u_h}{\partial x_1} + \frac{\partial w_0}{\partial x_2} \frac{\partial u_h}{\partial x_2} \right] dx \\ &= h_1 h_2 \int_0^1 \int_0^1 \left[ -\frac{1}{h_1} (1-\eta) \frac{1}{h_1} \{(y_1 - y_0)(1-\eta) + (y_5 - y_2)\eta\} \right. \\ &\quad \left. - \frac{1}{h_2} (1-\xi) \frac{1}{h_2} \{(y_2 - y_0)(1-\xi) + (y_5 - y_1)\xi\} \right] d\xi d\eta \\ &= h_1 h_2 \left[ -\frac{1}{h_1^2} \left\{ \frac{1}{3}(y_1 - y_0) + \frac{1}{6}(y_5 - y_2) \right\} \right. \\ &\quad \left. - \frac{1}{h_2^2} \left\{ \frac{1}{3}(y_2 - y_0) + \frac{1}{6}(y_5 - y_1) \right\} \right].\end{aligned}$$

Indexcserével megkapjuk a többi,  $x^0$ -csúcsú téglalap járulékait :

$$\begin{aligned} (A_{\text{loc}}^{(2)}y)_0 &:= h_1 h_2 \left[ -\frac{1}{h_1^2} \left\{ \frac{1}{3}(y_3 - y_0) + \frac{1}{6}(y_6 - y_2) \right\} \right. \\ &\quad \left. - \frac{1}{h_2^2} \left\{ \frac{1}{3}(y_2 - y_0) + \frac{1}{6}(y_6 - y_3) \right\} \right], \\ (A_{\text{loc}}^{(3)}y)_0 &:= h_1 h_2 \left[ -\frac{1}{h_1^2} \left\{ \frac{1}{3}(y_3 - y_0) + \frac{1}{6}(y_7 - y_4) \right\} \right. \\ &\quad \left. - \frac{1}{h_2^2} \left\{ \frac{1}{3}(y_4 - y_0) + \frac{1}{6}(y_7 - y_3) \right\} \right], \\ (A_{\text{loc}}^{(4)}y)_0 &:= h_1 h_2 \left[ -\frac{1}{h_1^2} \left\{ \frac{1}{3}(y_1 - y_0) + \frac{1}{6}(y_8 - y_4) \right\} \right. \\ &\quad \left. - \frac{1}{h_2^2} \left\{ \frac{1}{3}(y_4 - y_0) + \frac{1}{6}(y_8 - y_1) \right\} \right]. \end{aligned}$$

A lokális járulékok összeadásával megkapjuk az  $A$  globális mátrix szorzatát az ismeretlenek  $y$  vektorával :

$$\begin{aligned} (Ay)_0 &= h_1 h_2 \left\{ -\frac{2}{3} \left[ \frac{y_1 - y_0}{h_1^2} + \frac{y_3 - y_0}{h_1^2} + \frac{y_2 - y_0}{h_2^2} + \frac{y_4 - y_0}{h_2^2} \right] \right. \\ &\quad \left. - \frac{1}{6} \left[ \frac{y_5 - y_2}{h_1^2} + \frac{y_6 - y_2}{h_1^2} + \frac{y_6 - y_3}{h_2^2} + \frac{y_7 - y_3}{h_2^2} \right. \right. \\ &\quad \left. \left. + \frac{y_7 - y_4}{h_1^2} + \frac{y_8 - y_4}{h_1^2} + \frac{y_8 - y_1}{h_2^2} + \frac{y_5 - y_1}{h_2^2} \right] \right\}. \end{aligned}$$

Differencia-jelölésekkel és az  $x^0 = x^{i,j} = (x_{1i}, x_{2j})$  helyen ez rövidebben felírható mint

$$(Ay)_{ij} = -h_1 h_2 \left\{ y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} + \frac{h_1^2 + h_2^2}{6} y_{\bar{x}_1 x_1 \bar{x}_2 x_2} \right\}_{ij}. \quad (15.244)$$

Ha ebben a képletben nem 6, hanem 12 lenne a nevező, akkor ez a Laplace-operátor negyedrendű differencia-approximációja lenne, ld. (15.36)-et 15.3.2-ben.

Következőnek számítsuk ki még az első téglalpra vonatkozó

$$(B_{\text{loc}}^{(1)}y)_0 := \int_{x_{1,0}}^{x_{1,0}+h_1} \int_{x_{2,0}}^{x_{2,0}+h_2} w_0(x) u_h(x) dx_1 dx_2 \quad (15.245)$$

integrált! Ez egyrészt a sajátérték feladat (ld. 15.9.1.) miatt érdekes, de a jobboldali  $\varphi(w_0) = (f, w_0)_0$  érték approximációja miatt is. Ez utóbbi esetben  $u_h$  helyett a

$$f_h(x) := f_0 w_0(x) + f_1 w_1(x) + f_2 w_2(x) + f_3 w_3(x) \approx f(x), \quad x \in \square_1 \quad (15.246)$$



képletet be kell helyettesítenünk (15.245)-ba (ehhez ld. a 52. feladot). Felhasználva  $u_h$  és  $w_0$  explicit alakját, azt találjuk, hogy

$$\begin{aligned} (B_{\text{loc}}^{(1)}y)_0 &= h_1 h_2 \int_0^1 \int_0^1 (1-\xi)(1-\eta) [y_0(1-\xi)(1-\eta) + y_1\xi(1-\eta) \\ &\quad + y_2(1-\xi)\eta + y_5\xi\eta] d\xi d\eta \\ &= \frac{h_1 h_2}{36} [4y_0 + 2y_1 + 2y_2 + y_5]. \end{aligned}$$

A többi résztéglalap járulékait indexcserével kapjuk meg úgy, hogy az  $y_0 = y_{ij}$  stb. relációkat is figyelembe véve azonnal a globális mátrixot fel tudjuk írni :

$$\begin{aligned} (By)_{ij} &:= \frac{h_1 h_2}{36} [16y_{ij} + 4(y_{i+1,j} + y_{i,j+1} + y_{i-1,j} + y_{i,j-1}) \\ &\quad + y_{i+1,j+1} + y_{i-1,j+1} + y_{i-1,j-1} + y_{i+1,j-1}]. \end{aligned} \quad (15.247)$$

A (15.244), (15.247) jelölésekkel tehát a diszkrét sajátérték feladat alakja

$$(Ay)_{ij} = \lambda(By)_{ij}, \quad x = x_{ij} \in \omega_h; \quad y(x) = 0, \quad x \in \gamma_h, \quad (15.248)$$

és a Poisson-egyenlet elsőfajú peremérték feladatának bilineáris végeselem diszkretizációja

$$(Ay)_{ij} = (B\vec{f})_{ij}, \quad x = x_{ij} \in \omega_h; \quad y(x) = 0, \quad x \in \gamma_h,$$

ahol  $\vec{f}$  az  $f$  függvény rácspontbeli értékeiből összeállított vektor.

A példa tárgyalását azzal fejezzük be, hogy ismételten a következőre mutatunk rá : a végeselem módszerre nem tipikus a különféle integrálok kézi kiszámítása, mint fent. Ezt a munkát megfelelő kvadratúra képletek segítségével a számítógép végzi, ld. 15.7.6. Mi itt egyrészt a 15.7.7. szerinti általános munkamenetet akartunk szemléltetni, másrészt az eredő egyenleteket a speciális esetekben konkrétan felírni – és mindkét esetben láttuk a kapcsolatot ismert differencia képletekkel.

### 15.7.9 Végeselem és véges térfogat módszer

Tekintsük a következő általánosabb elliptikus peremérték feladatot!

$$\begin{aligned} \operatorname{div}(k \operatorname{grad} u) + f &= 0, \quad (x_1, x_2) \in \Omega, \\ k = k(x_1, x_2) &\geq k_0 = \operatorname{const} > 0, \\ u &= 0, \quad (x_1, x_2) \in \Gamma, \end{aligned} \quad (15.249)$$

ahol  $\Omega$  korlátos tartomány.  $\Gamma$  legyen töröttvonal, úgy hogy  $\Omega$  konform háromszöges triangulációja elkészíthető. A háromszögeket  $\Delta_i$ -vel jelöljük ( $1 \leq i \leq M$ ), csúcsait  $x^\ell := (x_{1\ell}, x_{2\ell})$ -vel. Ezek közül főként az  $\Omega$  szempontjából belső pontok ( $1 \leq \ell \leq N$ ) érdekelnek minket, mivel (15.249) peremérték feladatunk elsőfajú.

Legyen  $V_h = \text{span}\{w_1, \dots, w_N\}$  a Courant–elemeknek megfelelő véges-elem tér, ahol  $w_\ell = w_\ell(x) = w_\ell(x_1, x_2)$ . A (15.249) feladat közelítő variációs megfogalmazása :

Keressünk olyan

$$u_h := \sum_{\ell=1}^N u_{h,\ell} w_\ell \in V_h \quad (15.250)$$

függvényt, hogy

$$\iint_{\Omega} (k \text{ grad } u_h) \cdot \text{grad } w_j \, dx = \iint_{\Omega} f w_j \, dx, \quad \text{minden } 1 \leq j \leq N\text{-re.} \quad (15.251)$$

(15.250) behelyettesítése (15.251)-ba az

$$A \vec{u}_h = b$$

egyenletrendszert adja, ahol

$$A = (a_{i\ell}), \quad a_{i\ell} := \iint_{\Omega} k \text{ grad } w_\ell \cdot \text{grad } w_i \, dx,$$

$$b_i := \iint_{\Omega} f w_i \, dx, \quad \vec{u}_h := (u_{h,1}, \dots, u_{h,N})^T,$$

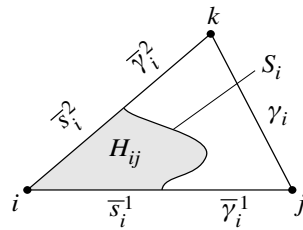
vagyis  $\vec{u}_h$  az  $u_h$  függvény projekciója  $\mathbb{R}^N$ -be.

Alkalmazzuk (15.249)-ra a boxmódszert is! Ehhez az  $x^i$  csomóponthoz a 30. ábra szerinti  $H_{ij}$  részt rendeljük hozzá. Legyen

$$H_i = H(x^i) := \bigcup_{j=1}^{m_i} H_{ij},$$

ahol  $m_i$  azon háromszögek száma, amelyeknek  $x^i$  csúcsa. Ezeket a háromszögeket  $\Delta_{ij}$ -vel is jelöljük.  $H_i$  pereme legyen  $S_i$ . Csak annyit teszünk fel, hogy  $H_{ij} \subset \Delta_{ij}$ , és hogy  $H_{ij}$  pereme a következő 3 részből áll (ha  $x^j$  és  $x^\ell$  a  $\Delta_{ij}$  másik két csúcsa) :

$$\begin{aligned} \bar{s}_i^1 : \quad a \quad \bar{\gamma}_i^1 &:= \overline{x^i x^j} \text{ szakasz } x^i\text{-hez kapcsolódó fele,} \\ \bar{s}_i^2 : \quad a \quad \bar{\gamma}_i^2 &:= \overline{x^i x^\ell} \text{ szakasz } x^i\text{-hez kapcsolódó fele,} \\ s_i &:= \partial H_{ij} / (\bar{s}_i^1 \cup \bar{s}_i^2) = \partial H_{ij} \cap \Delta_{ij}. \end{aligned}$$



ábra 15.29: Elsőrendű boxmódszer

Ennek pl. a boxmódszer 15.6-ban részletezett háromszöges változata tesz eleget (ld. ott a 18. ábrát, ahol  $H_{0,1}$  a háromszög azon része, amely az elemi térfogathoz hozzátartozik). Ezekkel a jelölésekkel a

$$0 = \int_{S_i} k(\text{grad } u) \cdot n \, ds + \int_{H_i} f \, dx \quad (15.252)$$

mérlegegyenletet írjuk fel (ld. 15.6.). Az egyenletet azáltal diszkrétizáljuk, hogy  $u$  helyett az (15.250) mintájára definiált

$$v_h = \sum_{\ell=1}^N v_{h,\ell} w_\ell$$

függvényt (15.252)-be helyettesítjük be. Ekkor (15.252)-ből keletkezik a

$$B\vec{v}_h = c$$

egyenletrendszer, ahol

$$B = (b_{i\ell}), \quad b_{i\ell} := - \int_{S_i} k \frac{\partial w_\ell}{\partial n} \, ds, \quad c_i := \int_{H_i} f \, dx.$$

Vegyük észre, hogy a Courant–elemek segítségével definiált  $v_h$  függvény behelyettesítése (15.252)-be olyan diszkrétizációra vezet, amely a 15.6.3-ban vizsgáltat tartalmazza, mint speciális esetet! Ha ugyanis  $v_h$  háromszögenként lineáris, akkor a 15.6.3. pontban, a (15.173) egyenlet alatt használt mindkét grad  $u$ -közelítés egyenlőségbe megy át :  $v_h$  gradiense konstans vektor, és a lineáris  $v_h$  függvényre alkalmazott trapézszabály pontos eredményt ad. A specializálás abból adódik (v.ö. a 30. és 18. ábrát), hogy milyen  $S_i$  vonal mentén kell a  $k(\text{grad } u) \cdot n$  integrálját kiszámítani. A következő tétel olyan feltételt fogalmaz meg, hogy ez mikor nem hat ki a mátrixra.

**15.36. Tétel** (Hackbusch; a boxmódszer és a Courant–elemes végeelem módszer eltérése).

a) Amennyiben a boxmódszerben a (15.249) feladat  $k$  együtthatóját háromszögenként  $\bar{k}$  átlagértékével helyettesítjük, akkor megegyezik a két módszer mátrixa :  $A = B$ ;

b) ha  $f \in L_2(\Omega)$ , akkor a módszerek két jobboldali vektorának,  $c$ -nek és  $b$ -nek a különbsége elsőrendű  $h$ -ban és így a két megoldás különbsége elsőrendű az  $\|\cdot\|_1$ -normában :

$$\|u_h - v_h\|_1 \leq \frac{h}{k_0} \|f\|_0.$$

**Bizonyítás.** Az alábbiakban csak az a) részt bizonyítjuk be, a b) rész miatt ld. Hackbusch 1989-es cikkét. Mivel

$$a_{i\ell} = \iint_{\Delta_i} k \operatorname{grad} w_\ell \cdot \operatorname{grad} w_i \, dx = \sum_{j=1}^{m_i} \iint_{\Delta_{ij}} k \operatorname{grad} w_\ell \cdot \operatorname{grad} w_i \, dx,$$

csak az  $x^i$  csúcsú tetszőleges  $\Delta_{ij}$  háromszöggel és  $H_{ij}$  részével foglalkozunk.  $\Delta_{ij}$  felett  $\operatorname{grad} w_\ell$  és  $\operatorname{grad} w_i$  konstans vektorok, ezért

$$\begin{aligned} \iint_{\Delta_{ij}} k \operatorname{grad} w_\ell \cdot \operatorname{grad} w_i \, dx &= \bar{k} |\Delta_{ij}| \operatorname{grad} w_\ell \cdot \operatorname{grad} w_i \\ &= \bar{k} \iint_{\Delta_{ij}} \operatorname{grad} w_\ell \cdot \operatorname{grad} w_i \, dx, \end{aligned}$$

ahol  $|\Delta_{ij}|$  a háromszög területe és

$$\bar{k} := \frac{1}{|\Delta_{ij}|} \iint_{\Delta_{ij}} k \, dx$$

a hővezetési együttható átlagértéke. Parciális integrálással kapjuk

$$\iint_{\Delta_{ij}} \operatorname{grad} w_\ell \cdot \operatorname{grad} w_i \, dx = \int_{\partial\Delta_{ij}} w_i \frac{\partial w_\ell}{\partial n} \, ds - \iint_{\Delta_{ij}} w_i \Delta w_\ell \, dx,$$

de  $\Delta w_\ell = 0$ . A visszamaradó vonalintegrálban vegyük figyelembe, hogy az  $x^i$  ponttal szembelevő  $\gamma_i$  oldalon érvényes  $w_i = 0$ , míg a másik két oldalon ( $\bar{\gamma}_i^1 \cup \bar{\gamma}_i^2 =: \bar{\gamma}_i$ ) a  $w_i$  függvény  $x^i$ -beli 1 értékéről lecsökken nullára. Emiatt, és mivel  $\bar{\gamma}_i^1$ -n és  $\bar{\gamma}_i^2$ -n teljesül  $\frac{\partial w_\ell}{\partial n} = \text{const}$ , így

$$\int_{\partial\Delta_{ij}} w_i \frac{\partial w_\ell}{\partial n} \, ds = \frac{1}{2} \int_{\bar{\gamma}_i} \frac{\partial w_\ell}{\partial n} \, ds.$$

A  $H_{ij}$  résztartomány  $\bar{s}_i := \bar{s}_i^1 \cup \bar{s}_i^2 = \partial H_{ij} \cap \bar{\gamma}_i$  peremének hossza  $\bar{\gamma}_i$ -ének a fele. Ezért

$$\frac{1}{2} \int_{\bar{\gamma}_i} \frac{\partial w_\ell}{\partial n} ds = \int_{\bar{s}_i} \frac{\partial w_\ell}{\partial n} ds.$$

Egyébként annak alapján, hogy

$$\iint_{H_{ij}} \Delta w_\ell dx = \int_{\partial H_{ij}} \frac{\partial w_\ell}{\partial n} ds,$$

és  $H_{ij}$ -ben  $\Delta w_\ell = 0$ , érvényes

$$\iint_{D_i} k \operatorname{grad} w_\ell \cdot \operatorname{grad} w_\ell dx = - \int_{s_i} \bar{k} \frac{\partial w_\ell}{\partial n} ds.$$

(Emlékeztetünk arra, hogy  $s_i = \partial H_{ij} \cap \Delta_{ij} = \partial H_{ij} / \bar{s}_i$ .) Innen, az  $x^i$  csúcsú háromszögek feletti összegzéssel,

$$a_{i\ell} = - \sum_{j=1}^{m_i} \int_{s_i} \bar{k} \frac{\partial w_\ell}{\partial n} ds,$$

azaz valóban  $A = B$ .  $\square$

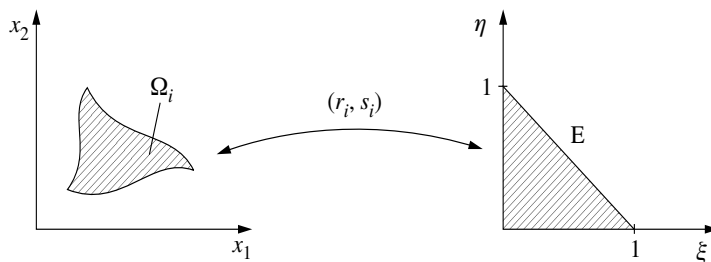
A tétel következménye, hogy a (15.249) feladat esetén a boxmódszer a Courant-módszerrel együtt stabil és elsőrendben konvergens, hiszen mátrixai megegyeznek és 15.7.6. szerint a  $k$  együttható helyettesítése résztartományonként konstans függvénnyel nem csökkenti a Courant-elemes megoldás konvergenciarendjét. A mátrix képlete ilyenkor (15.177) 15.6.3-ban.

Ha (15.249) megoldása elég sima ( $u \in H^2(\Omega)$ ,  $f \in H^1(\Omega)$ ) és ha minden  $x^i$ -ra a  $H(x^i)$  elemi térfogat minden  $H_{ij}$  részére igaz  $|H_{ij}| = \frac{1}{3} |\Delta_{ij}|$ , akkor a boxmódszer megoldása csak  $O(h^2)$ -tel tér el a végeelem módszer megoldásától.

### 15.7.10 Az izoparametrikus módszer \*\*\* 15.7.9.

A felsorolt példák még nem mutatják a végeelem módszer teljes erejét, hiszen úgy tűnik, hogy itt csak téglalapokról, háromszögekről van szó. Viszont ezeket a standard résztartományokat a  $\xi, \eta$ -síkba helyezhetjük, ahonnan egy sima, invertálható

$$\begin{aligned} r_i(\xi, \eta) &= x_1, \\ s_i(\xi, \eta) &= x_2, \end{aligned}$$



ábra 15.30: Izoparametrikus leképezés

leképezéssel az  $x_1, x_2$ -síkbeli  $\Omega_i$  görbevonalú résztartományba ( $1 \leq i \leq M$ ) vihetjük át. Ilyen leképezést az izoparametrikus módszer biztosít konstrukatív, lejjebb tárgyalt módon. Ezzel egyben azt a problémát is oldjuk meg, hogy az  $x_1, x_2$ -sík görbevonalú három- vagy négyszögeihez tartozó, és más módon nehezen megszerezhető formafüggvényeket is kezelhető formában kapjuk a  $\xi, \eta$ -síkbeli standard résztartományok formafüggvényeiből. Valójában az  $x_1, x_2$ -síkbeli formafüggvényekre nincs is szükség! Az összes művelet (integrálás, a lineáris egyenletrendszerből kapott megoldási értékek interpolációja az egész  $\Omega$ -ra) a  $\xi, \eta$ -síkbán végezhető el.

Legyen  $E$  a standard tartomány a  $\xi, \eta$ -síkbán. Akkor, ha pl. kiszámítandó az  $f(x_1, x_2)$  integrálja  $\Omega$  felett, a következőképpen járhatunk el :

$$\begin{aligned} \int_{\Omega} f(x_1, x_2) dx &= \sum_{i=1}^M \int_{\Omega_i} f(x_1, x_2) dx \\ &= \sum_{i=1}^M \int_E f(r_i(\xi, \eta), s_i(\xi, \eta)) D_i(\xi, \eta) d\xi d\eta, \end{aligned}$$

ahol

$$D_i := |\det J_i|, \quad \text{és } J_i := \frac{\partial(x_1, x_2)}{\partial(\xi, \eta)} = \begin{pmatrix} \frac{\partial r_i}{\partial \xi} & \frac{\partial r_i}{\partial \eta} \\ \frac{\partial s_i}{\partial \xi} & \frac{\partial s_i}{\partial \eta} \end{pmatrix}$$

az  $\Omega_i$ -hez tartozó Jacobi-mátrix. Ezek a szokásos matematikai összefüggések. A végeelem módszer viszont még az  $(r_i, s_i)$  leképezésre is kínál konstruktív eljárást. Ezt részletezzük a következőkben, ahol az áttekinthetőség érdekében az  $\Omega_i$  sorszámára utaló  $i$  indexet elhagyjuk.

A  $\xi, \eta$ -síkbeli  $E$  tartomány csomópontjai legyenek  $(\xi_j, \eta_j)$ , a formafüggvényei legyenek  $p_j(\xi, \eta)$ ,  $1 \leq j \leq n$ . Ekkor nemcsak a keresett megoldást  $\Omega_i$ -ben számíthatjuk ki

$$u_h(x_1, x_2) = u_h(r(\xi, \eta), s(\xi, \eta)) = \sum_1^n u_{h,j} p_j(\xi, \eta)$$

alapján, hanem kézenfekvő, ezt minden más kívánt függvényre is megtenni, pl.

$$x_1 = r(\xi, \eta) = \sum_1^n x_1(\xi_j, \eta_j) p_j(\xi, \eta), \quad x_2 = s(\xi, \eta) = \sum_1^n x_2(\xi_j, \eta_j) p_j(\xi, \eta). \quad (15.253)$$

Innen az „izoparametrikus” elnevezés : a megoldást és a koordinátákat ugyanolyan parametrikus formában keressük. (15.253) egyértelmű megadásához a  $(\xi_j, \eta_j)$  csomópontjainak  $(x_{1j}, x_{2j}) = (x_1(\xi_j, \eta_j), x_2(\xi_j, \eta_j))$  képeit az  $x_1, x_2$ -síkon,  $\Omega_i$ -ben elegendő megadni.

Most legyen

$$W(\xi, \eta) := \frac{\partial(p_1, \dots, p_n)}{\partial(\xi, \eta)},$$

akkor a Jacobi-mátrix

$$J = XW, \quad X := \begin{pmatrix} x_1(\xi_1, \eta_1), \dots, x_1(\xi_n, \eta_n) \\ x_2(\xi_1, \eta_1), \dots, x_2(\xi_n, \eta_n) \end{pmatrix}.$$

Innen a  $D_i(\xi, \eta) = |\det J_i|$  könnyen kiszámítható, hiszen a determináns mérete  $2 \times 2$ .

Szükségünk van még (másodrendű feladatnál, ld. pl. a (15.192) egyenlethez tartozó (15.199) bilineáris formát) tetszőleges  $P(\Omega_i)$ -beli  $q(x_1, x_2)$  polinom első deriváltjaira is. Ezeket mint (15.225)-ben kapjuk meg :

$$\nabla_{\xi, \eta} q = J^T \nabla_{x_1, x_2} q,$$

ahol a polinom  $q = \sum_1^n q(\xi_j, \eta_j) p_j(\xi, \eta)$  képlete alapján :

$$\nabla_{\xi, \eta} q := \left( \frac{\partial q}{\partial \xi}, \frac{\partial q}{\partial \eta} \right)^T = W^T \vec{q}, \quad \vec{q} := (q(\xi_1, \eta_1), \dots, q(\xi_n, \eta_n))^T.$$

Így végül is

$$\nabla_{x_1, x_2} q = J^{-T} \nabla_{\xi, \eta} q = (W J^{-1})^T \vec{q}.$$

Az itt felvázolt izoparametrikus eljárás különösen érdekes a  $T_6$ - és  $R_8$ -elem esetén : ha a perem mentén alkalmazzuk, akkor mód nyílik a  $\Gamma$  peremnek parabola darabokkal való közelítésére.

Konstruáljuk pl. a  $T_6$ -elemhez azt az „izoparametrikus” leképezést, amely a  $\xi, \eta$ -síkbeli egységssimplexet ( $0 \leq \xi, \eta \leq 1, 0 \leq \xi + \eta \leq 1$ ) az  $x_1, x_2$ -síkbeli  $(0, 0), (h_1, 0), (0, h_2)$  csúcsú háromszögbe viszi át, ahol a  $(0, 0) - (h_1, 0)$  és  $(0, 0) - (0, h_2)$  oldalak egyenesek, a  $(h_1, 0) - (0, h_2)$  oldal viszont görbe lehet : az  $(\frac{1}{2}, \frac{1}{2})$  pontnak az  $x_1, x_2$ -síkon feleljen meg az  $\frac{1}{2}(\mu h_1, \nu h_2)$  pont. (Ld. az 53. feladot is.)

Pontosabban, ha a  $(\xi, \eta) \rightarrow (x_1, x_2)$  leképezés jellemezhető a következő megfeleltetésekkel :

$$\begin{aligned} (0,0) &\rightarrow (0,0), & (1,0) &\rightarrow (h_1,0), & (0,1) &\rightarrow (0,h_2), \\ (\frac{1}{2},0) &\rightarrow \frac{1}{2}(h_1,0), & (\frac{1}{2},\frac{1}{2}) &\rightarrow \frac{1}{2}(\mu h_1, \nu h_2), & (0,\frac{1}{2}) &\rightarrow \frac{1}{2}(0, h_2), \end{aligned}$$

akkor (15.253) szerint

$$x_1 = h_1 \left[ p_2 + \frac{1}{2}p_4 + \frac{\mu}{2}p_5 \right], \quad x_2 = h_2 \left[ p_3 + \frac{\nu}{2}p_5 + \frac{1}{2}p_6 \right],$$

azaz (ld.  $T_6$  definícióját)

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} h_1[\lambda_2(2\lambda_2 - 1) + 2\lambda_1\lambda_2 + 2\mu\lambda_2\lambda_3] \\ h_2[\lambda_3(2\lambda_3 - 1) + 2\nu\lambda_2\lambda_3 + 2\lambda_3\lambda_1] \end{pmatrix}. \quad (15.254)$$

Az egységshimplexen érvényes

$$\lambda_1 = 1 - \xi - \eta, \quad \lambda_2 = \xi, \quad \lambda_3 = \eta.$$

Mivel a  $\Gamma$  perem a  $\lambda_1 = 0$  vonalnak felel meg, ahol  $\eta = 1 - \xi$ , a peremnek a (15.254) által létrehozott parametrikus közelítése

$$x_1 = h_1\xi[2\xi - 1 + 2\mu(1 - \xi)], \quad x_2 = h_2(1 - \xi)[1 - 2\xi + 2\nu\xi]. \quad (15.255)$$

A  $T_6$ -elem bázisa tulajdonságának megfelelően ez a leképezés már általt meghatározott, hogy

$$(1,0) \rightarrow (h,0), \quad (\frac{1}{2},\frac{1}{2}) \rightarrow \frac{1}{2}(\mu h_1, \nu h_2), \quad (0,1) \rightarrow (0, h_2),$$

és abban a speciális esetben, amikor  $\mu = 1$ , a (15.255) parametrikus alakból adódik

$$x_2 = h_2(1 - \xi)(1 + 2(\nu - 1)\xi), \quad 0 \leq \xi = \frac{x_1}{h_1} \leq 1,$$

ami egy parabola darab.

Az izoparametrikus módszer pontosságával közvetve már 15.7.5-ben foglalkoztunk, hiszen a 15.31. tétel bizonyítása támaszkodik a  $\xi, \eta$ -síkon elhelyezkedő standard tartományra való affin leképezésre. Ha a leképezések nem affinak, akkor a bizonyítás utáni 2. megjegyzés mérvadó. Egy további, az izoparametrikus leképezéssel felmerülő kérdés az, vajon milyen jól közelíthető az általános tartomány pereme, és ennek milyen kihatása van a pontos megoldásra és az izoparametrikus megoldásra. Ilyen vizsgálat eredménye az, hogy  $O(h^{k+1})$  nagyságrendű mind a pontos tartomány és a közelítő tartomány peremének eltérése, mind ennek kihatása a pontos megoldásra (ez utóbbi



témához ld. Blair cikkét), mind az izoparametrikus közelítő megoldásra (ehhez ld. Strang és Fix könyvét). Így a Courant–elem adta töröttvonalas peremközelítés másodrendű hibát okoz a végeelem megoldásban, és a fent szemléltetett parabolikus peremközelítések harmadrendű hibát.

Az izoparametrikus módszer eredményeit az  $x_1, x_2$ -síkon megmutatni viszont valamivel körülményesebb, mert az összes adat csak a  $\xi, \eta$ -sík standard tartományán könnyen hozzáférhető. Ha pl. egy adott  $(x_1, x_2)$  pontban akarjuk adni a megoldás értékét, akkor a (15.253) leképezést invertálnunk kell: az adott  $(x_1, x_2)$  koordinátákhoz keressük azt a  $(\xi, \eta)$  párt, amellyel (15.253) teljesül, mert csak annak birtokában ismerjük a hozzátartozó megoldási értéket.

Ehhez a Newton–módszert lehet használni. Alkalmazását megkönnyíti, hogy egyrészt csak  $2 \times 2$ -es méretű a nemlineáris egyenletrendszer, másrészt a deriváltakat analitikusan ki tudjuk számítani, harmadrészt, mivel általában sorozatosan kell ilyen megoldási értéket kiszámítani közeli pontokban, mindig jó kezdeti közelítés áll rendelkezésre – kivéve a legelején.

Ha speciálisan izovonalakat ki akarunk rajzolni, akkor ezt a közönséges differenciálegyenleteknél szokásos módszereknek a mintájára tehetjük meg. Az egyenlet ekkor  $u_h(\xi, \eta) = \text{const}$ , vagyis: a

$$\frac{\partial u_h}{\partial \xi} + \frac{\partial u_h}{\partial \eta} \frac{d\eta}{d\xi} = 0$$

differenciálegyenlet megoldando  $\eta = \eta(\xi)$ -re ismert  $u_h(\xi, \eta)$  és  $\eta_0 = \eta(\xi_0)$  alapján, ahol feltesszük, hogy a  $(\xi_0, \eta_0)$  pont rajta van az  $u_h(\xi, \eta) = \text{const}$  izovonalon.

Ehhez a ponthoz ld. az 54–58. feladatokat is.

### 15.7.11 Rácyszerkesztés \*\*\* 15.7.10.

Eltételezve a legegyszerűbb esetektől, a tartomány felosztása (triangulációja, a rácyszerkesztés) egy igen időigényes és logikailag egyáltalán nem egyszerű művelet. A legfontosabb szabály az, hogy itt szorosan együtt kell működni a számítógépes grafikai programokkal, amelyekkel egyrészt a megoldási tartományt létrehozzuk, másrészt a kész számítási eredményeket szándékozzuk bemutatni.

A rác szerkesztéséhez abból a ponteloszlásból induljunk ki, amellyel a számítógépes grafikai program ábrázolja a tartomány peremét. A képernyő előtt, egérrel feloszthatjuk a tartományt néhány  $\square_i$  négyszögre, amelyek a perem mentén görbevonalúak, máskülönben egyenesoldalúak. Minden  $\square_i$  négyszög  $\Gamma_i$  peremén rácspontokat jelölünk ki megfelelő számban úgy, hogy

minden négyszög szemben lévő oldalain ugyanannyi rácspont legyen. Legyen ez  $m_i \geq 3$  ill.  $n_i \geq 3$  pont (a sarokpontokat beleszámítva).

Ezután igyekezünk  $\square_i$ -n olyan koordinátarendszert bevezetni, amelynek vonalai a szemben lévő pontokat összeköti. Ezt konstruktív módon a következő eljárással tehetjük. A  $\xi, \eta$ -síkbeli  $E$  egységnégyzeten, téglalapos ekvidisztáns  $\omega_h$  rácson  $((m_i-2) \times (n_i-2)$  belső ponttal), a diszkrét Laplace-egyenletet közelítőleg oldjuk meg, ehhez mint elsőfajú peremfeltételt a  $\Gamma_i$ -n kijelölt rácspontok  $x_1 = x_1(\xi, \eta)$ -koordinátáit megadva :

$$u_{\bar{\xi}\xi} + u_{\bar{\eta}\eta} = 0, \quad (\xi, \eta) \in \omega_h \subset E; \quad u(\xi, \eta) = x_1(\xi, \eta), \quad (\xi, \eta) \in \gamma_h \in \partial E. \quad (15.256)$$

Itt  $\gamma_h$  az  $E$  diszkrét pereme, a  $\Gamma_i$ -n megadott rácspontok számának megfelelően.

Ezután a  $\square_i$ -beli  $x^{k\ell}$  rácspontok  $x_1(\xi_k, \eta_\ell)$ -koordinátáit az  $u$  megoldás értékeiből kapjuk meg :

$$(x_1)_{k\ell} := u(\xi_k, \eta_\ell), \quad 1 \leq k \leq m_i, \quad 1 \leq \ell \leq n_i.$$

A fenti eljárás végülis azt jelenti, hogy  $\square_i$  belső rácspontjainak  $x_1$ -koordinátáit megkapjuk a  $\Gamma_i$ -n kijelölt rácspontok  $x_1$ -koordinátáiból.

Hasonlóan a rácspontok  $x_2$ -koordinátáit egy második diszkrét Laplace-egyenlet megoldásával számítjuk ki, ehhez mint peremfeltételt a perempontok  $x_2$ -koordinátáit megadva :

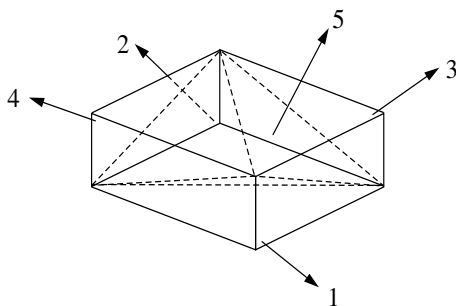
$$v_{\bar{\xi}\xi} + v_{\bar{\eta}\eta} = 0, \quad (\xi, \eta) \in \omega_h \subset E; \quad v(\xi, \eta) = x_2(\xi, \eta), \quad (\xi, \eta) \in \gamma_h \in \partial E, \quad (15.257)$$

és ezután

$$(x_2)_{k\ell} := v(\xi_k, \eta_\ell), \quad 1 \leq k \leq m_i, \quad 1 \leq \ell \leq n_i.$$

Összekötve a kapott rácspontokat, keletkezik az  $\square_i$  felosztása négyszögekre. A létrehozott négyszöges alaprács általában nem lesz ortogonális (bár a Laplace-egyenlet megoldásai és a konform leképezés – ez szögtartó – között ismert összefüggés van), de nagyon alkalmas.

Ehelyett egyszerűbben is lehet eljárni, egy irányt koordinátairányként választva és párhuzamosan ehhez az irányhoz, egyenesek segítségével, összekötve a görbevonallú négyzet perempontjait, az összekötéseken lineárisan változtatva a másik koordinátát – de ekkor a perem gyors változásai elég durván terjednek a négyzet belsejébe. Ezért célszerű, az így előállított  $(x_1)_{k\ell}$  és  $(x_2)_{k\ell}$  értékeket a következőképpen javítani : néhányszor végigmegyünk az összes belső rácsponton, egy-egy pontban először  $x_1$ -, majd  $x_2$ -koordinátáját a négy szomszédpontja  $x_1$ - ill.  $x_2$ -koordinátáinak átlagával helyettesítve. Ez semmi más, mint a Gauss-Seidel eljárás alkalmazása az (15.256), ill. (15.257) egyenletek megoldására!



ábra 15.31: Kocka szétvágása 5 tetraéderre

Ha háromszöges elemekkel akarunk dolgozni – erre a döntésre lejjebb visszatérünk – akkor pl. a négyszögekbe az egyik átlót húzzuk be (ld. Egidi és Maponi cikkét is). Háromszöges rács közvetlen szerkesztése miatt ld. pl. Friedrich és Bowyer cikkeit, a háromszögek minőségének méréséről pl. Pébay és Baker cikkét.

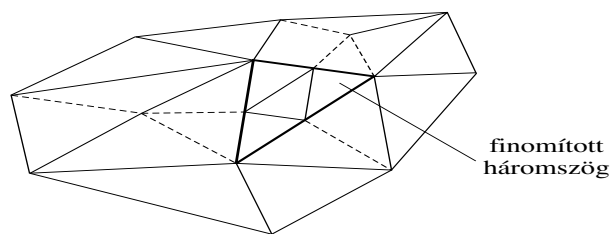
A kétdimenziós rácsszerkesztéshez ld. Godunov és Prokopov cikkét is vagy Liseikin könyveit.

Háromdimenziós rács szerkesztése bonyolultabb. Például tetraéderekkel tölthetjük ki a háromdimenziós tartományt. A tetraéderekre való felbontást úgy állíthatjuk elő, hogy először prizmákra vágjuk szét a tartományt (kitüntetett irányban), majd a prizmákat kockákra és maradékdarabokra, a kockákat viszont 5-5, a maradékot pedig 4, 3, vagy 2 tetraéderre. Vagy pedig egy koordinátát tüntetünk ki, amelyhez merőlegesen kétdimenziós rácsot szerkesztünk. A kitüntetett koordináta irányába eltolva a rácsot, prizmatikus elemeket kapunk. Ezen módszer fejlettebb változata az, hogy alkalmas térgörbe mentén forgatunk egy kétdimenziós rácsot, a rács méreteit a peremhez igazítva.

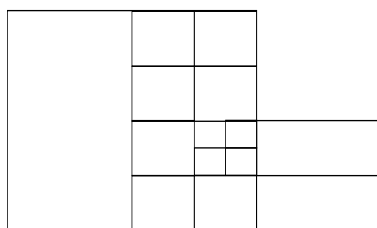
Az eredeti differenciálegyenlet megoldása után szükséges lehet a rács változtatása : ahol a megoldás nagyobb gradiensekkel rendelkezik, ott célszerű sűrűbben felosztani a tartományt. Ezt a folyamatot automatizálva, *adaptív módszert* kapunk. Ha minden résztartományban rendelkezünk a megoldás hibájának becslésével, akkor átrendezhetjük a rácsot úgy, hogy kis pontszám mellett adott pontossági korlát alatti hibát érünk el. Ehhez a témához ld. Babuška és Szabó, Bank és társai, Gáspár és Józsa, W.F. Mitchell, Rivara, valamint Zienkiewicz és Zhu munkáit.

Most arra a kérdésre térünk rá, hogy vegyünk háromszögeket vagy téglalapokat? Hasonló a „tetraéder vagy kocka?” kérdés. A válasz nem egyértelmű, és különböző oldalokról világítjuk meg a problémát.

A *háromszögek* előnye az, hogy jól alkalmazkodnak mindenféle tartomány-



ábra 15.32: Háromszögek konform finomítása



ábra 15.33: Téglalapok nemkonform finomítása

hoz. Könnyen megvalósítható a *háromszögek* konform (vagy legalább stabil) *finomítása* is (ehhez ld. Bank és társai vagy Rivara dolgozatait. A *tetraéderek stabil finomításához* ld. Liu és Joe, Zhang, Bey, valamint Kröger és Preusser cikkeit.) Az elméleti (pontossági becslések stb.) vizsgálatok valamivel egyszerűbbek.

A háromszögek hátránya viszont, hogy sok adatot kell nyilvántartani, nehéz több háromszög csúcaiban adott függvényértékek alapján magasabb rendű interpolációt készíteni, nem egyszerű (de megcsinálható) a többrácsos módszerben szükséges simító iterációk párhuzamosítása. A háromszöges elemeket inkább az elméleti vizsgálatoknál veszik alapul, a gyakorlatban gyakran csak egyenesoldalú háromszög elemeket használnak.

A *téglalapok* előnye, hogy nem veszítjük el az esetlegesen meglévő szimetriát, kevés adat kell a nyilvántartáshoz, a párhuzamosítás kézenfekvő, magasabb rendű interpoláció nem jelent problémát. Nehezebb viszont a finomítása : ha lokálisan akarjuk elkészíteni, akkor elveszítjük a 90-fokos szögeket, vagy keletkeznek *nemkonform csomópontok*. Ehhez ld. a 59. fejeletet.

Ez a többrácsos módszer közbülső rácsein nem jelent problémát : elég, ha a legfinomabb rácson minden pont konform. Viszont több szakember máskor is dolgozik nemkonform pontokkal és csak arra ügyel, hogy ezeket elsőnek eliminálja. A végelem konvergencia elméletet kiterjesztették erre az esetre, hogy nemkonform pontok is előfordulnak. Téglalapos nemkonform trianguláción véges térfogat módszert fejlesztett ki Gáspár Cs., a keletkező

lineáris egyenletrendszer többrácsos megoldásával együtt. A téma véges differencia vonatkozásban szerepel *FAC-algoritmus* név alatt Briggs, Henson és McCormick többrácsos könyvében is.

A téglalapos elemekre vonatkozó elméleti vizsgálatok bonyolultabbak, ld. Girault és Raviart könyvét. Ciarlet affín végeelem családokra fejlesztett elmélete legtöbbször nem alkalmazható. A téglalapos elemek kevésbé gazdaságosak mint a háromszögesek abban az értelemben, hogy több szabadsági fokkal rendelkeznek, mint az egyenlő rendű háromszöges elemek, ahogyan a 15.7.5-beli táblázatból azonnal kitűnik. Így több aritmetikai művelet kíván a megoldás előállítására. Ugyanakkor jobb lesz a téglalapos rácshoz tartozó megoldás hibakonstansa egyenlő rend esetén. Ezenkívül nem vesznek kárba a „kiegészítő” szabadsági fokok, hanem a megoldásoknak többoldalú viselkedési lehetőséget kölcsönöznek. Például a bilineáris elem esetén az  $x_1x_2$  jelenléte az elem polinomjai között az eredmények kiértékelésénél előnyös, mert (szintvonalak rajzolásakor) így az elemen belül is lehet nyeregpontra a megoldásnak – míg a Courant–elemnél az nem fordulhat elő.

Ezenkívül – legalább az  $R_8$  és  $R_9$  esetében – ismeretes, hogy az  $R_9$  „kiegészítő” szabadsági foka biztosítja, hogy a konvergenciarend nem csökken általános rácssorozaton, ld. Arnold, Boffi, Falk cikkét.

Az általánosabb tartományhoz akár háromszöges, akár téglalapos görbevonallú elemek esetén lesz jó a közelítés, az izoparametrikus módszert alkalmazva (ld. 15.7.10). Ekkor viszont természetesebb is lehet a téglalap, mint a háromszög, mert bonyolult tartomány esetén a legjobb a testre szabott koordinátarendszer használata. Ehhez olyan ortogonális (vagy ahhoz közeli) rendszert használjunk, amelynek egyik koordináta tengelye a perem, a másik erre (majdnem) ortogonális. Háromdimenziós esetben hasonló alapon prizmatikus elemeket lehet javasolni.

Konvekció–diffúzió feladatok megoldásakor a lineáris háromszöges elem esetén a legtöbb csomópontban csak 6 pontos karakterisztikát lehet garantálni, míg téglalapok esetén 8 (a koordináta irányok mellett az átlók irányait is), megfelelő diszkretizációt alkalmazva, ld. pl. a \*\*\* 17.4.0. pontbeli (48) sémát.

Végül is, éppen téglalapok esetén alkalmazhatjuk a ritka rácsok módszerét (ld. Zenger, ill. Bungartz és Griebel cikkét), a háromdimenziós esetet ld. pl. Balder és Zenger cikkében.

## 15.8 A végeelem módszerből keletkezett lineáris rendszerek megoldása

A végeelem módszer által eredményezett speciális, ritka mátrixú egyenletrendszerek megoldására legalkalmasabb eljárás a többrácsos módszer. A sávos Gauss-elimináció erre a célra sok helyen még használatban van, de sem tárígegy szerint, sem műveletigény szerint nem versenyképes, amikor néhány száz ismeretlennél több van (ld. az I. 1.6.1-ben közölt kis táblázatot, amely a háromdimenziós Poisson-egyenletből keletkezett egyenletrendszer megoldására vonatkozik, amikor minden irányban 100 belső pontunk van, összesen tehát  $10^6$  ismeretlen). A táblázatot a következőképpen egészíthetjük ki a teljes,  $V$ -ciklusos többrácsos módszer adataival, amikor a tartomány az egységkocka,  $h_\ell = 1/N_\ell = \frac{1}{2}h_{\ell-1}$ , 1 elő- és 1 utósímító iterációval, tri-lineáris interpolációval, az ehhez adjungált restriktcióval, és 6 iterációval a legfelsőbb szinten dolgozunk, az  $\varepsilon = 10^{-4}$  pontosságnak megfelelően – mert arra számíthatunk ilyenkor, hogy minden iterációban a hiba  $1/5$  részére esik :

módszer	teljes többrácsos módszer
tárhely	17.4 Mb
műveletek száma	$124 \cdot 10^6$

A következőkben a többrácsos módszernek néhány, a végeelem módszerrel kapcsolatos jellemző vonását mutatjuk meg. Máskülönb a már korábban, 15.4-ben bemutatott algoritmus és elmélet érvényes. Ezután a módszer teljes konvergencia bizonyítását adjuk, majd párhuzamos módszerekre is kitérünk.

Egyébként a 15.4.10. pontot is érdemes elolvasni, ha a végeelem módszer alkalmazásából keletkezett lineáris rendszerek más mint a többrácsos módszerrel akarjuk megoldani.

### 15.8.1 Beágyazott végeelem terek

Lagrange-féle elemekkel dolgozunk. Az  $\ell$ -edik szinten úgy konstruáljuk a  $\mathcal{T}_\ell$  triangulációt, hogy a  $\mathcal{T}_{\ell-1}$  triangulációt alapnak vesszük és  $\mathcal{T}_{\ell-1}$  részterületeit tovább osztjuk, pl. felezéssel (amikor is  $h_{\ell-1} = 2h_\ell$  érvényes a triangulációkat jellemző lépéstávolságokra).  $\mathcal{T}_\ell$  konstrukciójának köszönhetően teljesül  $V_h(\mathcal{T}_{\ell-1}) \subset V_h(\mathcal{T}_\ell)$ , azaz *beágyazott végeelem terekről* lehet beszélni. Ekkor a többrácsos módszerben szükséges interpolációt és leszűkítést (azaz restriktciót) természetes módon kapjuk meg. Ennek alapja, hogy ilyenkor a

$V_h(\mathcal{T}_\ell)$  végeelem tér  $\{w_m^\ell\}_{m=1}^{N_\ell}$  bázisfüggvényeiből kiszámíthatók a  $V_h(\mathcal{T}_{\ell-1})$  tér  $\{w_n^{\ell-1}\}_{n=1}^{N_{\ell-1}}$  bázisfüggvényei :

$$w_n^{\ell-1}(x) = \sum_{m=1}^{N_\ell} \beta_{nm}^\ell w_m^\ell(x), \quad 1 \leq n \leq N_{\ell-1}. \quad (15.258)$$

Ez mátrixalakban foglalható össze mint

$$\vec{w}_{\ell-1} = B_\ell \vec{w}_\ell, \quad \text{ahol } \vec{w}_\ell := (w_1^\ell, \dots, w_{N_\ell}^\ell)^T, \quad B_\ell := (\beta_{nm}^\ell).$$

Ez a mátrix téglalapalakú és ritka (ld. a 61. feladatot is). Legyen  $x_i^{\ell-1}$  a  $\mathcal{T}_{\ell-1}$  trianguláció tetszőleges csomópontja és  $x_j^\ell$  ugyanez a csomópont  $\mathcal{T}_\ell$ -ben. Ekkor

$$w_n^{\ell-1}(x_i^{\ell-1}) = \delta_{ni}, \quad w_m^\ell(x_j^\ell) = \delta_{mj},$$

és (15.258)-ből következik

$$1 = w_i^{\ell-1}(x_i^{\ell-1}) = \sum_{m=1}^{N_\ell} \beta_{im}^\ell w_m^\ell(x_j^\ell) = \beta_{ij}^\ell,$$

míg  $\mathcal{T}_\ell$  tetszőleges más  $x_k^\ell$  csomópontjára érvényes

$$w_n^{\ell-1}(x_k^\ell) = \beta_{nk}^\ell.$$

Lineáris elem esetén ez azt jelenti, hogy  $1 \geq \beta_{nk}^\ell \geq 0$ . Ha  $\beta_{nk}^\ell \neq 0$ , akkor  $x_n^{\ell-1}$  és  $x_k^\ell$  szomszédok.

Most legyen  $u_h^{\ell-1} = \sum_{j=1}^{N_{\ell-1}} y_j^{\ell-1} w_j^{\ell-1} \in V_h(\mathcal{T}_{\ell-1})$ , akkor (15.258) szerint

$$u_h^{\ell-1}(x) = (\vec{y}_{\ell-1})^T \vec{w}_{\ell-1}(x) = (\vec{y}_{\ell-1})^T B_\ell \vec{w}_\ell(x) = (\vec{y}_\ell)^T \vec{w}_\ell(x),$$

a

$$y_{\ell-1} := (y_1^{\ell-1}, \dots, y_{N_{\ell-1}}^{\ell-1})^T, \quad y_\ell := B_\ell^T y_{\ell-1} \quad (15.259)$$

jelölésekkel. Ezzel definiáltuk az együtthatók vektoráinak az átszámítását az  $\ell - 1$ -edik szintről az  $\ell$ -edik szintre – amelyből a  $P_{\ell-1}^\ell : V_h(\mathcal{T}_{\ell-1}) \rightarrow V_h(\mathcal{T}_\ell)$  interpoláció képlete következik. Az  $u_h^{\ell-1} \in V_h(\mathcal{T}_{\ell-1})$  függvényre nézve ez csupán azt jelenti, hogy a  $V_h(\mathcal{T}_\ell)$  tér elemének is tekintjük, mivel  $V_h(\mathcal{T}_{\ell-1}) \subset V_h(\mathcal{T}_\ell)$ .

Következőnek belátjuk, hogy a restrikiót éppen a  $B_\ell$  mátrix adja meg. Ehhez abból indulunk ki, hogy a restrikiót a maradékvektorra fogjuk alkalmazni. Az utóbbinak része viszont a  $b_\ell$  jobboldal (v.ö. (15.104)-gyel és (15.105)-tel is 15.4.7. elején). A variációs feladatnak megfelelően

$$b_m^\ell = (f, w_m^\ell)_0, \quad 1 \leq m \leq N_\ell, \quad b_\ell := (b_1^\ell, \dots, b_{N_\ell}^\ell)^T,$$

azaz  $b_\ell = \Pi_\ell f$ , ahol  $\Pi_\ell f$  az  $f$  függvény *Galjorkin-projekciója* :

$$\Pi_\ell f := ((f, w_1^\ell)_0, \dots, (f, w_{N_\ell}^\ell)_0)^T. \quad (15.260)$$

Mivel (15.258) szerint

$$b_i^{\ell-1} = (f, w_i^{\ell-1})_0 = \sum_{m=1}^{N_\ell} \beta_{im}^\ell (f, w_m^\ell)_0 = \sum_{m=1}^{N_\ell} \beta_{im}^\ell b_m^\ell,$$

így tehát mátrixalakban

$$b_{\ell-1} = B_\ell b_\ell,$$

vagyis

$$\Pi_{\ell-1} f = B_\ell \Pi_\ell f. \quad (15.261)$$

Azaz  $B_\ell$  valóban határozza meg a restrikiót.

A beágyazott végeelem módszernek ezen restrikiós és interpolációs operátorait *kanonikusnak* hívjuk, és ezek az előzőek szerint egymásnak adjungáltjai :

$$P_{\ell-1}^\ell = B_\ell^T = (R_\ell^{\ell-1})^T, \quad R_\ell^{\ell-1} = B_\ell,$$

míg a véges differencia alapú többrácsos módszer esetén javasoltuk a (15.90) relációt a 15.4.3. pontban.

A végeelem alapú többrácsos módszer programozásánál azt kell figyelembe venni, hogy – a differencia egyenletekhez képest – a  $h^d$ -nel megszorított egyenletekkel állunk szemben ( $d = 2, 3$ ). Így amikor a maradékvektort az  $\ell$ -edik szintről (amelynek jellemző lépéstávolsága  $h_\ell$ ) az  $\ell - 1$ -edik szintre visszük át (ahol  $h_{\ell-1} = 2h_\ell$  a jellemző lépéstávolság), hozzá kell tenni még egy  $2^d$  szorzót. Ez a szorzó automatikusan figyelembe van véve a kanonikus (15.259), (15.261) interpolációval és restrikióval, és a differencia esetben a (15.90) relációval.

(15.258) további következménye az, hogy az  $\ell$ -edik és  $\ell - 1$ -edik szinthez tartozó mátrixok között van egy egyszerű összefüggés :

$$(A_{\ell-1})_{ij} = a(w_j^{\ell-1}, w_i^{\ell-1}) = \sum_{m,n} \beta_{im}^\ell a(w_n^\ell, w_m^\ell) \beta_{jn}^\ell = (B_\ell A_\ell B_\ell^T)_{ij},$$

tehát

$$A_{\ell-1} = B_\ell A_\ell B_\ell^T = R_\ell^{\ell-1} A_\ell P_{\ell-1}^\ell. \quad (15.262)$$

Ezzel a (15.96) relációt 15.4.6-ból tetszőleges Lagrange-típusú beágyazott végeelem approximációra általánosítottuk. Az összefüggés neve *Galjorkin-reláció*. Segítségével a durva rácshoz tartozó mátrixokat akkor számíthatjuk



ki, ha eredetileg csak a legfinomabb rácson adott a mátrix (v.ö. a 15.4.10. ponttal).

Megemlítjük, hogy vannak helyzetek, amikor a nem beágyazott terek használata célszerű, ld. pl. Adams, vagy Mohr és Wienands, továbbá John, Knobloch és társai cikkét.

### 15.8.2 A végeelem alapú többrácsos módszer konvergenciája \*\*\* 15.8.1.

Elsőnek a többrácsos módszer konvergenciájához szükséges approximációs tulajdonságot bizonyítjuk be. Ebben döntő szerepet játszik az

$$a(u, v) = (f, v) \quad \text{minden } v \in V\text{-re} \quad (15.263)$$

és

$$a(u_h^\ell, v_h) = (f, v_h) \quad \text{minden } v_h \in V_h(\mathcal{T}_\ell)\text{-re} \quad (15.264)$$

variációs egyenletek megoldásainak különbségére ismert

$$\|u - u_h^\ell\|_0 \leq \bar{c} h_\ell^2 \|f\|_0 \quad (15.265)$$

becslés. Ez a 15.32. következtetés (15.232) egyenlőtlenségéből a *Nitsche-fogás* segítségével létrehozható

$$\|u - u_h^\ell\|_0 \leq \bar{c}_* h_\ell^2 |u|_2$$

becslés és (15.209) kombinációja, és a 15.30. és 15.32. tételekben felsorolt feltételek mellett érvényes. (A Nitsche-fogáshoz ld. II. 11.6.3.)

**15.37. Lemma** (végeelem alapú többrácsos módszer approximációs tulajdonsága). Ha az  $u_h^\ell$  végeelem megoldás eleget tesz a (15.265) becslésnek, akkor igaz az approximációs tulajdonság, azaz van olyan  $c$  konstans, hogy a  $C_\ell := A_\ell^{-1} - P_{\ell-1}^\ell A_{\ell-1}^{-1} R_\ell^{\ell-1}$  mátrixra teljesül az euklideszi normában a  $\|C_\ell\| \leq c$  becslés.

**Bizonyítás.** Becsüljük a  $C_\ell = A_\ell^{-1} - B_\ell^T A_{\ell-1}^{-1} B_\ell$  mátrix euklideszi normáját! Ehhez (v.ö. (15.104)-gyel és (15.105)-tel 15.4.7-ben) a tetszőleges  $g_\ell \in \mathbb{R}^{N_\ell}$  vektorral felírt

$$A_\ell y_\ell = g_\ell \quad \text{és} \quad A_{\ell-1} y_{\ell-1} = R_\ell^{\ell-1} g_\ell =: g_{\ell-1} \quad (15.266)$$

egyenletrendszerek megoldásaival képzett

$$y_\ell - P_{\ell-1}^\ell y_{\ell-1}$$

vektor normáját kell becsülnünk. Az adott (15.266) egyenletrendszer  $g_\ell$  jobb oldalához gyártunk először egy alkalmas  $\psi_\ell = (\psi_1^\ell, \dots, \psi_{N_\ell}^\ell)^T$  vektor segítségével az  $f_\ell = \psi_\ell^T \vec{w}_\ell \in V_h(\mathcal{T}_\ell)$  függvényt úgy, hogy teljesüljön

$$\Pi_\ell f_\ell = g_\ell.$$

Ehhez (15.260) szerint a

$$g_i^\ell = (\psi_\ell^T \vec{w}_\ell, w_i^\ell)_0 = \sum_{j=1}^{N_\ell} \psi_j^\ell(w_j^\ell, w_i^\ell)_0 = (G_\ell \psi_\ell)_i$$

összefüggést kell biztosítanunk, tehát a

$$G_\ell \psi_\ell = g_\ell$$

egyenletrendszert kell megoldanunk.

Mivel itt  $G_\ell$  a  $\{w_i^\ell\}$  rendszer Gram-féle mátrixa, ezen egyenletrendszer megoldása létezik, egyértelműen meghatározott és érvényes a

$$(G_\ell \psi_\ell, \psi_\ell) = (g_\ell, \psi_\ell) \leq \|g_\ell\| \|\psi_\ell\| \leq \frac{c_0}{h_\ell} \|g_\ell\| \|f_\ell\|_0 \quad (15.267)$$

becslés. Itt az  $L_2$ -féle és az euklideszi normák ekvivalenciáját használtuk fel :

$$\frac{h_\ell}{c_0} \|\psi_\ell\| \leq \|f_\ell\|_0 \leq c_1 h_\ell \|\psi_\ell\|, \quad \text{ha } f_\ell := \psi_\ell^T \vec{w}_\ell, \quad (15.268)$$

kétdimenziós tartomány esetén (az egydimenziós esethez ld. a 62. feladatot), valamint ezen bizonyítás utáni lemmát). Ha azt is figyelembe vesszük, hogy

$$\|f_\ell\|_0^2 = (G_\ell \psi_\ell, \psi_\ell), \quad (15.269)$$

akkor (15.267)-ből következik

$$\|f_\ell\|_0 \leq \frac{c_0}{h_\ell} \|g_\ell\|. \quad (15.270)$$

Az így jellemzett  $f_\ell$  függvény és (15.263) határozzák meg  $u$ -t. Közelítése az  $\ell$ -edik szinten a (15.264) megoldása :  $u_h^\ell = y_\ell^T \vec{w}_\ell$ , ahol  $y_\ell$  eleget tesz (15.266)-nak. Az  $\ell - 1$ -edik szinten  $u$  közelítése  $u_h^{\ell-1} = y_{\ell-1}^T \vec{w}_{\ell-1}$ , ld. (15.266), ahol  $y_{\ell-1} = A_{\ell-1}^{-1} g_{\ell-1}$  és  $g_{\ell-1} = B_\ell g_\ell = \Pi_{\ell-1} f_\ell$ , (15.261) szerint. (15.265) és (15.270) alapján következik

$$\|u - u_h^\ell\|_0 \leq \bar{c} h_\ell^2 \|f_\ell\|_0 \leq \bar{c} c_0 h_\ell \|g_\ell\|$$

és

$$\|u - u_h^{\ell-1}\|_0 \leq \bar{c}h_{\ell-1}^2\|f_\ell\|_0 = 4\bar{c}h_\ell^2\|f_\ell\|_0 \leq 4\bar{c}c_0h_\ell\|g_\ell\|,$$

amennyiben  $h_{\ell-1} = 2h_\ell$ . Ekkor

$$\|u_h^\ell - u_h^{\ell-1}\|_0 \leq 5\bar{c}c_0h_\ell\|g_\ell\|. \quad (15.271)$$

Ezen egyenlőtlenség baloldalára szeretnénk (15.268) alsó becslését alkalmazni. Míg az  $u_h^\ell$  és  $u_h^{\ell-1}$  függvények közvetlenül összehasonlíthatók, az  $u_h^{\ell-1}$ -et meghatározó  $y_{\ell-1}$  vektort még (15.259) segítségével át kell számítani az  $\ell$ -edik szintre :

$$u_h^{\ell-1} = (B_\ell^T y_{\ell-1})^T \vec{w}_\ell.$$

Ezután (15.268) és (15.271) adja a

$$\frac{h_\ell}{c_0}\|y_\ell - B_\ell^T y_{\ell-1}\| \leq \|u_h^\ell - u_h^{\ell-1}\|_0 \leq 5\bar{c}c_0h_\ell\|g_\ell\|$$

becslést, vagyis

$$\|(A_\ell^{-1} - B_\ell^T A_{\ell-1}^{-1} B_\ell)g_\ell\| \leq 5\bar{c}c_0^2\|g_\ell\|,$$

és így (mivel a  $g_\ell$ -t tetszőleges vektornak tételeztük fel)  $\|C_\ell\| \leq 5\bar{c}c_0^2$ .  $\square$

A fenti bizonyítás egyik lényeges lépése az  $L_2$ -norma és a súlyozott euklideszi norma ekvivalenciájának a hasznosítása. Ezt a következő lemmában egy speciális, de fontos esetben mutatjuk meg.

**15.38. Lemma** (két norma ekvivalenciája a Courant–elemek esetén). Legyen  $\Omega \in \mathbb{R}^2$  poligonális tartomány és  $\mathcal{T}_h$  annak egyenletesen reguláris triangulációja. Legyen  $V_h = \text{span}\{w_i\}_{i=1}^N$  a Courant–elemek által meghatározott végeselem tér,  $f := \sum_{i=1}^N \psi_i w_i$  a tetszőleges eleme és  $\|f\|_0$  annak  $L_2(\Omega)$ -normája. Akkor igaz a következő egyenlőtlenség :

$$\frac{1}{c_0}\|\psi\|_{(0,h)} \leq \|f\|_0 \leq c_1\|\psi\|_{(0,h)}, \quad c_0 := \frac{2\kappa}{27^{1/4}}, \quad c_1 := \frac{\sqrt{\kappa\pi}}{12^{1/4}},$$

ahol  $\kappa$  az egyenletesen reguláris trianguláció (15.217) konstansa, továbbá  $\psi := (\psi_1, \dots, \psi_N)^T$  és  $\|\psi\|_{(0,h)}^2 := \sum_{i=1}^N \psi_i^2 h^2$ .

**Bizonyítás.** Az  $\psi$  vektor helyett tekintjük az olyan hosszabb  $\bar{\psi}$  vektort, amely az  $\Omega$ -beli csomópontokon megegyezik  $\psi$ -fel, a tartomány peremén fekvő csomópontokon viszont nulla értéket vesz fel. A  $\bar{\psi}$  vektor komponenseit ugyancsak  $\psi_i$ -vel jelöljük (ahol  $\psi_i = f(x_i)$ ), ill. amikor  $x_i = x^{k,j}$  a  $k$ -edik háromszög  $j$ -edik csúcsa,  $\psi_i = f(x^{k,j})$  – mivel Lagrange–elemről van szó).

Ennek eredménye az, hogy minden  $\mathcal{T}_h$ -beli  $\Delta_k$  háromszög felírhatjuk, hogy

$$\begin{aligned} \int_{\Delta_k} f^2(x) \, dx &= \frac{|\Delta_k|}{6} \left( \sum_{j=1}^3 f^2(x^{k,j}) \right. \\ &\quad \left. + f(x^{k,1})f(x^{k,2}) + f(x^{k,2})f(x^{k,3}) + f(x^{k,3})f(x^{k,1}) \right) \\ &= \frac{|\Delta_k|}{6} \varphi_k^T \mathcal{A} \varphi_k, \\ \varphi_k &:= (f(x^{k,1}), f(x^{k,2}), f(x^{k,3}))^T, \quad \mathcal{A} := \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}. \end{aligned}$$

Mivel ezen (szimmetrikus, pozitív definit) mátrix sajátértékei  $\frac{1}{2}, \frac{1}{2}, 2$ , így

$$\frac{1}{2} \|\varphi_k\|^2 \leq \varphi_k^T \mathcal{A} \varphi_k \leq 2 \|\varphi_k\|^2,$$

vagyis

$$\frac{|\Delta_k|}{12} \sum_{j=1}^3 f^2(x^{k,j}) \leq \int_{\Delta_k} f^2(x) \, dx \leq \frac{|\Delta_k|}{3} \sum_{j=1}^3 f^2(x^{k,j}).$$

Innen következik a lemma állítása az összes háromszög feletti összegzéssel, figyelembe véve, hogy tetszőleges  $\Delta_k$  háromszög esetén (ha  $\rho_k$  a beírt körének sugara és  $h_k$  a leghosszabb oldala) érvényes

$$3\sqrt{3}\rho_k^2 \leq |\Delta_k| \leq \frac{\sqrt{3}}{4}h_k^2,$$

és a trianguláció egyenletes regularitása miatt igaz  $h = \max_i h_i \leq \kappa\rho_k$ , valamint (a felső becslésnél), hogy egy sarokpontban legfeljebb  $\frac{2\pi}{\vartheta_{\min}}$  háromszög találkozhat, ahol  $\vartheta_{\min} > \sin \vartheta_{\min} \geq \frac{1}{\kappa}$ , ld. (15.218) 15.7.5-ben.  $\square$

**Megjegyzés.** A lemma eredménye úgy is kimondható, hogy a Gram-féle mátrix jól kondicionált, hiszen  $\text{cond } G_\ell \leq c_0 c_1$  következik függetlenül  $h_\ell$ -től és  $\ell$ -től, ld. (15.269).  $\square$

Az approximációs tulajdonság ilyenformán megvan. A simítási tulajdonság levezetése úgy történik, mint azelőtt 15.4.7-ben. A különbség csak annyi, hogy a végeelem módszernél, kétdimenziós másodrendű feladat esetén,

$$\lambda_{\max}(A_\ell) \leq \|A_\ell\| \leq c = c(\kappa), \quad (15.272)$$

míg  $d$ -dimenziós, másodrendű feladatnál  $\|A_\ell\| = O(h^{d-2})$ . Ehhez ld. pl. (15.242)-t 15.7.8-ban, valamint a 63. feladatot, majd ezután 64. feladatot is.

Az előzők alapján a többrácsos módszer konvergenciája is következik.

**15.39. Tétel** (Hackbusch; végeelem alapú többrácsos módszer konvergenciája). A végeelem alapú többrácsos módszer (kanonikus restrikcióval és interpolációval, csillapított Jacobi-iterációval mint simító iterációval,  $\omega = 1/c$ , ahol  $c$  a (15.272) korlát) konvergens. A kétrácsos módszerre érvényes a

$$\|K_\ell(\nu)\| = \|(A_\ell^{-1} - B_\ell^T A_{\ell-1}^{-1} B_\ell) A_\ell (I - \omega A_\ell)^\nu\| \leq \frac{5\bar{c}c_0^2 c}{2(1 + \nu)} \quad (15.273)$$

becslés.

**Bizonyítás.** A kétrácsos módszer  $K_\ell(\nu)$  átmeneti mátrix normájára a (15.273) becslést kapjuk a 15.11. és 15.37. lemmákból, figyelembe véve (15.272)-t. Ezután alkalmazzuk a 15.13. vagy 15.14. tételt.  $\square$

**Megjegyzés.** A 15.12. lemma szerint a fenti eredmény ugyancsak igaz akkor is, ha a simító iteráció a Gauss–Seidel módszer.  $\square$

A végeelem alapú többrácsos módszer (tetszőleges trianguláció esetén) az elmélet szempontjából igen, de a program szempontjából nem annyira hatékony, mint a véges differencia-alapú többrácsos módszer. Ugyanis az utóbbi esetben az egyszerűbb szomszédsági viszonyok miatt a fehér-fekete, illetve négy szín Gauss–Seidel iterációt alkalmazhatjuk, ezenfelül magasabbrendű interpolációkat és restrikciókat is, melyek következtében a többrácsos program lényegesen jobban működik (65. feladat).

### 15.8.3 A szimmetrikus eset

A többrácsos módszernek 15.4.8-ban adott konvergencia bizonyításában hiányzott a  $\gamma = 1$  eset, azaz a  $V$ -ciklus. Az alábbiakban ezt pótoljuk abban az esetben, amikor az iterációs mátrix szimmetrikus, ami a legegyszerűbben a végeelem diszkretizációval biztosítható.

Az  $M_\ell$  iterációs mátrix ezen és más tulajdonságainak tisztázásához abból indulunk ki, hogy az  $\ell$ -edik szinten,  $\nu_1$  előiteráció és  $\nu_2$  utóiteráció esetén  $M_\ell$  15.4.8-beli rekurzív képlete (15.114) szerint :

$$M_\ell(\nu_1, \nu_2) = K_\ell(\nu_1, \nu_2) + S_\ell^{\nu_2} P_{\ell-1}^\ell (M_{\ell-1}(\nu_1, \nu_2))^\gamma A_{\ell-1}^{-1} R_\ell^{\ell-1} A_\ell S_\ell^{\nu_1}, \quad (15.274)$$

amikor  $\ell \geq 1$ , és  $M_0(\nu_1, \nu_2) := 0$ . Itt  $\gamma \geq 1$ , ezenkívül

$$K_\ell(\nu_1, \nu_2) = S_\ell^{\nu_2} (I_\ell - P_{\ell-1}^\ell A_{\ell-1}^{-1} R_\ell^{\ell-1} A_\ell) S_\ell^{\nu_1} \quad (15.275)$$

a kétrácsos módszer iterációs mátrixa.

Tekintsük most a szimmetrikus esetet, azaz legyen érvényes

$$A_\ell = A_\ell^T > 0, \quad R_\ell^{\ell-1} = (P_{\ell-1}^\ell)^T, \quad (15.276)$$

$$S_\ell = I_\ell - W_\ell^{-1}A_\ell, \quad W_\ell = W_\ell^T \geq A_\ell. \quad (15.277)$$

Ahogy tudjuk 15.4.7-ből, a csillapított Jacobi- és a szimmetrikus Gauss-Seidel eljárás iterációs mátrixai adhatók meg (15.277) alakban.

Továbbá, tegyük fel, hogy teljesül a (15.262) Galjorkin-reláció. Ez 15.8.1. szerint pl. beágyazott végeselem terek esetén biztosított. Abban az esetben, amikor  $\nu_1 = \nu_2$ , az  $M_\ell(\nu_1, \nu_2)$ , ill.  $K_\ell(\nu_1, \nu_2)$  mátrixokat  $M_\ell(\nu)$ -vel, ill.  $K_\ell(\nu)$ -vel jelöljük,  $\nu_1 = \nu_2 =: \frac{\nu}{2}$ . (Később fogjuk értelmezni, hogy  $\nu$  akár páratlan is lehet.)

Mielőtt a konvergencia bizonyításához térnénk át, ezen mátrixok tulajdonságaival foglalkozunk.

**15.40. Lemma** (Hackbusch; a többrácsos iterációs mátrix tulajdonságai, 1). A (15.276)-(15.262) feltételek mellett érvényesek a következők :

- a) az  $\widetilde{M}_\ell = \widetilde{M}_\ell(\nu) := A_\ell^{1/2} M_\ell(\nu) A_\ell^{-1/2}$  és a  $\widetilde{K}_\ell = \widetilde{K}_\ell(\nu) := A_\ell^{1/2} K_\ell(\nu) A_\ell^{-1/2}$  mátrixok szimmetrikusak és pozitív szemidefinitek;
- b)  $\widetilde{M}_\ell(\nu_1, \nu_2) := A_\ell^{1/2} M_\ell(\nu_1, \nu_2) A_\ell^{-1/2} = (\widetilde{M}_\ell(\nu_2, \nu_1))^T$ ;
- c) ha ezenkívül  $\gamma = 1$ , akkor  $\widetilde{M}_\ell(0, \nu_2) \widetilde{M}_\ell(\nu_1, 0) = \widetilde{M}_\ell(\nu_1, \nu_2)$ .

**Bizonyítás.** Legyen

$$\widetilde{S}_\ell := A_\ell^{1/2} S_\ell A_\ell^{-1/2} = I_\ell - X_\ell, \quad \text{ahol } X_\ell := A_\ell^{1/2} W_\ell^{-1} A_\ell^{1/2}.$$

Ez utóbbi – szimmetrikus – mátrixról (15.276), (15.277) alapján következik

$$0 < X_\ell \leq I_\ell,$$

így  $\widetilde{S}_\ell$  szimmetrikus és pozitív szemidefinit :  $0 \leq \widetilde{S}_\ell < I_\ell$ . Emiatt  $\widetilde{S}_\ell^{\nu/2}$  akkor is jól definiált, amikor  $\nu$  páratlan. Elsőnek vizsgáljuk a kétrácsos eljárás  $\widetilde{K}_\ell$  transzformált mátrixát :

$$\begin{aligned} \widetilde{K}_\ell &= \widetilde{S}_\ell^{\nu/2} A_\ell^{1/2} (I_\ell - P_{\ell-1}^\ell A_{\ell-1}^{-1} R_{\ell-1}^{\ell-1} A_\ell) A_\ell^{-1/2} \widetilde{S}_\ell^{\nu/2} \\ &= \widetilde{S}_\ell^{\nu/2} \widetilde{Q}_\ell \widetilde{S}_\ell^{\nu/2}, \end{aligned} \quad (15.278)$$

ahol

$$\widetilde{Q}_\ell := I_\ell - \widetilde{P}_{\ell-1}^\ell \widetilde{R}_\ell^{\ell-1}, \quad \widetilde{P}_{\ell-1}^\ell := A_\ell^{1/2} P_{\ell-1}^\ell A_{\ell-1}^{-1/2}, \quad \widetilde{R}_\ell^{\ell-1} := A_{\ell-1}^{-1/2} R_{\ell-1}^{\ell-1} A_\ell^{1/2}.$$

Ez a  $\widetilde{Q}_\ell$  mátrix (15.276) miatt szimmetrikus. Mutassuk meg, hogy a következő további tulajdonságai vannak  $\widetilde{Q}_\ell$ -nek :

- $\alpha)$   $\tilde{Q}_\ell$  projekció :  $\tilde{Q}_\ell^2 = \tilde{Q}_\ell$  (v.ö. (15.97)-tel 15.4.6-ban);  
 $\beta)$   $\tilde{R}_\ell^{\ell-1} \tilde{P}_{\ell-1}^\ell = I_{\ell-1}$ ;  
 $\gamma)$   $\tilde{R}_\ell^{\ell-1} \tilde{Q}_\ell = 0$ ,  $\tilde{Q}_\ell \tilde{P}_{\ell-1}^\ell = 0$ .

Kezdjük  $\alpha)$ -val : (15.262) szerint

$$\begin{aligned}
 (\tilde{P}_{\ell-1}^\ell \tilde{R}_\ell^{\ell-1})^2 &= A_\ell^{1/2} P_{\ell-1}^\ell A_{\ell-1}^{-1} (R_\ell^{\ell-1} A_\ell^{1/2} A_\ell^{1/2} P_{\ell-1}^\ell) A_{\ell-1}^{-1} R_\ell^{\ell-1} A_\ell^{1/2} \\
 &= A_\ell^{1/2} P_{\ell-1}^\ell A_{\ell-1}^{-1} R_\ell^{\ell-1} A_\ell^{1/2} = \tilde{P}_{\ell-1}^\ell \tilde{R}_\ell^{\ell-1},
 \end{aligned}$$

és  $\tilde{P}_{\ell-1}^\ell \tilde{R}_\ell^{\ell-1}$ -lel együtt  $\tilde{Q}_\ell = I_\ell - \tilde{P}_{\ell-1}^\ell \tilde{R}_\ell^{\ell-1}$  is projekció. Mint projekció  $\tilde{Q}_\ell$  egyébként a

$$0 \leq \tilde{Q}_\ell \leq I_\ell \quad (15.279)$$

egyenlőtlenségeket teljesíti, tehát  $\tilde{Q}_\ell$  szimmetrikus és pozitív szemidefinit.

$\beta)$  Felhasználva megint a (15.262) feltételt, azt kapjuk meg, hogy

$$\tilde{R}_\ell^{\ell-1} \tilde{P}_{\ell-1}^\ell = A_{\ell-1}^{-1/2} (R_\ell^{\ell-1} A_\ell^{1/2} A_\ell^{1/2} P_{\ell-1}^\ell) A_{\ell-1}^{-1/2} = I_{\ell-1},$$

ahonnan  $\gamma)$  is következik :

$$\tilde{R}_\ell^{\ell-1} \tilde{Q}_\ell = \tilde{R}_\ell^{\ell-1} - \tilde{R}_\ell^{\ell-1} \tilde{P}_{\ell-1}^\ell \tilde{R}_\ell^{\ell-1} = 0, \quad \tilde{Q}_\ell \tilde{P}_{\ell-1}^\ell = \tilde{P}_{\ell-1}^\ell - \tilde{P}_{\ell-1}^\ell R_\ell^{\ell-1} \tilde{P}_{\ell-1}^\ell = 0.$$

Itt jegyezzük meg, hogy (15.276) miatt igaz

$$\tilde{R}_\ell^{\ell-1} = (\tilde{P}_{\ell-1}^\ell)^T. \quad (15.280)$$

Most a lemma a) állítását láthatjuk be : először is  $\tilde{Q}_\ell$ -lel együtt  $\tilde{K}_\ell$  is szimmetrikus és pozitív szemidefinit, ld. (15.278). Ezután foglalkozzunk  $\tilde{M}_\ell$ -lel.  $\ell = 1$ -re már igaz a hiányzó állítás :  $\tilde{M}_1 = \tilde{K}_1 = \tilde{K}_1^T = \tilde{M}_1 \geq 0$ . Legyen igaz az állítás  $\ell - 1$ -re is. Ekkor az

$$\tilde{M}_\ell = \tilde{K}_\ell + \tilde{S}_\ell^{\nu/2} \tilde{P}_{\ell-1}^\ell (\tilde{M}_{\ell-1})^\gamma \tilde{R}_\ell^{\ell-1} \tilde{S}_\ell^{\nu/2} \quad (15.281)$$

mátrix (15.280) alapján szimmetrikus. Mivel  $\tilde{M}_{\ell-1}$ -mal együtt  $(\tilde{M}_{\ell-1})^\gamma$  is szimmetrikus és pozitív szemidefinit, így

$$(\tilde{S}_\ell^{\nu/2} \tilde{P}_{\ell-1}^\ell (\tilde{M}_{\ell-1})^\gamma \tilde{R}_\ell^{\ell-1} \tilde{S}_\ell^{\nu/2} y, y) = ((\tilde{M}_{\ell-1})^\gamma v, v) \geq 0,$$

ahol  $v := \tilde{R}_\ell^{\ell-1} \tilde{S}_\ell^{\nu/2} y$ , tehát  $\tilde{M}_\ell \geq \tilde{K}_\ell \geq 0$ .

Hasonlóan bizonyíthatjuk be indukcióval a b) és c) azonosságokat, ugyanis (15.275)-ből és  $\alpha)$ -ból következik

$$\begin{aligned}
 (\tilde{K}_\ell(\nu_1, \nu_2))^T &= \tilde{S}_\ell^{\nu_1} \tilde{Q}_\ell \tilde{S}_\ell^{\nu_2} = \tilde{K}_\ell(\nu_2, \nu_1), \\
 \tilde{K}_\ell(0, \nu_2) \tilde{K}_\ell(\nu_1, 0) &= \tilde{S}_\ell^{\nu_2} \tilde{Q}_\ell \tilde{Q}_\ell \tilde{S}_\ell^{\nu_1} = \tilde{S}_\ell^{\nu_2} \tilde{Q}_\ell \tilde{S}_\ell^{\nu_1} = \tilde{K}_\ell(\nu_1, \nu_2).
 \end{aligned}$$

Felhasználva az  $\alpha$ – $\gamma$ ) relációkat kapjuk

$$\begin{aligned}
& \widetilde{M}_\ell(0, \nu_2) \widetilde{M}_\ell(\nu_1, 0) \\
&= \widetilde{S}_\ell^{\nu_2} \left[ \widetilde{Q}_\ell + \widetilde{P}_{\ell-1}^\ell M_{\ell-1}(0, \nu_2) \widetilde{R}_\ell^{\ell-1} \right] \left[ \widetilde{Q}_\ell + \widetilde{P}_{\ell-1}^\ell \widetilde{M}_{\ell-1}(\nu_1, 0) \widetilde{R}_\ell^{\ell-1} \right] \widetilde{S}_\ell^{\nu_1} \\
&= \widetilde{S}_\ell^{\nu_2} \left[ \widetilde{Q}_\ell + \widetilde{P}_{\ell-1}^\ell \widetilde{M}_{\ell-1}(0, \nu_2) \widetilde{M}_{\ell-1}(\nu_1, 0) \widetilde{R}_\ell^{\ell-1} \right] \widetilde{S}_\ell^{\nu_1} \\
&= \widetilde{M}_\ell(\nu_1, \nu_2),
\end{aligned}$$

hiszen az indukció szerint  $\ell - 1$  esetén már igaz az állítás, és az  $\ell = 1$  esetben  $\widetilde{M}_\ell(\nu_1, \nu_2) = \widetilde{K}_\ell(\nu_1, \nu_2)$ .  $\square$

Hadd mutassunk rá, hogy az  $\widetilde{S}^{\nu/2}$  mátrixok csak a konvergencia bizonyítás és nem a tényleges számítás alatt szerepelnek. A  $\nu_1 \neq \nu_2$  esetre visszatérünk.

A konvergencia bizonyításához szükségünk van  $\widetilde{M}_\ell$  felső becslésére is. Ehhez tegyük fel, hogy teljesül az approximációs tulajdonság a következő alakban (v.ö. a 15.15. lemma előtt) :

$$\|W_\ell^{1/2}(A_\ell^{-1} - P_{\ell-1}^\ell A_{\ell-1}^{-1} R_\ell^{\ell-1})W_\ell^{1/2}\| \leq c_a. \quad (15.282)$$

**15.41. Lemma** (Braess, Hackbusch; többrácsos iterációs mátrix tulajdonságai, 2). Teljesüljön (15.276)-(15.262) és (15.282), legyen  $\nu_1 = \nu_2 = \nu/2$  és

$$0 \leq \widetilde{M}_{\ell-1} \leq \zeta_{\ell-1} I_{\ell-1}, \quad \zeta_{\ell-1} \in [0, 1]. \quad (15.283)$$

Ekkor

$$0 \leq \widetilde{M}_\ell \leq \zeta_\ell I_\ell, \quad \zeta_\ell := \min_{\eta \in [\zeta_{\ell-1}^\gamma, 1]} \max_{\xi \in [0, 1]} f(\xi, \eta), \quad (15.284)$$

ahol

$$f(\xi, \eta) := (1 - \xi)^\nu [\eta + (1 - \eta)c_a \xi].$$

**Bizonyítás.** Feltételeink mellett érvényes  $0 \leq \widetilde{M}_\ell$  és (15.280). Vezessünk le  $\widetilde{Q}_\ell$ -ről, (15.279) mellett és (15.282) alapján egy további felső becslést! Mivel  $\widetilde{Q}_\ell$  szimmetrikus és pozitív szemidefinit, így

$$W_\ell^{1/2}(A_\ell^{-1} - P_{\ell-1}^\ell A_{\ell-1}^{-1} R_\ell^{\ell-1})W_\ell^{1/2} = W_\ell^{1/2} A_\ell^{-1/2} \widetilde{Q}_\ell A_\ell^{-1/2} W_\ell^{1/2}$$

is az, tehát (15.282)-ből következik

$$0 \leq W_\ell^{1/2} A_\ell^{-1/2} \widetilde{Q}_\ell A_\ell^{-1/2} W_\ell^{1/2} \leq c_a I_\ell.$$



Ezzel ekvivalens

$$0 \leq \tilde{Q}_\ell \leq c_a A_\ell^{1/2} W_\ell^{-1} A_\ell^{1/2} = c_a X_\ell. \quad (15.285)$$

Kombinálva a (15.279) és (15.285) egyenlőtlenségeket, megkapjuk a kívánt becslést

$$0 \leq \tilde{Q}_\ell \leq \alpha c_a X_\ell + (1 - \alpha) I_\ell, \quad \alpha \in [0, 1], \quad (15.286)$$

alakban, így nyerve egy szabad paramétert. Összevetve a (15.278) és (15.281) relációkat, írhatunk

$$\begin{aligned} \tilde{M}_\ell &= \tilde{S}_\ell^{\nu/2} \left[ \tilde{Q}_\ell + \tilde{P}_{\ell-1}^\ell (\tilde{M}_{\ell-1})^\gamma \tilde{R}_\ell^{\ell-1} \right] \tilde{S}_\ell^{\nu/2} \\ &\leq \tilde{S}_\ell^{\nu/2} \left[ \tilde{Q}_\ell + \tilde{P}_{\ell-1}^\ell \zeta_{\ell-1}^\gamma \tilde{R}_\ell^{\ell-1} \right] \tilde{S}_\ell^{\nu/2} \\ &= \tilde{S}_\ell^{\nu/2} \left[ \tilde{Q}_\ell + \zeta_{\ell-1}^\gamma (I_\ell - \tilde{Q}_\ell) \right] \tilde{S}_\ell^{\nu/2}. \end{aligned} \quad (15.287)$$

Itt felhasználtuk a (15.283) feltételt és  $\tilde{Q}_\ell$  definícióját.

(15.287)-ből következik (15.286) segítségével

$$0 \leq \tilde{M}_\ell \leq \tilde{S}_\ell^{\nu/2} \left[ (1 - \zeta_{\ell-1}^\gamma) (\alpha c_a X_\ell + (1 - \alpha) I_\ell) + \zeta_{\ell-1}^\gamma I_\ell \right] \tilde{S}_\ell^{\nu/2}. \quad (15.288)$$

Itt a kapcsos zárójelben  $I_\ell$  szorzója összesen

$$\eta := (1 - \zeta_{\ell-1}^\gamma)(1 - \alpha) + \zeta_{\ell-1}^\gamma = 1 - \alpha + \alpha \zeta_{\ell-1}^\gamma \in [\zeta_{\ell-1}^\gamma, 1],$$

és ekkor

$$(1 - \zeta_{\ell-1}^\gamma) \alpha = 1 - \eta.$$

Ezzel (15.288) felírható mint

$$\tilde{M}_\ell \leq \tilde{S}_\ell^{\nu/2} [(1 - \eta) c_a X_\ell + \eta I_\ell] \tilde{S}_\ell^{\nu/2}.$$

Most vegyük észre, hogy  $\tilde{S}_\ell = I_\ell - X_\ell$  miatt ezen egyenlőtlenség jobb oldalán mind a három szorzó felcserélhető, tehát

$$\tilde{M}_\ell \leq (I_\ell - X_\ell)^\nu [\eta I_\ell + (1 - \eta) c_a X_\ell] = f(X_\ell, \eta).$$

Innen és  $0 < X_\ell \leq I_\ell$ -ből következik a lemma állítása.  $\square$

**15.42. Tétel** (Braess, Hackbusch; szimmetrikus többrácsos módszer konvergenciája). A 15.41. lemma feltételei mellett a többrácsos módszer konvergens az  $A_\ell$  által generált energetikai normában : a  $\nu \geq 1$ ,  $\gamma \geq 1$  esetben érvényes

$$\|M_\ell(\nu)\|_{A_\ell} \leq \frac{c_a}{\nu + c_a}.$$

**Bizonyítás.** Mivel

$$\|M_\ell(\nu)\|_{A_\ell} = \|A_\ell^{1/2} M_\ell(\nu) A_\ell^{-1/2}\| = \|\widetilde{M}_\ell(\nu)\|,$$

így – figyelembe véve a 15.41. lemma eredményeit – elegendő  $\zeta_\ell$ -re  $\ell$ -től független felső becslést levezetni. Kiindulhatunk abból, hogy  $\zeta_0 = 0$  és  $\zeta_{\ell-1} \leq c_a/(\nu + c_a) =: \eta^*$ . Ekkor  $\eta^* \in [\zeta_{\ell-1}^\gamma, 1]$ , és így (15.284) szerint

$$\zeta_\ell \leq \max_{\xi \in [0,1]} f(\xi, \eta^*).$$

Itt

$$f(\xi, \eta^*) = \eta^*(1 - \xi)^\nu(1 + \nu\xi)$$

felveszi maximumát  $\xi = 0$ -nál, értéke ott éppen  $\eta^*$ , amivel lezártuk az indukciós bizonyítást.  $\square$

**Megjegyzés.** Ezzel beláttuk a konvergenciát a szimmetrikus esetben, akár  $V$ - akár  $W$ -ciklusról lenne szó. A bizonyítás szerint konvergensek olyan többrácsos eljárások is, amelyeknél  $\gamma \geq 1$  és  $\gamma$  függhet  $\ell$ -től.  $\square$

Ehhez a tételhez ld. a 66. feladatot is.

Most vizsgáljuk a  $\nu_1 \neq \nu_2$  esetet.

**15.43. Tétel** (Braess, Hackbusch;  $V$ -ciklus konvergenciája szimmetrikus esetben). A 15.41. lemma feltételei mellett (de megengedve  $\nu_1 \neq \nu_2$ -t, ha  $\nu_1 + \nu_2 \geq 1$  és  $\gamma = 1$ ) teljesül

$$\|M_\ell(\nu_1, \nu_2)\|_{A_\ell} \leq \left( \frac{c_a}{\nu_1 + c_a} \cdot \frac{c_a}{\nu_2 + c_a} \right)^{1/2} < 1.$$

**Bizonyítás.** Az 15.40. lemma c) azonosságának köszönhetően

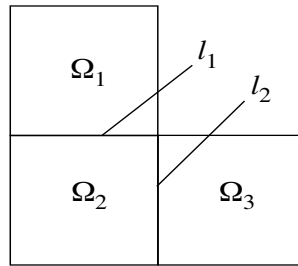
$$\begin{aligned} \|M_\ell(\nu_1, \nu_2)\|_{A_\ell}^2 &= \|\widetilde{M}_\ell(\nu_1, \nu_2)\|^2 = \|\widetilde{M}_\ell(0, \nu_2)\widetilde{M}_\ell(\nu_1, 0)\|^2 \\ &\leq \|\widetilde{M}_\ell(0, \nu_2)\|^2 \|\widetilde{M}_\ell(\nu_1, 0)\|^2. \end{aligned}$$

Használjuk most azt, hogy az euklideszi normában és tetszőleges  $A$  mátrix esetén (ld. a 67. feladatot)

$$\|A\|^2 = \lambda_{\max}(A^T A) = \lambda_{\max}(A A^T) = \|A^T A\| = \|A A^T\|.$$

Hivatkozva a 15.40. lemma b) és c) azonosságára, valamint a 15.42. tételre is, felírhatjuk, hogy

$$\begin{aligned} \|\widetilde{M}_\ell(\nu_1, 0)\|^2 &= \|(\widetilde{M}_\ell(\nu_1, 0))^T \widetilde{M}_\ell(\nu_1, 0)\| = \|\widetilde{M}_\ell(0, \nu_1)\widetilde{M}_\ell(\nu_1, 0)\| \\ &= \|\widetilde{M}_\ell(\nu_1, \nu_1)\| \leq \frac{c_a}{\nu_1 + c_a}, \end{aligned}$$



ábra 15.34:  $L$ -alakú tartomány

és hasonlóan

$$\|\widetilde{M}_\ell(0, \nu_2)\|^2 = \|\widetilde{M}_\ell(0, \nu_2)(\widetilde{M}_\ell(0, \nu_2))^T\| = \|\widetilde{M}_\ell(\nu_2, \nu_2)\| \leq \frac{c_a}{\nu_2 + c_a},$$

amivel az állítást beláttuk.  $\square$

**Megjegyzés.** A 15.39. és 15.42. tételek fontos feltétele (15.282), amely teljesül, ha alkalmazható a 15.30. tétel. Anélkül is meg lehet mutatni a konvergenciát, de akkor csak olyan eredmény ismeretes, mint

$$\|M_\ell\|_{A_\ell} \leq q < 1,$$

ahol a  $q$  konstans független  $\ell$ -től. Viszont a 15.39. és 15.42. tétel szerinti gyors konvergencia ( $\nu_1$ -et és  $\nu_2$ -t megfelelően nagyoknak választva) nem következik.  $\square$

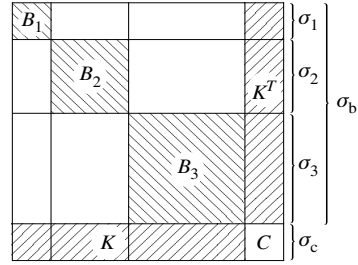
#### 15.8.4 Tartomány dekompozíció és párhuzamos módszerek, bevezetés

Ebben a pontban és a következő kettőben a diszkretizált egyenleteknek olyan megoldási módszereivel foglalkozunk, amelyeket nagy, összetett tartományok esetén célszerű használni. Egy idevonatkozó, klasszikus feladat a repülőgép testében fellépő mechanikai feszültségek, ill. sajátfrekvenciák számítása. Ekkor célszerű a törzs és a sok figyelmet kívánó szárnyak szétválasztása. Példának tekintjük a kétdimenziós  $L$ -alakú  $\Omega$  tartományt a 15.34. ábrán.

Az  $l_1$ ,  $l_2$  vonalak segítségével a tartományt három részre osztottuk.  $\Omega$ -ban megoldandó az

$$\mathcal{L}u = f, \quad x \in \Omega; \quad u|_\Gamma = 0 \quad (15.289)$$

alakú önadjungált elliptikus peremértékfeladat. Erre alkalmazunk véges differencia vagy végeselem eljárást úgy, hogy az  $l_1$ ,  $l_2$  vonalak a rácsvonalakhoz tartozzanak. Célunk az, hogy az  $\Omega_1$ ,  $\Omega_2$ ,  $\Omega_3$  résztartományokban függetlenül egymástól megoldható feladatokra bontsuk (15.289)-et.



ábra 15.35: Alstruktúrák: blokkmátrix

A nagyméretű feladatok szétbontása kisebb részfeladatokra gyakran előnyös. Ezenkívül az egyes részfeladatok függetlensége a többprocesszoros gépeken lehetséges párhuzamos számítást is megalapozhatja.

Az elliptikus peremértékfeladatok jellemző vonása viszont az, hogy a megoldás a tartomány belső pontjában az összes peremértéktől függ. Ez teszi problematikusá az ötlet megvalósíthatóságát.

### 15.8.5 Alstruktúra-eljárás

Először is a Gauss-elimináció segítségével megmutatjuk, hogy mégis bizonyos mértékig (és már ez is előnyös) szétbontható a (15.289)-hez tartozó diszkrét feladat, amit

$$Ay = g, \quad y, g \in \mathbb{R}^N, \quad (15.290)$$

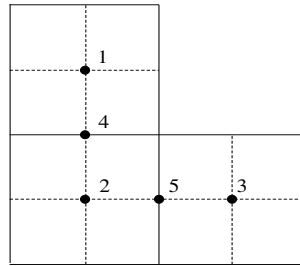
alakban írunk fel. Feltehetjük, hogy  $A$  nagyméretű ritka, szimmetrikus és pozitív definit mátrix. Az  $y$  vektor első komponensei legyenek azok, amelyek belső pontokhoz tartoznak. Tegyük fel, hogy ebből  $N_b$  darab van. Majd ezután következzenek az összecsatoló vonalakhoz tartozó ismeretlenek, amelyeknek száma legyen  $N_c = N - N_b$ . Ekkor (15.290) blokkalakja

$$Ay = g, \quad A = \begin{pmatrix} B & K^T \\ K & C \end{pmatrix}, \quad y = \begin{pmatrix} y_b \\ y_c \end{pmatrix}, \quad g = \begin{pmatrix} g_b \\ g_c \end{pmatrix}. \quad (15.291)$$

Mivel az  $\Omega_\ell$  „alstruktúrák” ( $1 \leq \ell \leq L$ ) csak az elválasztó vonalakon keresztül csatlakoznak egymáshoz, de közvetlenül nem, így a  $B$  mátrix blokk-diagonális,

$$B = \text{diag}_{1 \leq \ell \leq L}(B_\ell).$$

Példánk esetében  $L = 3$ , és ekkor a 15.35. ábra a blokkalakat mutatja (fehér területek : nullák). Ott  $\sigma_\ell$ -lrel jelöltük az  $\Omega_\ell$  belső rácspontjaihoz tartozó ismeretlenek indexeit az  $y$  vektorban ( $\sigma_b := \cup_{1 \leq \ell \leq L} \sigma_\ell$ ), míg  $\sigma_c$  tartalmazza a vonalakon fekvő ismeretlenek indexeit. Ezen halmazok elemeinek számát  $|\cdot|$ -vel jelöljük. Így  $N_b = \sum_{\ell=1}^L |\sigma_\ell|$  és  $N_c = |\sigma_c|$ .



ábra 15.36:  $L$ -alakú tartomány diszkretizációja

Legyen például (15.289) a Poisson-egyenlet és (15.290) annak diszkretizációja a szokásos ötpontos differenciasémával (vagy derékszögű háromszöges lineáris végelemekkel), az  $\Omega_\ell$  alstruktúrák legyenek egységnégyzetek, a lépéstávolság  $h = 1/2$ . Ekkor minden  $\Omega_\ell$ -hez csak egy belső ismeretlen tartozik, összesen 5 ismeretlen van (15.36. ábra).

Így  $N_b = 3$ ,  $N_c = 2$ ,  $\sigma_\ell = \{\ell\}$ ,  $\ell = 1, 2, 3$ ;  $\sigma_b = \{1, 2, 3\}$ ,  $\sigma_c = \{4, 5\}$ , és  $A$  felírható mint

$$A = \begin{pmatrix} 4 & 0 & 0 & -1 & 0 \\ 0 & 4 & 0 & -1 & -1 \\ 0 & 0 & 4 & 0 & -1 \\ -1 & -1 & 0 & 4 & 0 \\ 0 & -1 & -1 & 0 & 4 \end{pmatrix}, \quad K = \begin{pmatrix} -1 & -1 & 0 \\ 0 & -1 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}.$$

tehát  $B_\ell = 4$ ,  $\ell = 1, 2, 3$ ,

Most kezdjük a Gauss-eliminációt, és elsőnek a belső ismeretleneket kiküszöböljük.  $A$ -ra nézve ez a következő LDU-blokkfelbontással írható fel (ld. I. 1.3.2, a jobboldali vektor megfelelő módon változik meg. Az  $A$  mátrix szimmetriáját figyelmen kívül hagyhatjuk – ekkor is  $U = L^T$ ):

$$A = \begin{pmatrix} I_b & 0 \\ KB^{-1} & I_c \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & S_c \end{pmatrix} \begin{pmatrix} I_b & B^{-1}K^T \\ 0 & I_c \end{pmatrix}. \quad (15.292)$$

Itt  $I_b$ , illetve  $I_c$  az  $N_b$ -, ill  $N_c$ -dimenziójú egységmátrix és  $S_c$  az összecsatoló vonalak ismeretlenjeihez tartozó Schur-féle komplementer (ld. I. 1.3.2):

$$S_c = C - KB^{-1}K^T.$$

Mivel a  $B$  mátrix  $B_\ell$  blokkjai és  $K$  között nullblokkok vannak, a belső ismeretlenek eliminációja minden résztartományban független a többi belső ismeretlenek eliminációjától. Így ez akár egymásután, akár párhuzamosan vagy különböző gépeken történhet, csak az egyes eredményekből az  $S_c$  Schur-féle komplementert kell összeállítani. Ennek a műveletnek a képleteit lent

részletezzük. Az  $S_c$ -vel kapcsolatos (jobban kondicionált, kisméretű de általában telt mátrixú) egyenletrendszer megoldása után, az összekapcsolódó vonalakhoz tartozó ismeretlenek birtokában az egyes alstruktúrák részére külön-külön végezhetjük a visszahelyettesítést.

A Schur-féle komplementer előállításához először is emlékeztetünk, hogy az a Gauss-elimináció során képződik, nem kell hozzá a  $B^{-1}$  inverz mátrixot kiszámítani. A belső ismeretlenek eliminációja alatt a  $C$  mátrix helyén az

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - \frac{a_{ik}^{(k)} a_{kj}^{(k)}}{a_{kk}^{(k)}}, \quad i, j \in \sigma_c, \quad 1 \leq k \leq N_b - 1,$$

képlet szerint változnak az elemek (v.ö. I. 1.3.1.). A  $C \rightarrow S_c$  átmenet (vagyis az előző folyamat  $k = 1$ -től  $k = N_b$ -ig) összefoglalható mint

$$\begin{aligned} a_{ij}(S_c) &= a_{ij}(C) - \sum_{k=1}^{N_b} \frac{a_{ik}^{(k)} a_{kj}^{(k)}}{a_{kk}^{(k)}} \\ &= a_{ij}(C) - \sum_{\ell=1}^L \sum_{k \in \sigma_\ell} \frac{a_{ik}^{(k)} a_{kj}^{(k)}}{a_{kk}^{(k)}} = a_{ij}(C) + \sum_{\ell=1}^L a_{ij}(\sigma_\ell), \end{aligned}$$

vagyis mátrixalakban

$$S_c = C + \sum_{\ell=1}^L S(\sigma_\ell).$$

Itt  $S(\sigma_\ell)$ -lel jelöltük az  $\ell$ -edik alstruktúra ismeretleneinek az eliminációja során a  $C$  blokk helyén képződő mátrixokat – ha az elimináció elején ott a nullmátrix volt. Az egyes alstruktúrák eliminációjának függetlensége miatt ezek az  $S(\sigma_\ell)$  mátrixok külön-külön elkészíthetők.

Speciálisan a 15.36. ábra példájának esetén az alstruktúra eljárás a következőképpen folyik :

1)  $\sigma_1$ -indexű ismeretlenek (azaz  $x_1$ ) eliminációja,  $S(\sigma_1)$  kiszámítása :

$$\text{tehát } S(\sigma_1) = -\frac{1}{4} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{és } g_4 \rightarrow g_4 + \frac{1}{4}g_1;$$

2)  $\sigma_2$ -indexű ismeretlenek (példánkban  $x_2$ ) eliminációja,  $S(\sigma_2)$  kiszámítása :

$$S(\sigma_2) = -\frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad g_4 \rightarrow g_4 + \frac{1}{4}g_2, \quad g_5 \rightarrow g_5 + \frac{1}{4}g_2;$$

3)  $\sigma_3$ -indexű ismeretlenek (azaz  $x_3$ ) eliminációja :

$$S(\sigma_3) = -\frac{1}{4} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad g_5 \rightarrow g_5 + \frac{1}{4}g_3.$$

$$\begin{array}{|c|c|c|c|c|} \hline 4 & 0 & 0 & -1 & 0 \\ \hline 0 & 4 & 0 & -1 & -1 \\ \hline 0 & 0 & 4 & 0 & -1 \\ \hline -1 & -1 & 0 & 0 & 0 \\ \hline 0 & -1 & -1 & 0 & 0 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|c|c|} \hline 4 & 0 & 0 & -1 & 0 \\ \hline 0 & 4 & 0 & -1 & -1 \\ \hline 0 & 0 & 4 & 0 & -1 \\ \hline 0 & -1 & 0 & -1/4 & 0 \\ \hline 0 & -1 & -1 & 0 & 0 \\ \hline \end{array}, \quad \begin{array}{|c|c|c|c|c|} \hline 4 & 0 & 0 & -1 & 0 \\ \hline 0 & 4 & 0 & -1 & -1 \\ \hline 0 & 0 & 4 & 0 & -1 \\ \hline -1 & -1 & 0 & 0 & 0 \\ \hline 0 & -1 & -1 & 0 & 0 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|c|c|} \hline 4 & 0 & 0 & -1 & 0 \\ \hline 0 & 4 & 0 & -1 & -1 \\ \hline 0 & 0 & 4 & 0 & -1 \\ \hline -1 & 0 & 0 & -1/4 & -1/4 \\ \hline 0 & 0 & -1 & -1/4 & -1/4 \\ \hline \end{array}$$

ábra 15.37:  $\sigma_1$ -, ill.  $\sigma_2$ -indexű ismeretlenek eliminációja

$$\begin{array}{|c|c|c|c|c|} \hline 4 & 0 & 0 & -1 & 0 \\ \hline 0 & 4 & 0 & -1 & -1 \\ \hline 0 & 0 & 4 & 0 & -1 \\ \hline -1 & -1 & 0 & 0 & 0 \\ \hline 0 & -1 & -1 & 0 & 0 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|c|c|} \hline 4 & 0 & 0 & -1 & 0 \\ \hline 0 & 4 & 0 & -1 & -1 \\ \hline 0 & 0 & 4 & 0 & -1 \\ \hline 0 & -1 & 0 & 0 & 0 \\ \hline 0 & -1 & 0 & 0 & -1/4 \\ \hline \end{array}$$

ábra 15.38:  $\sigma_3$ -indexű ismeretlenek eliminációja

Összefoglalva:

$$\sum_{\ell=1}^3 S(\sigma_\ell) = -\frac{1}{4} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad S_c = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} - \frac{1}{4} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 14 & -1 \\ -1 & 14 \end{pmatrix},$$

– és valóban,

$$\begin{aligned} S_c = C - KB^{-1}K^T &= \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} - \begin{pmatrix} -1 & -1 & 0 \\ 0 & -1 & -1 \end{pmatrix} \frac{1}{4} \begin{pmatrix} -1 & 0 \\ -1 & -1 \\ 0 & -1 \end{pmatrix} \\ &= \frac{1}{4} \begin{pmatrix} 14 & -1 \\ -1 & 14 \end{pmatrix}. \end{aligned}$$

4) Az  $S_c$ -vel kapcsolatos kis rendszernek a megoldása után a kapott  $y_4, y_5$  értékeket behelyettesítjük a  $\sigma_\ell$ -indexű egyenletekbe és ekkor onnan, függetlenül egymástól, kiszámíthatjuk az  $y_1, y_2, y_3$  komponenseket.

Az  $S(\sigma_\ell)$  mátrixok kiszámításakor nyilván nem kell, hogy  $C$  helyén legyen az elején a nullmátrix : ha szekvenciálisan dolgozunk, akkor az  $S(\sigma_\ell)$  előállításakor  $C$  helyén legyen  $C + \sum_{k=1}^{\ell-1} S(\sigma_k)$ , ha párhuzamosan dolgozunk, akkor – amelyik processzor éppen elkészült a maga  $S(\sigma_\ell)$  mátrixának (vagy akár

annak egy elemének) a kiszámításával – az adja hozzá a  $C$  helyén éppen álló mátrixhoz.

Az alstruktúra-eljárás előnye, hogy egyes lépései függetlenek és az egész mátrix csak kis részének jelenlétét tételezi fel (így az egész mátrixot nem is állítjuk össze), összességében nem kell több tárhely, mint ha a teljes mátrixot összeállítottuk volna (a blokkdiagonális körül álló nullblokkokat nyilván ekkor sem tároljuk), eltérés csak a  $C$  mátrix helyén adódhat.

Az eljárás hátrányos akkor, amikor igen nagymértékű feladat esetén vagy az egyes alstruktúrák vagy pedig a kapott  $S_c$  eliminációja már tárgondokat okoz, ill. ezeknek során indokolt volna iterációs módszerekhez folyamodnunk. Ekkor különösen az okoz problémát, hogy  $S_c$  eléggé telt mátrix.

Így két feladattal állunk szemben :

a) az eljárást úgy kell megváltoztatnunk, hogy a  $B$  mátrix blokkjaival kapcsolatos rendszereket csak közelítőleg oldjuk meg (pl. többrácsos módszerrel);

b) az  $S_c$  kiszámítását, tárolását és invertálását (azaz  $LDU$ -felbontását) megkerülve olyan prekondicionálási mátrixot kell javasolni helyette, amely könnyen kiszámítható és ritka, és amely az  $S_c$ -t jól közelíti meg (prekondicionáló értelmében, ld. I., 166. o.).

Az a) és b) feladatok megoldásával  $B$  és  $S_c$  helyett olyan  $P_b$ , ill.  $Q_c$  mátrixot nyerünk, amelyet mint prekondicionálási mátrixot alkalmazhatunk a prekondicionált konjugált gradiens módszer keretében.

Emlékeztetünk rá (ld. I. 146. oldalon az (1.158) képletet és a módszer algoritmusát), hogy a prekondicionáló mátrix az a szimmetrikus, pozitív definit mátrix, amellyel valóban egzaktul meg kell oldanunk egy egyenletrendszert a prekondicionált konjugált gradiens módszer minden lépésében.

### 15.8.6 Dirichlet tartomány dekompozíciós módszer

Az előző pont végén említett a) és b) feladatokat a következőképpen oldjuk meg :

a) A  $B_\ell$  mátrix helyett azt a  $P_\ell$  mátrixot vesszük ( $\ell = 1, \dots, L$ ), amely a többrácsos módszer  $m$  lépéséből adódik, nulla közelítésből kiindulva az

$$\mathcal{L}u = f, \quad x \in \Omega_\ell; \quad u = 0, \quad x \in \partial\Omega_\ell,$$

Dirichlet-feladat

$$B_\ell y_\ell = g_\ell \tag{15.293}$$

alakú diszkrét egyenleteit közelítőleg megoldva. Itt is a teljes feladat megoldásvektorát a (15.291) blokkalokban írjuk fel,  $y_b = (y_1, \dots, y_L)^T$ .



b) Az  $S_c$  Schur-féle komplementer prekondicionálási mátrixaként használjuk az úgynevezett *Dryja-féle prekondicionálási mátrixot* (lényegében, egy konstans szorzót nem számítva, a  $-y_{\bar{x}}$  differenciahányadoshoz tartozó tridiagonális mátrixnak a négyzetgyökéről van szó).

c) A teljes  $A$  mátrix számára ezután a következő  $P$  prekondicionálási mátrixot vesszük (v.ö. (15.292)-vel) :

$$P := \begin{pmatrix} I_b & 0 \\ KP_b^{-1} & I_c \end{pmatrix} \begin{pmatrix} P_b & 0 \\ 0 & Q_c \end{pmatrix} \begin{pmatrix} I_b & P_b^{-1}K^T \\ 0 & I_c \end{pmatrix}, \quad (15.294)$$

ahol

$$P_b := \text{diag}_{1 \leq \ell \leq L}(P_\ell). \quad (15.295)$$

Mielőtt  $P$ -t összevetjük az eredeti  $A$  mátrixszal (ld. (15.292)), a  $P_\ell$  és  $Q_c$  mátrixokat még közelebbről megvizsgáljuk.

$P_\ell$ -re érvényes a következő képlet :

$$P_\ell = P_\ell(m) := B_\ell(I_\ell - M_\ell^m)^{-1} \quad (15.296)$$

ahol  $I_\ell$  a  $|\sigma_\ell|$ -méretű egységmátrix és  $M_\ell$  a (15.293)-ra alkalmazott többrácsos módszer iterációs mátrixa. Ugyanis ezekkel a jelölésekkel a többrácsos módszernek egy iterációja felírható a következő alakban :

$$y_\ell^{(m)} = M_\ell y_\ell^{(m-1)} + (I_\ell - M_\ell)B_\ell^{-1}g_\ell \quad (15.297)$$

(v.ö. 15.4.8-ban (15.112)-vel és (15.113)-mal, valamint az ott tett megjegyzéssel), a hiba egyenlete

$$y_\ell^{(m)} - y_\ell = M_\ell(y_\ell^{(m-1)} - y_\ell) = M_\ell^m(y_\ell^{(0)} - y_\ell),$$

és mivel az iteráció kezdeti közelítéseként  $y_\ell^{(0)} = 0$ -t vesszük, így

$$y_\ell^{(m)} = y_\ell - M_\ell^m y_\ell = (I_\ell - M_\ell^m)B_\ell^{-1}g_\ell.$$

Ez azt jelenti, hogy az  $m$  iterációs lépés segítségével a

$$P_\ell y_\ell^{(m)} = B_\ell(I_\ell - M_\ell^m)^{-1}y_\ell^{(m)} = g_\ell$$

alakú rendszert oldottuk meg pontosan.

Szükségünk lesz arra, hogy  $P_\ell$  önadjungált legyen a  $B_\ell$  által definiált skalárszorzatban, mert akkor  $P_b$  szimmetrikus és pozitív definit.

Ehhez olyan restriktiót kell alkalmaznunk, amely az interpoláció transzponáltja (mint mátrix), és ugyanannyi Gauss–Seidel utósimítást kell, hogy

elvégezzünk, mint elősimítást. Ezenkívül itt az utósimításnál a megfordított sorrendben iteráljuk a pontokat.

Feltehetjük, hogy  $M_\ell$  normája a diszkretizációtól függetlenül kisebb egynél (ld. a 15.41. lemmát).

Legyen  $x, y \in \mathbb{R}^{|\sigma_\ell|}$  és  $(x, y)_{B_\ell} := (B_\ell x, y)$  a szimmetrikus és pozitív definit  $B_\ell$  mátrix által definiált skalárszorzat, valamint  $\|x\|_{B_\ell} := (B_\ell x, x)^{1/2}$  a hozzátartozó norma.

**15.44. Lemma** (U. Langer és társai; önadjungált többrácsos operátorok spektrális ekvivalenciája). Ha  $M_\ell$  önadjungált a  $B_\ell$ -skalárszorzatban és érvényes

$$\|M_\ell\|_{B_\ell} \leq \eta < 1,$$

akkor  $P_\ell = B_\ell(I_\ell - M_\ell^m)^{-1}$  jól definiált, szimmetrikus, pozitív definit, és spektrálisan ekvivalens  $B_\ell$ -l-el:

$$(1 - \eta^m)P_\ell \leq B_\ell \leq (1 + \eta^m)P_\ell.$$

[Kommentár. A 15.40. lemma és a 15.42. tétel biztosítják az  $M_\ell$  itt követelt tulajdonságait.]

**Bizonyítás.** Első állításunk következik a perturbációs lemmából (I. 1.2.), amely szerint  $I_\ell - M_\ell$  és  $I_\ell - M_\ell^m$  reguláris mátrixok (és érvényes  $\|(I_\ell - M_\ell^m)^{-1}\|_{B_\ell} \leq 1/(1 - \eta^m)$ ). Az  $M_\ell$  mátrix önadjungált a  $B_\ell$ -skalárszorzatban, ha

$$(M_\ell y, v)_{B_\ell} = (B_\ell M_\ell y, v) = (y, M_\ell v)_{B_\ell} = (B_\ell y, M_\ell v) = (y, B_\ell M_\ell v)$$

minden  $|\sigma_\ell|$ -dimenziójú  $y$  és  $v$  vektorra, vagyis ha  $(B_\ell M_\ell)^T = B_\ell M_\ell$ , azaz ha

$$B_\ell^{-1} M_\ell^T = M_\ell B_\ell^{-1}, \quad (15.298)$$

ill. még egy ekvivalens alakban : ha  $B_\ell^{1/2} M_\ell B_\ell^{-1/2}$  szimmetrikus mátrix (ld. a 15.40. lemmát).

Ekkor  $P_\ell$  szimmetrikus. Lássuk ugyanis be, hogy  $P_\ell^{-1}$  szimmetrikus :

$$P_\ell^{-1} = (I_\ell - M_\ell^m) B_\ell^{-1} = B_\ell^{-1} - B_\ell^{-1} (M_\ell^T)^m = B_\ell^{-1} (I_\ell - (M_\ell^T)^m) = P_\ell^{-T}.$$

Itt használtuk (15.298)-at, pontosabban az ebből következő  $B_\ell^{-1} (M_\ell^T)^m = M_\ell^m B_\ell^{-1}$  egyenlőséget és a  $(P_\ell^T)^{-1} = P_\ell^{-T}$  jelölést. Ezután az euklideszi skalárszorzatban érvényes

$$(P_\ell^{-1} y, y) = ((I_\ell - M_\ell^m) B_\ell^{-1} y, y) = \|y\|_{B_\ell^{-1}}^2 - (B_\ell^{1/2} M_\ell^m B_\ell^{-1/2} v, v),$$

ahol  $v := B_\ell^{-1/2}y$ , és itt

$$|(B_\ell^{1/2}M_\ell^m B_\ell^{-1/2}v, v)| \leq \|B_\ell^{1/2}M_\ell^m B_\ell^{-1/2}\| \|v\|^2 = \|M_\ell^m\|_{B_\ell} \|y\|_{B_\ell^{-1}}^2.$$

Innen és az előző összefüggésből azt kapjuk, hogy

$$(1 - \eta^m)(B_\ell^{-1}y, y) \leq (P_\ell^{-1}y, y) \leq (1 + \eta^m)(B_\ell^{-1}y, y),$$

vagyis  $B_\ell^{-1}$  és  $P_\ell^{-1}$  spektrálisan ekvivalensek :

$$(1 - \eta^m)B_\ell^{-1} \leq P_\ell^{-1} \leq (1 + \eta^m)B_\ell^{-1}.$$

Állításunk most onnan következik, hogy szimmetrikus, pozitív definit  $A$  és  $B$  mátrixok esetén

$$0 < A \leq B \iff 0 < B^{-1} \leq A^{-1}. \quad \square$$

**Megjegyzés.** Amennyiben ismert, hogy

$$0 \leq (M_\ell y, y)_{B_\ell} \leq \eta \|y\|_{B_\ell}^2, \quad \eta < 1,$$

vagyis ha a lemma feltételein kívül  $M_\ell$  még nemnegatív definit is a  $B_\ell$ -skalárszorzatban, akkor következik

$$(1 - \eta^m)P_\ell \leq B_\ell \leq P_\ell.$$

Az  $M_\ell \geq 0$  feltétel nélkül is kaphatjuk ezt az eredményt, ha  $m$  páros. Ugyanis a  $B_\ell^{1/2}M_\ell^m B_\ell^{-1/2}$  mátrix (15.298) miatt szimmetrikus, így valóságosak a sajátértékei – és azok  $(\lambda(M_\ell))^m$ -nel egyeznek meg, tehát páros  $m$  esetén nemnegatívak.  $\square$

Most foglalkozunk a Dryja-féle prekondicionáló mátrixszal, amely akkor működik hatékonyan, ha az egyes elválasztó vonalaknak nincsenek közös belső pontjaik. Csak ezt az egyszerűbb esetet fogjuk vizsgálni. Legyen összesen  $J$  ilyen vonal, és a  $j$ -edik vonalra illeszkedő pontok száma  $m_j$ . (A vonalak végpontjai  $\Gamma$ -n vannak és nem tartoznak ide.) Akkor a Dryja-féle prekondicionáló mátrix

$$Q_c := \text{diag}_{1 \leq j \leq J}(Q_j),$$

ahol

$$Q_j := 2A_j^{1/2} \quad \text{és} \quad A_j := \text{tridiag}(-1, 2, -1)$$

$m_j \times m_j$ -méretű mátrixok. Tudjuk, hogy az  $A_j$  szimmetrikus és pozitív definit mátrix, így egyértelműen meghatározott a szimmetrikus és pozitív definit négyzetgyöke. Ez utóbbit nem érdemes kiszámítani, ugyanis  $A_j$  sajátértékei

és sajátvektorai ismertek (I. 3.1.1., (3.6)). Ezzel ismert az  $U_j$  unitér mátrix is, amely  $A_j$ -t a  $\Lambda_j$  diagonális alakjára hozza :

$$A_j = U_j \Lambda_j U_j^T.$$

Ekkor

$$Q_j = 2A_j^{1/2} = 2U_j \Lambda_j^{1/2} U_j^T,$$

és a  $Q_j$ -vel kapcsolatos,  $Q_j y = g$  alakú rendszereket a következőképpen oldjuk meg :

a) kiszámítjuk a  $q := U_j^T g$  szorzatot :

$$q_\ell = \sqrt{2} \sum_{k=1}^{m_j} \sin \frac{\ell k \pi}{m_j + 1} g_k, \quad \ell = 1, 2, \dots, m_j.$$

Ahogy látjuk, ez semmi más, mint egy diszkrét sinus-transzformáció és így  $O(m_j \ln m_j)$  művelettel a megfelelő gyors Fourier-algoritmus segítségével elintézhető (ld. I. 4.7. Előnyös, ha  $m_j + 1$  kettőhatvány).

b) A kapott  $q$  vektort megszorozzuk a  $\Lambda_j^{-1/2}$  diagonális mátrixszal :  $x := \Lambda_j^{-1/2} q$ .

c) Az  $x$  vektort visszatranszformálva megkapjuk a keresett  $y$  vektort :

$$y_k = \sqrt{2} \sum_{\ell=1}^{m_j} \sin \frac{\ell k \pi}{m_j + 1} x_\ell.$$

Ez még egyszer  $O(m_j \ln m_j)$  művelet. Ily módon – annak ellenére, hogy mindegyik  $Q_j$  telt mátrix – olcsóbban jutunk hozzá a  $Q_c$ -mátrixú prekondicionáló egyenletrendszerek megoldásához mint LU-felbontás segítségével!

Egy újabb megközelítést ehhez a problémához ld. Kiss B. és Krebsz A. dolgozatában. Ott – nagyságrend szerint egyenlő műveletigény mellett – még a gyors Fourier-transzformáció sem kell (ami azzal az előnnyel is jár, hogy a pontszám eltérhet kettőhatványtól).

A  $P_\ell$  és  $Q_j$  mátrixok invertálásának ismeretében megtárgyalhatjuk a prekondicionált konjugált gradiens algoritmus (ld. I. 1.6.8.) egy iterációjának a végrehajtását.

1) Az iteráció során a  $P$  prekondicionálási mátrixszal, ld. (15.294), megoldandó a

$$P \begin{pmatrix} y_b \\ y_c \end{pmatrix} = \begin{pmatrix} r_b \\ r_c \end{pmatrix}$$

alakú egyenletrendszer. (15.294) szerint ez azt jelenti, hogy megoldandó

$$\begin{pmatrix} P_b & 0 \\ 0 & Q_c \end{pmatrix} \begin{pmatrix} z_b \\ z_c \end{pmatrix} = \begin{pmatrix} w_b \\ w_c \end{pmatrix},$$

ahol

$$\begin{pmatrix} z_b \\ z_c \end{pmatrix} := \begin{pmatrix} I_b & P_b^{-1}K^T \\ 0 & I_c \end{pmatrix} \begin{pmatrix} y_b \\ y_c \end{pmatrix} = \begin{pmatrix} y_b + P_b^{-1}K^T y_c \\ y_c \end{pmatrix}$$

és

$$\begin{pmatrix} w_b \\ w_c \end{pmatrix} := \begin{pmatrix} I_b & 0 \\ KP_b^{-1} & I_c \end{pmatrix}^{-1} \begin{pmatrix} r_b \\ r_c \end{pmatrix} = \begin{pmatrix} r_b \\ r_c - KP_b^{-1}r_b \end{pmatrix}.$$

Ez végül is a következőképpen kivitelezhető egymás után:

$$\begin{aligned} P_b z_b &= r_b \text{ megoldása, majd } w_c := r_c - K z_b, \\ Q_c y_c &= w_c \text{ megoldása, majd } v_b := r_b - K^T y_c, \\ P_b y_b &= v_b \text{ megoldása.} \end{aligned}$$

A  $P_b = \text{diag}(P_\ell)$  és  $Q_c = \text{diag}(Q_j)$  mátrixok invertálása viszont a  $P_\ell$ , ill.  $Q_j$  mátrixok invertálását jelenti, amit fent megbeszéltük, és ami ( $\ell$ -re, ill.  $j$ -re nézve) párhuzamosan végrehajtható.

2) A konjugált gradiens algoritmusban előforduló skalárszorzatok azon részösszegeit számíthatjuk ki párhuzamosan, amelyek az egyes  $\bar{\Omega}_\ell$  résztartományok ismeretleneivel kapcsolatosak. A belső ismeretlenek esetén ( $y_b$ ) ez világos, de az elválasztó vonalakon fekvő ismeretlenekkel ( $y_c$ ) kapcsolatos járulékokat a globális mátrix elemeihez feloszthatjuk aszerint, hogy melyik  $\bar{\Omega}_\ell$ -ből származnak. (Ez természetes módon úgy adódik, hogy az elválasztó vonalak egyben végelemeknek a peremvonalai, és a globális mátrixot  $\Omega$ -ra nem is állítjuk össze, hanem csak minden  $\bar{\Omega}_\ell$ -ra külön-külön.) Ezután a párhuzamosan előállítható részösszegek kiszámítása után viszont szükség van globális információcserére: vagy minden processzor közli a maga részösszegét egy kitüntetett processzorral, amely azokat  $\ell$  felett összegzi, vagy pedig minden processzor kapja az összes információt és összegzi. Ezen lépés befejezésekképpen még a prekondicionált konjugált gradiens algoritmus  $\lambda$ ,  $\mu$  számait is kiszámítja.

3) A  $\lambda$ ,  $\mu$  számok birtokában az iteráció új vektorait megint párhuzamosan számíthatjuk ki.

Az algoritmus leírása után most a konvergencia szempontjából döntő kérdéssel foglalkozunk, a  $P$  és az  $A$  mátrixok spektrális ekvivalenciájával (ld. I. 1.6.8., a prekondicionált konjugált gradiens módszer algoritmusát után), amit  $A \sim P$  alakban jelöljük.

**15.45. Tétel** (U. Langer és társai; Dirichlet-tartomány dekompozíciós prekondicioner becslése). Legyen  $A$ , ill.  $P$  a (15.291), ill. (15.294) mátrix prekondicionálási mátrixa. Teljesüljön

$$\underline{\gamma}_b P_b \leq B \leq \bar{\gamma}_b P_b, \quad \underline{\gamma}_c Q_c \leq S_c + T_c \leq \bar{\gamma}_c Q_c, \quad (15.299)$$

végül legyen

$$\rho := \rho(S_c^{-1}T_c), \quad T_c := K(B^{-1} - P_b^{-1})B(B^{-1} - P_b^{-1})K^T.$$

Akkor  $A \sim P$  :

$$\underline{\gamma} \left(1 - \sqrt{\frac{\rho}{1+\rho}}\right) P \leq A \leq \bar{\gamma} \left(1 + \sqrt{\frac{\rho}{1+\rho}}\right) P,$$

ahol

$$\underline{\gamma} := \min(\underline{\gamma}_b, \underline{\gamma}_c), \quad \bar{\gamma} := \max(\bar{\gamma}_b, \bar{\gamma}_c).$$

**Bizonyítás.** Az  $A$ , ill.  $P$  mátrixot a következő alakban írjuk fel (v.ö. (15.292)-vel) :

$$A = V \operatorname{diag}(B, S_c) V^T, \quad P = W \operatorname{diag}(P_b, Q_c) W^T,$$

ahol

$$V := \begin{pmatrix} I_b & 0 \\ KB^{-1} & I_c \end{pmatrix}, \quad W := \begin{pmatrix} I_b & 0 \\ KP_b^{-1} & I_c \end{pmatrix}.$$

Ezek a (reguláris) mátrixok a spektrális ekvivalencia bizonyítását csak azért nehezítik meg, mert különböznek. Egyébként, ha pl.  $A$  és  $B$  két tetszőleges szimmetrikus mátrix, akkor  $A \sim B \Leftrightarrow VAV^T \sim VBV^T$ . Ennek belátásához elegendő a

$$c_0(Ay, y) \leq (By, y) \leq c_1(Ay, y), \quad 0 < c_0 \leq c_1,$$

egyenlőtlenségekbe  $y = V^T u$ -t behelyettesíteni.

Emiatt a következőképpen járunk el :

- Kiszámítjuk az  $\tilde{A} := W^{-1}AW^{-T}$  mátrixot;
  - megmutatjuk :  $\tilde{A}$  spektrálisan ekvivalens diagonális részével,  $\tilde{D}$ -vel;
  - ezután belátjuk, hogy  $\tilde{D} \sim \tilde{P} := \operatorname{diag}(P_b, Q_c)$ ;
  - ekkor  $\tilde{A} \sim \tilde{P}$ ,  $A \sim W\tilde{P}W^T = P$  következik.
- a) Az  $\tilde{A}$  mátrix kiszámítása :

$$\begin{aligned} \tilde{A} &= \begin{pmatrix} I_b & 0 \\ -KP_b^{-1} & I_c \end{pmatrix} \begin{pmatrix} B & K^T \\ K & C \end{pmatrix} \begin{pmatrix} I_b & -P_b^{-1}K^T \\ 0 & I_c \end{pmatrix} \\ &= \begin{pmatrix} B & (I_b - BP_b^{-1})K^T \\ K(I_b - P_b^{-1}B) & X_c \end{pmatrix}, \end{aligned}$$

$$X_c := S_c + T_c = (C - KB^{-1}K^T) + K(B^{-1} - P_b^{-1})B(B^{-1} - P_b^{-1})K^T.$$

b)  $\tilde{A}$  és  $\text{diag}(B, X_c) =: \tilde{D}$  ekvivalenciájához megjegyezzük, hogy  $A$ -val együtt  $\tilde{A}$  és  $B$  is szimmetrikusak és pozitív definiték, és így az  $X_c = [\tilde{A}/B]$  Schur-féle komplementer is az (ld. I. 1.3.2; 1.4. lemma). Ennek alapján  $\tilde{D}$  is szimmetrikus és pozitív definit, így az

$$\tilde{A}x = \lambda\tilde{D}x$$

általánosított sajátérték feladat minden sajátértéke pozitív (ld. I. 3.7.2). Ezt a feladatot most egy

$$S_c^{-1}T_c x = \nu x \quad (15.300)$$

alakú feladatra vezetjük vissza (ehhez ld. a 68. feladatot is). Mivel

$$\tilde{D}^{-1}\tilde{A} = I + \begin{pmatrix} 0 & (B^{-1} - P_b^{-1})K^T \\ X_c^{-1}K(B^{-1} - P_b^{-1})B & 0 \end{pmatrix},$$

így, ha  $\tilde{D}^{-1}\tilde{A}x = \lambda x$  és  $x = (x_b, x_c)^T$ , akkor

$$\begin{aligned} (B^{-1} - P_b^{-1})K^T x_c &= (\lambda - 1)x_b, \\ X_c^{-1}K(B^{-1} - P_b^{-1})B x_b &= (\lambda - 1)x_c. \end{aligned}$$

Innen következik, hogy

$$(\lambda - 1)^2 x_c = X_c^{-1}K(B^{-1} - P_b^{-1})B(B^{-1} - P_b^{-1})K^T x_c = X_c^{-1}T_c x_c.$$

Ha most

$$T_c x_c = \nu S_c x_c = \nu(X_c - T_c)x_c,$$

akkor

$$X_c^{-1}T_c x_c = \frac{\nu}{1 + \nu} x_c,$$

és így  $(\lambda - 1)^2 = \nu/(1 + \nu)$ , vagyis

$$\lambda_{\pm} = 1 \pm \sqrt{\frac{\nu}{1 + \nu}}.$$

A  $2 \times 2$ -es blokkaloknak megfelelően kaptunk 2 sajátértéket  $S_c^{-1}T_c$  minden egyes  $\nu$  sajátértékéhez.  $\rho = \nu_{\max} = \rho(S_c^{-1}T_c)$ -vel következik a

$$\lambda_{\min}(\tilde{D}y, y) \leq (\tilde{A}y, y) \leq \lambda_{\max}(\tilde{D}y, y)$$

egyenlőtlenségekből, hogy

$$\left(1 - \sqrt{\frac{\rho}{1 + \rho}}\right) \tilde{D} \leq \tilde{A} \leq \left(1 + \sqrt{\frac{\rho}{1 + \rho}}\right) \tilde{D}. \quad (15.301)$$

c)  $\tilde{D} \sim \tilde{P} = \text{diag}(P_b, Q_c)$  érvényes, hiszen

$$\begin{aligned} (\tilde{D}y, y) &= \left( \tilde{D} \begin{pmatrix} y_b \\ y_c \end{pmatrix}, \begin{pmatrix} y_b \\ y_c \end{pmatrix} \right) = (By_b, y_b) + (X_c y_c, y_c) \\ &\leq \bar{\gamma}_b(P_b y_b, y_b) + \bar{\gamma}_c(Q_c y_c, y_c) \leq \bar{\gamma}(\tilde{P}y, y), \end{aligned}$$

(15.299) segítségével. Pontosan így vezethető le az alsó becslés is, tehát

$$\underline{\gamma}\tilde{P} \leq \tilde{D} \leq \bar{\gamma}\tilde{P}. \quad (15.302)$$

d) Figyelembe véve a (15.301) és (15.302) becsléseket, következik

$$\underline{\gamma} \left( 1 - \sqrt{\frac{\rho}{1+\rho}} \right) (\tilde{P}y, y) \leq (\tilde{A}y, y) \leq \bar{\gamma} \left( 1 + \sqrt{\frac{\rho}{1+\rho}} \right) (\tilde{P}y, y),$$

és innen,  $y = W^T x$  behelyettesítésével, kapjuk a tétel állítását.  $\square$

**Megjegyzés.** A 15.44. lemmából következik, hogy a többrácsos módszer  $m$  lépésszámának növelésével a tétel  $\underline{\gamma}_b$  és  $\bar{\gamma}_b$  konstansait tetszőlegesen közel hozhatjuk az 1 értékhez (a gyakorlati számításnál viszont kettő lépésnél többet ne vegyünk).  $\square$

Bizonyítás nélkül említjük a következő eredményt, amelyből a tétel  $\underline{\gamma}_c$  és  $\bar{\gamma}_c$  konstansok becslése nyerhető.

**15.46. Lemma** (Dryja; a Dryja-féle prekondicionálási mátrix becslése). A Dryja-féle  $Q_c$  prekondicionálási mátrix esetén érvényes

$$\gamma_0 Q_c \leq S_c \leq \gamma_1 Q_c, \quad 0 < \gamma_0 \leq \gamma_1,$$

ahol  $\gamma_0$  és  $\gamma_1$  nem függenek  $h$ -től, a diszkretizáció lépéstávolságától. Továbbá

$$h\gamma_2 I_c \leq S_c \leq \gamma_3 I_c, \quad \gamma_4 I_c \leq B \leq \gamma_5 I_c, \quad 0 < \gamma_2, \gamma_4, \quad (15.303)$$

ahol a  $\gamma$  konstansok függetlenek  $h$ -től.  $\square$

Mivel  $T_c$  nemnegatív definit (ld. a 68. feladatot), így az előző lemmából következik

$$\gamma_0 Q_c \leq S_c \leq S_c + T_c = S_c(I_c + S_c^{-1}T_c) \leq (1 + \rho)S_c,$$

tehát a tétel konstansai

$$\underline{\gamma}_c = \gamma_0, \quad \bar{\gamma}_c = (1 + \rho)\gamma_1.$$



Mint utolsó kérdés maradt ezután  $\rho$  becslése. Ez a szám a (15.300) sajátérték feladat maximális sajátértéke, tehát

$$\rho = \sup_{0 \neq x \in \mathbb{R}^{N_c}} \frac{(T_c x, x)}{(S_c x, x)}. \quad (15.304)$$

A  $P_b, M, B$  mátrixokra vonatkozó (15.295), (15.296) és (15.298) relációk segítségével kapjuk, hogy

$$B^{-1} - P_b^{-1} = B^{-1} - P_b^{-T} = M^m B^{-1} = B^{-1} (M^T)^m,$$

ahol  $M := \text{diag}_{1 \leq \ell \leq L} (M_\ell)$ . Így, figyelembe véve a  $T_c$  mátrix 15.45. tételbeli definícióját és mégegyszer (15.298)-at, következik

$$\begin{aligned} T_c &= K B^{-1} (M^m) B M^m B^{-1} K^T = R_c^T R_c, \\ R_c &:= B^{1/2} M^m B^{-1} K^T = (B^{1/2} M^m B^{-1/2}) B^{-1/2} K^T. \end{aligned}$$

Ennek alapján

$$\begin{aligned} (T_c x, x) &= \|R_c x\|^2 \leq \|M^m\|_B^2 \|B^{-1/2} K^T x\|^2 \\ &\leq \eta^{2m} (K B^{-1} K^T x, x), \end{aligned}$$

ld. a 15.44. lemmát, és mivel  $S_c = C - K B^{-1} K^T$ , így

$$(T_c x, x) \leq \eta^{2m} [(C x, x) - (S_c x, x)].$$

Most (15.304) és (15.303) adja a következő becslést:

$$\rho \leq \eta^{2m} \left[ \sup_{0 \neq x \in \mathbb{R}^{N_c}} \frac{(C x, x)}{(S_c x, x)} - 1 \right] \leq \eta^{2m} \left( \frac{\gamma_5}{\gamma_2 h} - 1 \right).$$

Eszerint ahhoz, hogy a  $\rho \leq \delta = \text{const} < 1$  feltételt biztosítsuk, elegendő, ha az  $m$  lépésszámot az

$$m = \left\lceil \frac{\delta \gamma_2}{2 \gamma_5} \log \frac{1}{h} \Big/ \log \frac{1}{\eta} \right\rceil = O \left( \log \frac{1}{h} \right) \quad (15.305)$$

nagyságrendben választjuk. Numerikus kísérletek szerint viszont elegendő  $m = O(1)$  is.

Mivel a preconditionált konjugált gradiens módszer hibája úgy csökken, mint

$$q := (\sqrt{\kappa} - 1) / (\sqrt{\kappa} + 1), \quad \text{ahol } \kappa = \text{cond}(P^{-1}A),$$

és mivel tételünk szerint

$$\kappa \leq (\bar{\gamma}/\underline{\gamma}) \left(1 + \sqrt{\frac{\rho}{1+\rho}}\right) / \left(1 - \sqrt{\frac{\rho}{1+\rho}}\right) = (\bar{\gamma}/\underline{\gamma})(1+\rho) \left(1 + \sqrt{\frac{\rho}{1+\rho}}\right)^2$$

– így (15.305) esetén  $m$  igen, de  $q$  nem függ  $h$ -től : a konvergencia sebessége nem romlik a diszkretizáció finomításával.

Befejezésül a fenti párhuzamos módszerrel kapcsolatos néhány további ötletet, ill. általánosítást sorolunk fel.

1) A most megtárgyalt eset, hogy az elválasztó vonalak nem kereszteződnek, esetleg oda vezet, hogy az egyes processzorok nagyon egyenlőtlen módon vannak leterhelve. Ekkor a  $Q_c$  Dryja-féle prekondicionálási mátrix helyett a Bramble–Pasciak–Schatz-félét (ld. az irodalomjegyzéket) kell használni. Háromdimenziós esetben is ez javasolható, de itt is (ha az elválasztó felületek nem metszik egymást) előnyös a Kiss B., Krebsz A. által konstruált eljárás.

2) A  $P_b$  prekondicionálási mátrix (15.294)-ben három helyen szerepel. A fenti elméletet alig kell megváltoztatni, ha a (15.294) szorzat kettő külső blokk-háromszöges mátrixában más többrácsos módszert alkalmazunk, mint a blokk-diagonális mátrixban, akár nemszimmetrikus eljárás is megengedhető.

3) A  $P$  mátrixhoz tartozó blokk-háromszöges mátrixok egy transzformációt hajtanak végre a végeelem bázison. Ilyen szempontból az a kérdés vetődik fel, hogy itt (ahol éppen az elválasztó vonalakhoz és a belső ismeretlenekhez tartozó információ kicserélődik) miért kellene a belső ismeretlenekhez tartozó prekondicionálás esetén nulla közelítésből indítani a többrácsos módszert? Célszerűbb az elválasztó vonalakon meglévő információt olcsó, de hatékony módon a résztartományok belsejébe kiterjeszteni.

## 15.9 Elliptikus sajátérték feladatainak numerikus megoldása

Már a 3. fejezetben foglalkoztunk az algebrai sajátérték feladatok numerikus megoldásával, majd 11.6.4-ben röviden a közönséges differenciálegyenletek sajátérték feladatával.

A parciális differenciálegyenletekkel kapcsolatban érdemes visszatérni erre a témára, mert jelentős alkalmazások tartoznak ide: gépek és épületek, valamint azoknak alkotóelemei (lemezek, héjak), vagy akár molekulák rezgéseiről is lehet szó; a sajátértékek a frekvenciák négyzetét, a sajátfüggvények az alaprezgések alakját adják meg. Numerikus szempontból itt

1. igen nagyméretű feladatok keletkeznek;
2. többszörös sajátértékek elég gyakran előfordulnak;
3. a diszkretizált feladatok sajátvektorai approximálják a megfelelő sajátfüggvényeket, amelyeknek simasági tulajdonságait (a leginkább szükséges, alacsony rendszámú sajátfüggvényekre gondolunk itt) tiszta algebrai módon eljárva nem használhatjuk ki.

A többszörös sajátértékek miatt, amelyek a tartomány szimmetriái jelenléte esetén keletkeznek, javasoljuk a vektoriterációs módszerek (I. 3.4.) blokkváltozatait, kombinálva a Ritz-módszerrel (ld. I. 2.2. végén a megjegyzést, 3.9-ben a 18. feladatot, valamint lejjebb a 15.9.4. pontot). Az 1-3. problémák hatékonyabb kezelésére viszont a többrácsos módszer azon verzióját ajánljuk, amelyre a 15.9.3. pont végén hivatkozunk. Ez bonyolultabb ugyan, de többszörös (ill. igen közeli) sajátértékek esetén is kitűnően működik.

Tárgyalásunk elején felidézzük az elméleti tudnivalókat.

### 15.9.1 Elméleti háttér

Az alapvető elméleti eredmény ezen a területen nem az

$$\mathcal{L}u = \lambda u, \quad x \in \Omega; \quad Bu = 0, \quad x \in \Gamma \quad (15.306)$$

feladatra vonatkozik (ahol  $\mathcal{L}$  tipikusan páros rendű elliptikus operátor és  $B$  a peremfeltételek operátora), hanem az ebből létrehozható variációs feladatra :

Keressünk olyan  $u \in V$ -t, hogy

$$a(u, v) = \lambda(u, v)_0 \quad \text{minden } v \in V\text{-re.} \quad (15.307)$$

Itt  $V \subset L_2(\Omega) \subset V'$  komplex értékű függvények Hilbert-terei. Ezekről lejjebb még további feltételeket sorolunk, de  $V$  esetén mindenképp  $H_0^1(\Omega)$ -ra gondolunk; mint mindig,  $(u, v)_0$  jelöli az  $L_2$  tér skalárszorzatát, amelynek képlete most

$$(u, v)_0 := \int_{\Omega} u(x) \overline{v(x)} \, dx.$$

Továbbá,  $a(u, v)$  folytonos bilineáris – nem feltétlenül szimmetrikus – funkcionál  $V \times V$ -n; a norma  $V$ -ben  $\|u\|_V$ .

A sajátérték feladatok esetén nem célszerű a bilineáris funkcionál elliptikusságát megkövetelni. Ugyanis nem szabad, hogy nehézséget okozzon az, ha a (15.306) sajátérték feladat helyett az  $\mathcal{L}u - \alpha u = \lambda u$  egyenletre vonatkozó feladatot konstans  $\alpha$ -val vizsgáljuk, amelynek bilineáris formája

$$\bar{a}(u, v) := a(u, v) - \alpha(u, v)_0.$$

Ha most  $a$  elliptikus,

$$a(u, u) \geq m_a \|u\|_V^2, \quad (15.308)$$

akkor  $\bar{a}$ -re teljesül

$$\bar{a}(u, u) \geq m_a \|u\|_V^2 - \alpha \|u\|_0^2,$$

és ez elég nagy  $\alpha$  esetén már biztosan nem  $V$ -elliptikus: elegendő olyan nem-nulla  $u^*$  függvényt behelyettesíteni  $\bar{a}$ -be, amellyel (15.308) teljesül egyenlőséggel : ekkor  $\bar{a}(u^*, u^*) < 0$ , mihelyt  $\alpha > m_a \|u^*\|_V^2 / \|u^*\|_0^2$ .

**Definíció.** Az  $a$  folytonos bilineáris funkcionált akkor hívjuk *koercívnek*  $V$ -ben (avagy  $V$ -koercívnek), ha léteznek olyan  $c_V > 0$ ,  $c_0 \geq 0$  konstansok, amelyekkel teljesül

$$a(u, u) \geq c_V \|u\|_V^2 - c_0 \|u\|_0^2. \quad \square \quad (15.309)$$

A (15.307) feladatban keresett az  $u \in V$  sajátfüggvény és a  $\lambda \in \mathbb{C}$  sajátérték (amelynek konjugált komplex értékét  $\bar{\lambda}$ -tal jelöljük).

Ezt a feladatot érdemes az *adjungált sajátérték feladattal* együtt vizsgálni, amelynek variációs megfogalmazása így néz ki:

Keresett  $u \in V$  és  $\lambda \in \mathbb{C}$  úgy, hogy

$$a(v, u) = \bar{\lambda}(v, u)_0 \quad \text{minden } v \in V\text{-re.} \quad (15.310)$$

A (15.307), ill. (15.310) feladatok  $\lambda$  sajátértékeihez tartozó  $u(\lambda)$ , ill.  $u^*(\lambda)$  sajátfüggvények  $V$ -ben altereket feszítenek ki, amelyeket  $S(\lambda)$ -val, ill.  $S^*(\lambda)$ -val jelöljük.

**15.47. Tétel** (Riesz–Schauder elmélet). A  $V \hookrightarrow L_2(\Omega)$  beágyazás legyen folytonos, sűrű és kompakt, az  $a$  bilineáris funkcionál legyen folytonos és  $V$ -koercív. Ekkor

1. a (15.307) és (15.310) feladatokhoz léteznek megoldások; ha  $\lambda$  sajátértéke (15.307)-nek, akkor (15.310)-nek is – és megfordítva;
2. a (15.307) és (15.310) feladatoknak megszámlálhatóan sok sajátértéke van, egyetlen torlódási pontjuk a  $+\infty$ ;
3. ha  $\lambda$  sajátérték, akkor

$$1 \leq \dim S(\lambda) = \dim S^*(\lambda) < \infty,$$

tehát azonosságú lineárisan független sajátfüggvénye van – külön-külön – a (15.307) és (15.310) feladatoknak egy adott sajátértékhez;

4. amennyiben a bilineáris funkcionál szimmetrikus (ami igaz, ha (15.306) önadjungált), akkor a sajátértékek valósak, egyetlen torlódási pontjuk  $+\infty$ , és (15.307) és (15.310) megfelelő sajátalterei megegyeznek :  
 $S(\lambda) = S^*(\lambda)$   $\square$ .

**Megjegyzés.** A tétel lényeges feltétele a folytonos, sűrű és *kompakt* beágyazás  $V \hookrightarrow L_2(\Omega)$ . Ezért felsorolunk néhány kritériumot, amely ezt biztosítja (ld. pl. Hackbusch „elliptikus” könyvét) :

1. Legyen  $\Omega \subset \mathbb{R}^n$  korlátos, legyen  $s$  és  $t$  valós. Ekkor a  $H_0^s(\Omega) \hookrightarrow H_0^t(\Omega)$  Szoboljev–terek közti beágyazás folytonos, sűrű és kompakt, ha  $s > t$ .

2. Legyen  $\Omega$  mint 1. alatt, de ezenkívül  $\Gamma$  pereme legyen szakaszonként folytonosan differenciálható, sarkainak belső szöge kisebb  $2\pi$ -nél; a  $k$  és  $\ell$  egész számok legyenek nemnegatívak. Ekkor a  $H^k(\Omega) \hookrightarrow H^\ell(\Omega)$  beágyazás  $k > \ell$  esetén folytonos, sűrű és kompakt.

3. A  $C^s(\overline{\Omega}) \hookrightarrow C^t(\overline{\Omega})$  beágyazás folytonos, sűrű és kompakt, ha  $\Omega$  olyan, mint 1. alatt és  $s > t$ .

Ha ezek a kritériumok nem teljesülnek, akkor leginkább azzal lehet számolni, hogy a megszámlálható spektrum mellett folytonos spektrumrész is van.  $\square$

A (15.307) variációs feladat szokásos végeelem diszkretizációjával eljutunk a következő végesdimenziójú feladathoz, amelyben  $V_h$  a  $V$  tér végesdimenziós altere :

Keressünk olyan  $\lambda^h \in \mathbb{C}$ -t és  $u_h = u_h(\lambda^h) \in V_h \subset V$ -t, hogy

$$a(u_h, v_h) = \lambda^h (u_h, v_h)_0 \quad \text{minden } v_h \in V_h\text{-ra.} \quad (15.311)$$

Hasonló a (15.310) adjungált feladat diszkrét alakja :

Keresett  $\lambda^h \in \mathbb{C}$  és  $u_h^* = u_h^*(\lambda^h) \in V_h$  úgy, hogy

$$a(v_h, u_h^*) = \overline{\lambda^h} (v_h, u_h^*)_0 \quad \text{minden } v_h \in V_h\text{-ra.} \quad (15.312)$$

A  $V_h$  végeelem tér rendelkezzen a következő approximációs tulajdonsággal (v.ö. 15.7.5-tel):

$$\inf_{u_h \in V_h} \|u - u_h\|_V \rightarrow 0, \quad h \rightarrow 0, \quad \text{minden } u \in V\text{-re.} \quad (15.313)$$

Megfelelő módosítással a Riesz–Schauder tétel  $V_h$ -ban is alkalmazható, ami azt jelenti, hogy itt is a végeelem módszer előnyös olyan szempontból, hogy nem igényel új elméletet a diszkretizált feladat megoldhatóságáról.

A különbség az, hogy ekkor csak  $N := \dim V_h$  darab sajátértéket és „diszkrét” (azaz  $V_h$ -beli) sajátfüggvényt kapunk. A sajátértékek véges számából

következik, hogy az összes sajátérték egyenletes approximációjára nem gondolhatunk, csupán a (15.307), ill. (15.310) feladat rögzített  $\lambda$  sajátértékére kérdezhetjük, vajon (15.311)-nek, ill. (15.312)-nek olyan  $\lambda^h$  megoldásai vannak-e, hogy  $\lambda^h \rightarrow \lambda$ , amikor  $h \rightarrow 0$ . Ennek során nehezítő körülményt jelent az, hogy  $\lambda$ -nak esetleg 1-nél nagyobb multiplicitása van. Emiatt csak olyan konvergencia eredményt említünk, amely egyszeres sajátértékre vonatkozik (a többszeres sajátértékek esetét ld. Babuška és Osborn cikkében).

**15.48. Tétel** (egyszeres sajátérték végeelem approximációjának konvergenciája). Teljesüljenek a 15.47. tétel feltételei, valamint (15.313). Legyen  $\lambda$  a (15.307) feladat egyszeres sajátértéke :  $\dim S(\lambda) = 1$ . Legyen  $u(\lambda)$  a hozzátartozó sajátfüggvény és  $u^*(\lambda)$  a (15.310) adjungált feladat sajátfüggvénye. Végül, teljesüljön

$$\|u(\lambda)\|_V = 1, \quad (u(\lambda), u^*(\lambda))_0 = 1. \quad (15.314)$$

Ekkor a (15.311), ill. (15.312) diszkrét feladatoknak vannak a  $(\lambda^h, u_h(\lambda^h))$ , ill.  $(\lambda^h, u_h^*(\lambda^h))$  megoldásai. Ha  $h$  elég kicsi, akkor  $\lambda^h$  is egyszeres sajátérték. Ekkor vannak olyan  $c_0, c_1$  konstansok, amelyekkel teljesül

$$\begin{aligned} \|u(\lambda) - u_h(\lambda^h)\|_V &\leq c_0 \inf_{v_h \in V_h} \|u(\lambda) - v_h\|_V, \\ \|u^*(\lambda) - u_h^*(\lambda^h)\|_V &\leq c_0 \inf_{v_h \in V_h} \|u^*(\lambda) - v_h\|_V, \\ |\lambda - \lambda^h| &\leq c_1 \inf_{v_h \in V_h} \|u(\lambda) - v_h\|_V \inf_{v_h \in V_h} \|u^*(\lambda) - v_h\|_V. \end{aligned} \quad (15.315)$$

Vagyis: az approximáció kvázioptimális.  $\square$

**Megjegyzések. 1.** Nemönadjungált feladat esetén a (15.314) feltétel második része nemcsak normalizáció kérdése, hanem azt is biztosítja, hogy a  $\lambda$  sajátérték mértani és algebrai multiplicitása egybeesik. (A mértani multiplicitás adja a hozzátartozó sajátvektorok számát :  $\dim(A - \lambda I)$ ; az algebrai multiplicitás viszont  $\dim \bigcup_{k=1}^{\infty} (A - \lambda I)^k$  és így nagyobb lehet a mértaninál. Ha a kettő egybeesik, akkor tehát  $\dim(A - \lambda I) = \dim(A - \lambda I)^2$ . Ez a tulajdonság önadjungált feladatoknál automatikusan megvan, és ekkor  $u(\lambda) = u^*(\lambda)$ .)

**2.** Amennyiben  $\mathcal{L}$  önadjungált, másodrendű operátor,  $V = H_0^1(\Omega)$  vagy  $V = H^1(\Omega)$ , és  $\Omega$  és az  $a(u, v)$  olyanok, hogy a (15.307) feladat  $u(\lambda)$  megoldásai  $H^2(\Omega)$ -ban vannak, akkor a lineáris háromszöges  $V_h$  végeelem tér esetén léteznek (15.311)-nek és (15.312)-nek olyan  $(\lambda^h, u_h(\lambda^h))$ ,  $(\lambda^h, u_h^*(\lambda^h))$  megoldásai, hogy

$$\|u(\lambda) - u_h(\lambda^h)\|_V = O(h), \quad \|u(\lambda) - u_h(\lambda^h)\|_0 = O(h^2),$$

$$\|u^*(\lambda) - u_h^*(\lambda^h)\|_V = O(h), \quad \|u^*(\lambda) - u_h^*(\lambda^h)\|_0 = O(h^2),$$

és  $|\lambda - \lambda^h| = O(h^2)$ .

**3.** A feladat még nagyobb simasága esetén gyorsabb konvergencia elérhető vagy Richardson extrapoláció vagy magasabbrendű elemek segítségével.

**4.** Ezek az eredmények többszörös sajátértékek esetén is igazak; a „diszkrét” sajátérték multiplicitása viszont nem feltétlenül a megfelelő pontos sajátérték multiplicitására áll be, még elegendő kicsi  $h$ -ra sem. Az optimális konvergencia eredményeket ld. Chatelin, ill. Osborn és Babuška dolgozataiban.

**5.** Önadjungált feladat esetén a  $\lambda_k^h$  „diszkrét” sajátértékek fentről közelítik meg a  $\lambda_k$  „folytonos” sajátértékeket: a sajátértékek  $k$  sorszámától függetlenül van olyan konstans  $\delta$ , hogy

$$\lambda_k \leq \lambda_k^h \leq \lambda_k + \delta h^2 \lambda_k^2, \quad \text{amikor } h \leq h(k). \quad \square \quad (15.316)$$

Ilyen egyenlőtlenségekre II. 11.6.4-ben adtunk konkrét példákat egydimenziós differenciálegyenletek sajátérték feladatai kapcsán. Ott a Courant–Fischer tételt is megfogalmaztuk, amelyből a  $\lambda_k \leq \lambda_k^h$  alsó becslések következnek. A következő pontban egy kétdimenziós sajátérték feladat és végeselem diszkrétizációja példáján adjuk meg a (15.316) egyenlőtlenségeknek konkrét alakját.

### 15.9.2 A diszkrét Laplace–operátor sajátérték feladata

Ebben a pontban vizsgáljuk a Laplace–operátor sajátértékeit, legtöbbször abban a speciális esetben, amikor az  $\Omega$  tartomány az egységnyezet és a peremfeltételek elsőfajúak:

$$\Delta u + \lambda u = 0, \quad x = (x_1, x_2) \in \Omega; \quad u(x) = 0, \quad x \in \Gamma. \quad (15.317)$$

Itt  $\Gamma$  az egységnyezet pereme.

1) Ez a feladat egzaktul megoldható, ld. 15.3.3. Így az itteni eredmények hasznosak a numerikus eljárások tesztelésénél. Numerikus szempontból ez nem különösen egyszerű feladat azért, mert itt a sajátértékek igen gyakran többszörösek. Mivel a folytonos feladat sajátértékei  $\lambda_{k,\ell} = \pi^2(k^2 + \ell^2)$ , azért az a kérdés, vajon mekkora a multiplicitása egy adott  $\lambda = n\pi^2$  sajátértéknek, az az ismert számelméleti probléma (ld. Landau könyvét 101. o.), hogy  $n$  hányféleképpen felbontható  $k^2 + \ell^2$  alakú négyzetösszegre, pozitív  $k$  és  $\ell$

számokkal – de megengedve azt is, hogy  $k$ -nak és  $\ell$ -nek közös osztója legyen. Mint példát azt említjük meg, hogy

$$\begin{aligned} 5 &= 1^2 + 2^2, \text{ tehát } \lambda = 5\pi^2 \text{ multiplicitása } 2, \\ 325 &= 1^2 + 18^2 = 6^2 + 17^2 = 10^2 + 15^2, \text{ tehát } \lambda = 325\pi^2 \text{ multiplicitása } 6, \\ 5525 &= 7^2 + 74^2 = 14^2 + 73^2 = 22^2 + 71^2 = 41^2 + 62^2 = 50^2 + 55^2, \\ &\text{ tehát } \lambda = 5525\pi^2 \text{ multiplicitása } 10. \end{aligned}$$

Ld. a 69. -71. feladatokat is.

2) A következő táblázatban az ötpontos differenciaséma néhány sajátértékének közelítő értékét (kerekítés nélkül, két decimális számjeggyel) és multiplicitását adjuk meg az  $h_1 = h_2 = 1/N$  esetben ( $n = (N - 1)^2$  adja a belső rácspontok és a sajátértékek számát). Ezeket a sajátértékeket – amelyek a folytonos feladat sajátértékeit alulról közelítik – megkapjuk a 15.3.3-beli (15.42) képletből, ha ott az  $N_1 = N_2 = N = 1/h$  relációkat behelyettesítjük :

$$\lambda_{k_1 k_2, VD}^h := \frac{4}{h^2} \sin^2 k_1 \pi h_2 + \frac{4}{h^2} \sin^2 \frac{k_2 \pi h}{2}. \quad (15.318)$$

A táblázatban megadott  $k$  sorszám megfelel a „diszkrét” sajátértékek növekedési sorrendjének, így pl. a táblázat ötödik sorában megmutatjuk a 10. „diszkrét” sajátérték értékeit változó  $n$  mellett, valamint a hozzá tartozó  $m$  multiplicitást: ez  $n = 16$  esetén 4, ami ott konkrétan azt jelenti, hogy a 7-10. sajátérték egyenlő. A táblázatból látható, hogy a „diszkrét” sajátértékek multiplicitása lényegesen különbözhet a „folytonos”-étól; ehhez ld. a 72. feladatot is.

$k \setminus n$	1	4	9	16	25	36	49	64	$\infty$
1.	16	18	18.74	19.09	19.29	19.41	19.48	19.53	$2\pi^2 \approx 19.74$
$m$	1	1	1	1	1	1	1	1	1
2.	-	36	41.37	44.09	45.64	46.60	47.23	47.67	$5\pi^2 \approx 49.35$
$m$	-	2	2	2	2	2	2	2	2
4.	-	54	64	69.09	72	73.79	74.98	75.80	$8\pi^2 \approx 78.95$
$m$	-	1	3	1	1	1	1	1	1
10.	-	-	-	100	117.64	129.51	137.74	143.63	$17\pi^2 \approx 167.78$
$m$	-	-	-	4	2	2	2	2	2
13.	-	-	-	130.90	144	156.70	165.49	171.76	$20\pi^2 \approx 197.39$
$m$	-	-	-	1	5	2	2	2	2
16.	-	-	-	180.90	170.35	196	207.01	214.86	$26\pi^2 \approx 256.61$
$m$	-	-	-	1	2	6	2	2	2

Amennyiben a rács nemekvidisztáns (és a diszkretizáció pl. a Shortley–Weller approximáció (15.83) szimmetrikus változata), akkor egyrészt a „diszkrét” sajátértékek már nem feltétlenül alulról közelítik a „folytonosokat”, másrészt a  $h_1$ -nél ill.  $h_2$ -nél kisebb ráctávolságok megjelenésével gyorsan növekszik a legnagyobb sajátérték.

Ld. a 70. feladatot is.

3) Most a végeselem módszerhez térünk át; a következő eredménnyel szemléltetjük a 15.48. tételhez fűzött 2. és 5. megjegyzést (ld. ott a (15.316) becslést) az  $R_4$  elem esetén.



**15.49. Lemma** (bilineáris végeelem approximáció sajátértékei). Legyen  $\lambda_{k,\ell} = (k\pi)^2 + (\ell\pi)^2$  a (15.317) feladat sajátértéke és  $\lambda_{k,\ell}^h$  a 15.7.8. pontban szereplő (15.248) bilineáris végeelem approximáció hozzátartozó sajátértéke. Ekkor

a)  $\lambda_{k,\ell}^h$  értéke a  $\lambda_1 = \lambda_1(k)$ ,  $\lambda_2 = \lambda_2(\ell)$  „egydimenziós” diszkrét sajátértékek segítségével adható meg :

$$\lambda_{k,\ell}^h = \frac{\lambda_1 + \lambda_2 - \frac{h_1^2 + h_2^2}{6} \lambda_1 \lambda_2}{\left(1 - \frac{h_1^2}{6} \lambda_1\right) \left(1 - \frac{h_2^2}{6} \lambda_2\right)}, \quad \lambda_1 := \frac{4}{h_1^2} \sin^2 \frac{k\pi h_1}{2}, \quad \lambda_2 := \frac{4}{h_2^2} \sin^2 \frac{\ell\pi h_2}{2};$$

b)  $h_1, h_2 \leq 2/(\pi \max(k, \ell))$  esetén igaz a következő becslés :

$$\lambda_{k,\ell} \leq \lambda_{k,\ell}^h \leq \lambda_{k,\ell} + \frac{1}{2}(h_1^2 + h_2^2)\lambda_{k,\ell}^2. \quad (15.319)$$

**Bizonyítás.** a) A 15.7.8. pontban megadott (15.248) bilineáris végeelem approximáció (15.244) és (15.247) differencia-képleteibe behelyettesítjük a  $v^{k,\ell}(x) := \sin k\pi x_1 \sin \ell\pi x_2$  rácsfüggvényt. Az  $(Av^{k,\ell})_{ij}$  kifejezésről azonnal következik, hogy az az

$$(Av^{k,\ell})_{ij} = h_1 h_2 \left\{ \lambda_1 + \lambda_2 - \frac{h_1^2 + h_2^2}{6} \lambda_1 \lambda_2 \right\} v_{ij}^{k,\ell} \quad (15.320)$$

alakban írható fel, mert  $\sin k\pi x_1$ , ill.  $\sin \ell\pi x_2$  az

$$y_{\bar{x}_k x_k} + \lambda y = 0, \quad x \in \omega_{h_k}^{(k)}; \quad y(x) = 0, \quad x \in \gamma_{h_k}^{(k)},$$

egydimenzió sajátérték feladatok ( $k = 1, 2$ ) sajátfüggvényei, és a fenti  $\lambda_1$ , ill.  $\lambda_2$  a hozzátartozó ( $k$ -adik, ill.  $\ell$ -edik) sajátérték.

A  $Bv^{k,\ell}$  kifejezés előállításához a következőket számítjuk ki a sin-függvény ismert azonosságainak a segítségével :

$$\begin{aligned} v_{i+1,j}^{k,\ell} + v_{i-1,j}^{k,\ell} + v_{i,j+1}^{k,\ell} + v_{i,j-1}^{k,\ell} &= \\ &= 2(\cos k\pi h_1 + \cos \ell\pi h_2) \sin k\pi x_{1i} \sin \ell\pi x_{2j} \\ &= 2 \left( 1 - 2 \sin^2 \frac{k\pi h_1}{2} + 1 - 2 \sin^2 \frac{\ell\pi h_2}{2} \right) v_{ij}^{k,\ell} \\ &= (4 - h_1^2 \lambda_1 - h_2^2 \lambda_2) v_{ij}^{k,\ell}, \end{aligned}$$

és

$$\begin{aligned} v_{i+1,j+1}^{k,\ell} + v_{i-1,j+1}^{k,\ell} + v_{i-1,j-1}^{k,\ell} + v_{i+1,j-1}^{k,\ell} &= \\ &= 4 \cos k\pi h_1 \cos \ell\pi h_2 \sin k\pi x_{1i} \sin \ell\pi x_{2j} \\ &= 4 \left( 1 - 2 \sin^2 \frac{k\pi h_1}{2} \right) \left( 1 - 2 \sin^2 \frac{\ell\pi h_2}{2} \right) v_{ij}^{k,\ell} \\ &= (2 - h_1^2 \lambda_1)(2 - h_2^2 \lambda_2) v_{ij}^{k,\ell}. \end{aligned}$$

Ennek alapján

$$\begin{aligned} (Bv^{k,\ell})_{ij} &= h_1 h_2 36 (16 + 4(4 - h_1^2 \lambda_1 - h_2^2 \lambda_2) + (2 - h_1^2 \lambda_1)(2 - h_2^2 \lambda_2)) v_{ij}^{k,\ell} \\ &= h_1 h_2 \left(1 - \frac{h_1^2}{6} \lambda_1\right) \left(1 - \frac{h_2^2}{6} \lambda_2\right) v_{ij}^{k,\ell}, \end{aligned}$$

ami (15.320)-szal együtt azt jelenti, hogy  $v^{k,\ell}$  valóban a „diszkrét” sajátfüggvény, és hogy a hozzátartozó sajátértéket éppen a lemma a) képlete adja meg.

b) Mivel  $\lambda_1$ , ill.  $\lambda_2$  a diszkrét egydimenziós sajátérték feladatok sajátértékei, II. 11.6.4-ből tudjuk (ld. ott (246)), hogy ezek a hozzátartozó „folytonos” sajátértékeket alulról közelítik :

$$\lambda_1 < (k\pi)^2 =: p, \quad \lambda_2 < (\ell\pi)^2 =: q. \quad (15.321)$$

Ezenkívül igaz

$$0 < \lambda_1 < \frac{4}{h_1^2}, \quad 0 < \lambda_2 < \frac{4}{h_2^2}.$$

Vezessük be még a következő jelöléseket:  $u := h_1^2/6$ ,  $v := h_2^2/6$ . Ekkor :

$$0 < \lambda_1 u =: x < \frac{2}{3}, \quad 0 < \lambda_2 v =: y < \frac{2}{3}, \quad p+q = \lambda_{k,\ell}, \quad \lambda_1 + \lambda_2 = \lambda_{k,\ell}^h. \quad (15.322)$$

Továbbá azt kapjuk, hogy

$$\lambda_{k,\ell}^h = \frac{\lambda_1 + \lambda_2 - (x\lambda_2 + y\lambda_1)}{(1-x)(1-y)} < (\lambda_1 + \lambda_2)(1+3x)(1+3y). \quad (15.323)$$

Az utolsó összefüggés onnan következik, hogy

$$(1-x)(1+3x) = 1 + x(2-3x) > 1,$$

és hasonlóan  $(1-y)(1+3y) > 1$ . Most folytathatjuk a (15.323) felső becslést (15.321) és (15.322) segítségével :

$$\lambda_{k,\ell}^h < \lambda_{k,\ell}(1 + 3(up + vq) + (3pv)(3qu)).$$

Itt

$$(3pv)(3qu) \leq \frac{1}{2}[(3pv)^2 + (3qu)^2] \leq 3(pv + qu),$$

ha  $3pv = \frac{1}{2}(h_2 k \pi)^2 \leq 2$  és  $3qu = \frac{1}{2}(h_1 \ell \pi)^2 \leq 2$ . Ekkor végül is

$$\lambda_{k,\ell}^h < \lambda_{k,\ell}(1 + 3(u+v)(p+q)) = \lambda_{k,\ell} \left(1 + \frac{1}{2}(h_1^2 + h_2^2)\lambda_{k,\ell}\right).$$

Most rátérünk  $\lambda_{k,\ell}^h$  alsó becsléséhez, amely a  $h_1 \leq 2/(\ell\pi)$ ,  $h_2 \leq 2/(k\pi)$  korlátozások nélkül is igaz. Ennek levezetéséhez kiindulunk a 29. feladatban közölt elemi becslésből :

$$t^2 > \sin^2 t > \frac{t^2}{1 + \alpha t^2}, \quad \text{ha } 0 < t < \frac{\pi}{2}; \quad \alpha := 1 - \frac{4}{\pi^2},$$

amelyet  $\lambda_1$ -re és  $\lambda_2$ -re alkalmazunk :

$$\frac{p}{1 + \beta p u} < \lambda_1 < p, \quad \frac{q}{1 + \beta q v} < \lambda_2 < q. \quad (15.324)$$

Itt  $\beta := \frac{3}{2}\alpha < 1$ . Azt fogjuk megmutatni, hogy

$$uv\lambda_{k,\ell}^h = \frac{vx + uy - (u+v)xy}{(1-x)(1-y)} > \frac{vx}{1-\beta x} + \frac{uy}{1-\beta y}. \quad (15.325)$$

Valóban,  $0 < x, y < \frac{2}{3}$  és

$$\begin{aligned} & (1 - \beta x)(1 - \beta y)(vx + uy - (u+v)xy) \\ & - (1-x)(1-y)[vx(1-\beta y) + uy(1-\beta x)] \\ & = (1-\beta)\{u(1-x)(1-\beta x)y^2 + v(1-y)x^2(1-\beta y)\} > 0 \end{aligned}$$

miatt (15.324)-ből következik

$$\frac{\lambda_1}{1-\beta x} = \frac{\lambda_1}{1-\beta u\lambda_1} > p > \lambda_1, \quad \frac{\lambda_2}{1-\beta y} = \frac{\lambda_2}{1-\beta v\lambda_2} > q > \lambda_2,$$

és innen, valamint (15.325)-ből, kapjuk a kívánt

$$\lambda_{k,\ell}^h > \frac{\lambda_1}{1-\beta x} + \frac{\lambda_2}{1-\beta y} > p + q = \lambda_{k,\ell}$$

eredményt.  $\square$

Hangsúlyozzuk, hogy a végeselem módszer sajátértékei nem az  $A$  merevségi mátrix sajátértékei. A fenti példánkban, bilineáris elem esetén az utóbiakra a lemmából következik a

$$\lambda_{k,\ell}^h(A) = \lambda_1 + \lambda_2 - \frac{h_1^2 + h_2^2}{6}\lambda_1\lambda_2$$

képlet, amely megmutatja, hogy ezek alulról közelítik a pontos sajátértékeket. (15.317)-nél általánosabb differenciáloperátor végeselem mátrixának sajátértékeivel foglalkozik a 63. feladat.

4) A következő táblázatban a  $T_3$  végelem approximáció néhány sajátértékét adunk meg a  $h_1 = h_2 = 1/N$  esetben és a *standard rácson* (az egységnyezetet felosztjuk egyenletesen kis négyzetekre, az átlokat délnyugatról északkeletre húzzuk be); itt is  $n = (N - 1)^2$  adja a belső rácspontok és a sajátértékek számát, a  $k$  sorszám megfelel a „diszkrét” sajátértékek növekedési sorrendjének.

A diszkrét sajátérték feladat pontos megoldása (triviális esetektől eltekintve) ismeretlennek tűnik. Ennek oka a  $B$  mátrixban rejlik (ld. (15.243) – egyébként, ha megfelelően választjuk az egységnyezet felosztását háromszögekre, akkor a végelem sajátérték feladat pontos megoldása ismert, ld. a 72\*. feladatot). A multiplicitás lent nem szerepel, mivel az itt közölt sajátértékeknél ez mindig 1 volt; a sajátértékeket itt is kerekítés nélkül, két decimális számjeggyel adtuk meg.

$k \setminus n$	1	4	9	16	25	36	49	64	$\infty$
1.	32	25.37	22.86	21.72	21.10	20.74	20.50	20.34	$2\pi^2 \approx 19.74$
2.	-	72	62.56	57.79	55.20	53.63	52.62	51.93	$5\pi^2 \approx 49.35$
3.	-	86.4	71.55	63.32	58.88	56.27	54.60	53.47	$5\pi^2 \approx 49.35$
4.	-	145.15	120.55	107.03	99.03	93.99	90.62	88.26	$8\pi^2 \approx 78.95$
5.	-	-	153.6	137.63	126.04	118.74	113.98	110.73	$10\pi^2 \approx 98.70$
6.	-	-	165.45	145.06	130.11	121.04	115.35	111.59	$10\pi^2 \approx 98.70$
7.	-	-	206.23	182.21	166.64	156.79	150.28	145.77	$13\pi^2 \approx 128.30$
8.	-	-	257.58	223.48	197.90	180.00	167.90	159.51	$13\pi^2 \approx 128.30$
9.	-	-	319.95	259.24	241.84	223.69	210.59	201.39	$17\pi^2 \approx 167.78$
10.	-	-	-	269.79	248.54	227.57	212.93	202.92	$17\pi^2 \approx 167.78$
11.	-	-	-	282.43	256.01	237.73	225.27	216.47	$18\pi^2 \approx 177.65$

A fenti táblázat szerint, éppúgy mint a korábbi „véges differencia sajátértékeket” tartalmazó táblázat szerint a legelső sajátérték egyszeres. Ez nem véletlen, hanem a *Frobenius–Perron tétellel* megmagyarázható. Ugyanis a véges differencia fogalmazás az  $A_h y_h = \lambda_{VD} y_h$  standard sajátérték feladatra vezet, azaz itt

$$\frac{1}{\lambda_{VD}} y_h = A_h^{-1} y_h.$$

A Courant–elemekkel készített approximáció az  $Ax = \lambda_{VE} Bx$  általánosított sajátérték feladatot eredményezi, ahol  $A = h_1 h_2 A_h$ , vagyis ekkor

$$\frac{h_1 h_2}{\lambda_{VE}} x = A_h^{-1} Bx.$$

Itt  $A_h^{-1} > 0$  elemenként, ld. a 15.10. következményt 15.3.2-ben, de ekkor  $A_h^{-1} B > 0$  is pozitív elemű mátrix, ld. (15.243)-at 15.7.8-ban. Így  $A_h^{-1}$  és  $A_h^{-1} B$  maximális abszolútértékű sajátértéke egyenlő a spektrálsugarral (vagyis pozitív szám), és egyszeres sajátérték, az idézett tétel szerint (ld. Rózsa Pál könyvét, 478. o.), amely még azt is garantálja, hogy a hozzátartozó sajátvektornak minden komponense pozitív.

### 15.9.3 Gradiens módszerek az általánosított sajátérték feladat megoldására

A (15.311) és (15.312) végeelem sajátérték feladatok átírása mátrixalakba szokásos módon megvalósítható. Ehhez legyen  $\{w_i\}_{i=1}^N$  a  $V_h$  végeelem tér bázisa. Ekkor

$$u_h(\lambda^h) = \sum_{j=1}^N y_j w_j \quad \text{és} \quad v = w_i$$

behelyettesítésével (15.311)-be jutunk az

$$Ay = \lambda^h By \quad (15.326)$$

algebrai sajátérték feladathoz, amelyben  $y = (y_1, \dots, y_N)^T$ ,

$$A = (a_{ij}), \quad a_{ij} := a(w_j, w_i),$$

a merevségi mátrix és

$$B = (b_{ij}), \quad b_{ij} := (w_j, w_i)_0, \quad (15.327)$$

a tömegmátrix. Ez utóbbi – mint a  $\{w_i\}$  rendszer Gram-mátrixa – hermitikus és pozitív definit. Az  $A$  mátrix viszont nem feltétlenül rendelkezik ezekkel a tulajdonságokkal. A (15.309) koercivitásnak köszönhetően az euklideszi skalárszorzatban érvényes

$$(Ay, y) + c_0 \|y\|^2 \geq c_V (D_V y, y) \geq 0,$$

ahol  $D_V = (d_{ij})$  a  $V$ -skalárszorzattal kapcsolatos hermitikus és pozitív definit mátrix,  $d_{ij} := (w_j, w_i)_V$ .

A (15.312) feladat mátrixalakja

$$A^* z = \bar{\lambda}^h B z, \quad (15.328)$$

ha a (15.312)  $u_h^*(\lambda^h)$  megoldását az

$$u_h^*(\lambda^h) = \sum_{j=1}^N z_j w_j$$

formában keressük. A (15.328)-cal ekvivalens

$$z^* A = \lambda^h z^* B$$

egyenlet szerint a  $z^*$  vektorok az  $(A, B)$  mátrixpár baloldali sajátvektorai.

A gyakorlatban fontos rezgési feladatok szemidiszkrétizált alakja (ld. \*\*\* 18. fejezet 7. és 8. pontját)

$$Kx + C\dot{x} + M\ddot{x} = 0, \quad (15.329)$$

ahol  $x = x(t)$  a kitérések keresett vektora,  $\dot{x} = dx/dt$  stb.,  $K$  a merevségi,  $C$  a csillapítással kapcsolatos mátrix, míg  $M$  a tömegmátrix. Ez utóbbinak kiszámításával 16.4.8-ban foglalkozunk, azt a lehetőséget vizsgálva, hogy megfelelő kvadratúra képlet segítségével  $M$  diagonális alakban kapható-e. Ez a körülmény itt nagy munkamegtakarítást jelent és emellett Courant–elemek és (akár magasabbrendű) téglalapos elemek esetén ismert, hogy ettől az egyszeres sajátértékeknek és a hozzátartozó sajátfüggvényeknek a 15.48. tétel által biztosított konvergenciarendje továbbra is érvényes, ld. A.B. Andreev cikkét.

Amennyiben a (15.329) egyenlet megoldását harmonikus rezgés alakjában keressük :

$$x(t) = e^{i\omega t}y, \quad i := \sqrt{-1},$$

akkor a

$$(K + i\omega C - \omega^2 M)y = 0 \quad (15.330)$$

sajátérték feladatot kapjuk. Ha itt  $C$  elhanyagolható, akkor  $\lambda^h := \omega^2$  behelyettesítésével a (15.326) feladatra jutunk; egyébként viszont  $z := (y, \frac{1}{\omega}y)^T$  bevezetésével (15.330)-at a (15.326) alakban írhatjuk fel, ahol ekkor  $y$  helyett  $z$  áll és

$$A = \begin{pmatrix} iC & K \\ I & 0 \end{pmatrix}, \quad B = \begin{pmatrix} M & 0 \\ 0 & I \end{pmatrix}, \quad \lambda^h = \omega.$$

Ebben az esetben is  $B = B^* > 0$  (és  $B$  diagonális, ha  $M$  az), és  $A$  itt sem hermitikus, sem pozitív definit.

Az önadjungált, valós együtthatójú elliptikus sajátérték feladatok diszkrétizációjával viszont a következő nagyméretű, általánosított sajátérték feladathoz jutunk el (amelynek dimenzióját most az algebrai megközelítésnél szokásos  $n$  betűvel jelöljük) :

$$Av = \lambda Bv, \quad A = A^T > 0, \quad B = B^T > 0, \quad A, B \in \mathbb{R}^n. \quad (15.331)$$

Ennek sajátértékei pozitívak (ld. a 73. feladatot is) :

$$0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n; \quad (15.332)$$

sajátvektorai teljes rendszert alkotnak :

$$\text{span}\{v^{(i)}\}_{i=1}^n = \mathbb{R}^n, \quad Av^{(i)} = \lambda_i Bv^{(i)},$$

$$(Bv^{(i)}, v^{(j)}) = \delta_{ij}, \quad 1 \leq i, j \leq n. \quad (15.333)$$

Ezt a témát I. 3.7.2-ben már röviden érintettük. Ezért itt először emlékeztetünk arra, hogy  $B = B^T > 0$  miatt  $B$  Cholesky-felbontása létezik :  $B = LL^T$ , és a (15.331) egyenletet balról  $L^{-1}$ -gyel megszorozva, a  $v$  vektor helyett  $L^{-T}y$ -t írva megkapjuk a

$$Cy := L^{-1}AL^{-T}y = \lambda y$$

standard alakú sajátérték feladatot. Mivel  $C = C^T > 0$ , így következik (15.332); a hozzátartozó  $\{y^{(i)}\}_{i=1}^n$  sajátvektorok teljes rendszert alkotnak. Erről az  $\{y^{(i)}\}_{i=1}^n$  rendszerről feltehetjük, hogy ortonormált az euklideszi skalárszorzatban :

$$(y^{(i)}, y^{(j)}) = \delta_{ij}, \quad 1 \leq i, j \leq n.$$

A sajátértékek egyben (15.331) sajátértékei is; a sajátvektorokat visszatranszformálva nyerjük a (15.331) feladat sajátvektorait :  $v^{(i)} = L^{-T}y^{(i)}$ . Így

$$\delta_{ij} = (y^{(i)}, y^{(j)}) = (L^T v^{(i)}, L^T v^{(j)}) = (LL^T v^{(i)}, v^{(j)}) = (Bv^{(i)}, v^{(j)}).$$

A  $\{v^{(i)}\}_{i=1}^n$  sajátvektor rendszernek az ortogonalitása a  $B$  mátrix által definiált skalárszorzatban tehát algebrailag természetes módon adódik. Érdeemes észrevenni, hogy ez a hozzájuk tartozó  $V_h$ -beli („diszkrét”) sajátfüggvényekre nézve, azaz az

$$u_h^{(i)} = \sum_{k=1}^n v_k^{(i)} w_k, \quad i = 1, \dots, n,$$

függvények számára mit jelent : az ortogonalitást az  $L_2(\Omega)$  skalárszorzatában. Ugyanis

$$\begin{aligned} (u_h^{(j)}, u_h^{(i)})_0 &= \sum_{k=1}^n v_k^{(j)} \sum_{\ell=1}^n (w_k, w_\ell)_0 v_\ell^{(i)} \\ &= \sum_{\ell=1}^n \sum_{k=1}^n b_{\ell k} v_k^{(j)} v_\ell^{(i)} = (Bv^{(j)}, v^{(i)}) = \delta_{ij}, \end{aligned}$$

(15.327) és (15.333) miatt.

A (15.331) feladat numerikus megoldását szolgáló, a következőkben tárgyalásra kerülő gradiens módszerben alapvető szerepet játszik a Rayleigh-hányados (v.ö. I. 168. oldallal) :

$$\rho(x) := \frac{(Ax, x)}{(Bx, x)},$$

amelynek tulajdonságait az alábbi lemmában soroljuk fel.

**15.50. Lemma** (a Rayleigh–hányados tulajdonságai).

1. Ha  $x = v^{(i)}$  sajátvektor, akkor  $\rho(x) = \lambda_i$  a hozzátartozó sajátérték; megfordítva  $\rho(x) = \lambda_i$ -ből csak akkor következik  $x = v^{(i)}$ , ha  $i = 1$  vagy  $i = n$ .
2. Érvényes  $0 < \lambda_1 \leq \rho(x) \leq \lambda_n$  tetszőleges  $0 \neq x \in \mathbb{R}^n$ -re;
3.  $\rho(\alpha x) = \rho(x)$ , ahol  $0 \neq \alpha \in \mathbb{C}$  tetszőleges;
4.  $\nabla \rho(x) = (\rho'(x))^T = \frac{2}{(Bx, x)}(A - \rho(x)B)x$  a Rayleigh–hányados gradiense (deriváltjainak oszlopvektora);
5. Ha  $x \neq 0$ , akkor  $\nabla \rho(x) = 0$  (azaz  $x$  stacionárius pont)  $\Leftrightarrow x$  sajátvektor;
6.  $\rho(x)$  Hesse–mátrixa (ld. I. 6.5.)

$$H(x) = \frac{2}{(Bx, x)}(A - \rho(x)B - Bx(\nabla \rho)^T - \nabla \rho x^T B).$$

7. Ha  $x = v^{(i)} + \varepsilon w$ ,  $(Bw, w) = 1$ ,  $(Bv^{(i)}, w) = 0$ , akkor  $\rho(x) = \lambda_i + O(\varepsilon^2)$ .  $\square$

Ld. itt a 74. , 75. feladatokat is; a 7. ponthoz ld. I. 189. o.

Az 1. és 2. tulajdonság alapján  $\rho(x)$  globális minimumát, maximumát megkeresve kiszámíthatjuk a legkisebb, ill. legnagyobb sajátértéket és a hozzátartozó sajátvektorokat.

Erre a célra használhatjuk a gradiens módszereket (v.ö. I. 6.5.1-gyel; az ötlet Kantorovicsra megy vissza). Előnyük, hogy könnyen programozhatók és hogy minden lépésben csak az  $A$  és  $B$  mátrixszal kell megszorozni egy-két vektort. Ez  $n$  szerint lineáris lépésenkénti műveletigényt jelent azoknál az algebrai sajátérték feladatoknál, amelyek elliptikus sajátérték feladatok differencia, ill. végeelem diszkretizációjából származnak.

Mint kiindulási képletet tekinthetjük a következőt :

$$\begin{aligned} x_0 \text{ adott, } m = 0, 1, 2, \dots : \\ x_{m+1} &= x_m - t_m(A - \rho_m B)x_m, \quad (15.334) \\ \rho_m &:= \rho(x_m). \end{aligned}$$

Itt  $t_m$  alkalmasan kiválasztott pozitív szám – amire lejjebb visszatérünk.



A jelenlegi,  $\rho(x)$  szélsőértékeinek kiszámítására vonatkozó feladatunknak I. 6.5.1-hez képest a következő sajátságai vannak.

Bármely stacionárius pontba jutunk, a fenti lemma 5. és 1. pontja szerint ezzel egy sajátértéket és a hozzá tartozó sajátvektort megtaláltuk. Többszörös sajátérték alterében nem tud haladni a módszer, mert ott 5. alapján mindenütt  $\nabla\rho(x) = 0$ . A Hesse-mátrix sajnos nem pozitív definit a megoldásokon (v.ö. az I-ben található 6.5. lemmával, az előtte álló megjegyzéssel, valamint a 6.7. tétellel), ugyanis

$$H(v^{(i)}) = 2(A - \lambda_i B),$$

és így

$$(H(v^{(i)})v^{(i)}, v^{(i)}) = 0, \quad i = 1, \dots, n.$$

Kicsi  $x$ -re a Hesse-mátrix úgy viselkedik, mint  $2(A - \rho(x)B)/(Bx, x) = O(\|x\|_B^{-2})$ , így ott  $\nabla\rho$  Lipschitz-folytonossága nem garantálható (v.ö. a 6.6. tétellel). A szélsőérték helyek megkeresését nyilván az  $\|x\|_B = (Bx, x)^{1/2}$ -normához tartozó egységömbre kell korlátozni.

Ha nem a szélső, hanem közbülső sajátértékeket akarunk elérni, akkor az  $x_m$  közelítéseket mindig a már kiszámított sajátvektorokra kell ortogonalizálnunk.

Kiegészítve (15.334)-et normálással és ortogonalizációval, már majdnem üzemképes gradiens módszerhez jutunk. Ami még hiányzik, az a  $t_m$ -választás. Ennek megadásához kézenfekvő abból kiindulni, hogy teljesüljön a

$$\rho(x(t)) < \rho(x_m) \tag{15.335}$$

leereszkedési kritérium, ahol

$$x(t) := x_m - td_m, \quad d_m := (A - \rho_m B)x_m, \quad \|x_m\|_B = 1.$$

A Rayleigh-hányados definíciójából ekkor azt kapjuk (ld. a 76. feladatot), hogy teljesülnie kell a

$$2tc > t^2b \tag{15.336}$$

feltételnek, ahol

$$c := \|d_m\|^2, \quad b := \|d_m\|_B^2(\rho(d_m) - \rho_m). \tag{15.337}$$

Mivel feltehetjük, hogy  $d_m \neq 0$  (mert egyébként már  $x_m$  a sajátvektor és  $\rho_m$  a sajátérték), azért a  $c$  szám pozitív. Ennek alapján (15.336)-nak mindig van pozitív  $t$  megoldása :

$$\begin{aligned} \text{vagy } \rho_m &\geq \rho(d_m) \text{ - és ekkor tetszőleges } t > 0 \text{ jó,} \\ \text{vagy } \rho_m &< \rho(d_m) \text{ - és ekkor } 0 < t < 2c/b \text{ megfelelő.} \end{aligned}$$

Így bármely esetben pl.

$$t_m := c/(|b| + c\|d_m\|_B) \quad (15.338)$$

használható.

Az alábbiakban megadunk most olyan gradiens módszernek az algoritmusát – és megfogalmazzuk a konvergenciát biztosító tételt – amely az előbbi elgondolásokon alapszik, de javítja azokat a következő tekintetben :

1. A  $t_m$  lépésszámmat megválasztja a  $\rho(x(t)) = \min_t!$  követelményből;
2. A közbülső sajátértékek kiszámítását is célba veszi;
3. A  $d_m$  leereszkedési irányt definiálja egy  $P$  prekondicionálási mátrix segítségével :

$$Pd_m = (A - \rho_m B)x_m, \quad (15.339)$$

feltéve, hogy ez a  $P$  nemcsak szimmetrikus, pozitív definit és könnyen invertálható, hanem *spektrálisan ekvivalens* is az  $A$  mátrixszal: létezzenek  $0 < \gamma_1 \leq \gamma_2$  konstansok úgy, hogy

$$\gamma_1 P \leq A \leq \gamma_2 P. \quad (15.340)$$

$P$  lehet pl. a néhány többrácsos iterációnak megfelelő mátrix, nulla közelítésből kiindulva, az  $Ad_m = (A - \rho_m B)x_m$  egyenlet megoldására alkalmazva (ld. a 15.44. lemmát).

*Prekondicionált gradiens módszer algoritmus:*

Adott:  $k - 1 \geq 0$ , a már kiszámított  $v^{(i)}$  sajátvektorok száma,

$\varepsilon > 0$ , a pontosság (ld. az algoritmus 9. lépését),

max, a maximális iterációs szám,

$x_0$ , a  $k$ -adik sajátvektor nulladik közelítése;

keresett:  $v^{(k)}$ ,  $\lambda_k$ , a (15.331) általánosított sajátérték feladat  $k$ -adik megoldása

1.  $m := 0(1) \max$
2.  $[x_m := x_m - \sum_{i=1}^{k-1} (Bx_m, v^{(i)})v^{(i)}]$  projekció  $H_k^\perp := \text{span}\{v^{(i)}\}_{i=k}^n$ -be
3.  $r := (Bx_m, x_m)$ ,  $?r \leq 0?$  [válasszál új  $x_m$  közelítést,  $\rightarrow 1.$ ]
4.  $\rho_m := (Ax_m, x_m)/r$  a Rayleigh-hányados
5.  $r := \sqrt{r}$ ,  $x_m := x_m/r$  normálás
6.  $d_m := P^{-1}(A - \rho_m B)x_m$  a prekondicionálási egyenlet megoldása
7.  $d_m := d_m - \sum_{i=1}^{k-1} (Bd_m, v^{(i)})v^{(i)}$   $d_m$  projekciója  $H_k^\perp$ -ba

8.  $bd := (Bd_m, d_m)$ ,  $?bd \leq \varepsilon?$  [ $v^{(k)} := x_m, \rightarrow 12.$ ] a pontosság ellenőrzése
9.  $t_m := tv(x_m, d_m, \rho_m, bd)$   $t$ -választás, ld. lejjebb
10.  $x_{m+1} := x_m - t_m * d_m$   $m$ -ciklus vége
11.  $v^{(k)} := x_{m+1} / (Bx_{m+1}, x_{m+1})^{1/2}$ ,  $\lambda_k := \rho_m$
12. stop [eredmény:  $v^{(k)}$  a sajátvektor és  $\lambda_k$  a sajátérték]

A  $t$ -választást intézi a következő  $tv$  alprogram:

1.  $c := (d_m, (A - \rho_m B)x_m)$ ,  $ad := (Ad_m, d_m)$ ,  $rd := ad/bd$
2.  $b := bd * (rd - \rho_m)$ ,  $bx := (Bd_m, x_m)$ ,  $a := b * bx - c * bd$
3.  $?b < \min(0, c * (3bx + (bd)^{1/2}))?$  [ $tv := 1/(bd)^{1/2}, \rightarrow 5.$ ] itt az  $a = 0$ ,
4.  $tv := 2c / (b + \sqrt{b^2 - 4ac})$   $b < 0$  szingularitást is elkerüljük
5. kimenet a  $tv$  eredménnyel

A  $b^2 - 4ac \leq 0$  helyzettől egyébként nem kell tartanunk.

Ezen algoritmus programozásánál nyilván arra kell ügyelnünk, hogy a 2., 3. és 6. lépésben előforduló  $Bx_m$  vektort csak egyszer számítsuk ki, valamint arra is, hogy a  $\rho_m$  mennyiség kiszámítása során már a prekondicionálási egyenlet jobboldalát is összeállítjuk.

A  $bd$  és  $bx$  mennyiségek kiszámításához használt  $Bd_m$  vektor lehet még a 7. lépés elején készített  $Bd_m$  vektor, míg az  $ad$  és  $bd$  skalárszorzatok jobb oldalán álló  $d_m$  vektor kell, hogy már a 7. lépés utáni legyen (ahogyan ezt az algoritmus meg is adja).

Ezen megjegyzések figyelembevételével (és az  $x := x - \alpha * y$ -féle műveleteket – ha  $x, y$  vektorok,  $\alpha$  skalár – egy skalárszorzatnak számítva) azt látjuk, hogy a következő műveletigényes lépések vannak iterációnként a fenti algoritmusnak:

$$\begin{array}{lll}
 \text{skalárszorzatok:} & 4k + 4, & \\
 \text{szorzások } A\text{-val: } 2, \text{ szorzások } B\text{-vel:} & 2, & (15.341) \\
 \text{egyenletrendszerek megoldása:} & 1. &
 \end{array}$$

Ezen lépések mindegyikének  $n$ -ben lineáris műveletigénye van, ha az  $A$  és  $B$  ritka mátrixok abban az értelemben, hogy soronként korlátozott számú ( $n$ -től független) nemnulla elemmel rendelkeznek, és ha prekondicionáló eljárásaként a többrácsos módszert alkalmazzuk.

Más iterációs módszernek néhány lépésével is állíthatjuk elő a (15.339) egyenletrendszer  $P$  mátrixát és  $d_m$  megoldását. Ekkor viszont nem számíthatunk arra, hogy (15.340)-ben a  $\gamma_1, \gamma_2$  konstansok függetlenek lesznek  $n$ -től (amely összefügg a diszkretizáció lépéstávolságával).

Ez a körülmény ugyanakkor döntő az algoritmus konvergencia sebességének szempontjából, mint ahogyan ezt a következő tétel megmutatja.

**15.51. Tétel** (Prikazcsikov–Himics; prekondicionált gradiens módszer konvergenciája). Legyen  $1 \leq k \leq n - 1$ , teljesüljön (15.340) és

$$\lambda_k < \rho(x_0) < \lambda_{k+1} \quad (15.342)$$

adott  $x_0$  vektorral. Ekkor a fenti prekondicionált gradiens módszer konvergál a (15.331) feladat  $k$ -adik megoldásához:

$$\rho(x_m) \searrow \lambda_k, \quad x_m \rightarrow v^{(k)}, \quad m \rightarrow \infty.$$

Továbbá igaz a következő becslés:

$$0 \leq \rho(x_{m+1}) - \lambda_k \leq q_m(\rho(x_m) - \lambda_k), \quad m = 0, 1, \dots$$

ahol

$$q_m := 1 - \frac{\gamma_1 \lambda_{k+1} - \rho(x_m)}{\gamma_2 \lambda_{k+1}} \frac{\xi}{[\max(1, \xi - 1) + 2\sqrt{\xi(\xi - 1)}]}, \quad (15.343)$$

és itt  $\xi := \rho(x_m)/\lambda_k \searrow 1$ .  $\square$

**Megjegyzések. 1.** A tétel, amelynek többoldalas bizonyításától eltekintünk, csak a (15.342) feltétel teljesülése esetén garantálja a konvergenciát. Ez a feltétel bár ellenőrizhetetlen, de legtöbbször teljesül. Mindenesetre az  $x_0$  alkalmas választásának fontosságát aláhúzza. Ha pl. véletlenül  $x_0$  éppen  $v^{(k+2)}$ -nek az elég jó közelítése, akkor (15.342) nem igaz, hanem  $\rho(x_0) \approx \lambda_{k+2}$ , és  $d_m \approx 0$  miatt az algoritmust a 8. lépésben elhagyjuk. Ilyenkor tehát nem a  $(\lambda_k, v^{(k)})$ , hanem a  $(\lambda_{k+2}, v^{(k+2)})$  megoldást kapjuk.

**2.** Az itt vizsgált, többdimenziós elliptikus sajátérték feladatok diszkretizációjából származó  $A$  és  $B$  mátrixok esetén javasolható a következő algoritmus ahhoz, hogy  $v^{(k)}$  nulladik közelítését kiszámítsuk :

1.  $\ell := n/k, \ell 1 := (-1)^{k+1}$
2.  $i := 1(1)\ell [ x_i := 1, x_{n-i+1} := \ell 1 ]_i$
3.  $i := \ell + 1(1)n - \ell [ x_i := 0 ]_i$

Minden  $k$ -ra nézve ez nem lineárisan független rendszer, de itt a lényeg az, hogy a projekció  $H_k^\perp$ -be ne legyen a nullvektor.

A nulladik közelítések kiszámításánál a többrácsos gondolat is segíthet: oldjuk meg a feladatot először durva rácson ( $k$ -nál nem kevesebb rácsponttal), ahol jobbak a konvergencia viszonyok, majd ezután az eredményeket interpoláljuk a finomabb rácsra.

**3.** Az algoritmus 8. lépésében használt kilépési feltétel  $\varepsilon$ -ja a sajátértékek és nem a sajátvektorok pontosságát jellemzi, v.ö. a Rayleigh–hányados 7. tulajdonságával.

4. Amennyiben  $\lambda_k$  többszörös sajátérték, a tétel megint nem garantálja a konvergenciát. Ettől az algoritmus ( $k = 1, 2, \dots$ -re alkalmazva) a gyakorlatban mégis az  $S(\lambda_k)$  sajátaltérnek egy bázisát össze fogja állítani, az ortogonalizálásnak köszönhetően.

5. Az algoritmust mindenekelőtt arra használjuk, hogy a (15.331) feladat néhány legkisebb sajátértéket és hozzátartozó sajátvektort számítsuk ki olyan esetekben, amikor a feladat mérete túl nagy az I. 3.7.1-ben tárgyalt módszerek bevetésére: nem fér el  $A$  vagy  $B$  Cholesky-felbontása a gyors memóriában.

6. Az első – mondjuk  $j$  darab – sajátérték és sajátvektor kiszámítására célszerű úgy használni az algoritmust, hogy körülötte két ciklust szervezünk. A külsőben gyorsan csökkentjük  $\varepsilon$  értékét (pl. az előző  $\varepsilon$ -t mindig 10-re osztva míg el nem érjük a gépi epszilon 100-szorosának szintjét), a belsőben emeljük  $k$ -t (tehát:  $k := 1(1)j$ ). Amikor a  $k$ -ciklust egyszer végig jártuk, akkor a kapott  $\{v^{(k)}\}_{k=1}^j$  közelítő sajátvektorokat már nulladik közelítésnek használhatjuk a következő menethez (kihagyva a már konvergált sajátvektorok újraszámítását). Ilyenkor is – többszörös sajátértékek jelenlétében – elég jellemző az, hogy nem tisztán növekvő sorrendben állnak elő a sajátértékek.

7. Ha netán a legnagyobb sajátértékek kiszámítására akarjuk az algoritmust alkalmazni, akkor a  $t$ -választásnál át kell állítani a 3. lépést úgy, hogy alkalmas negatív értékét adjon és a 4. lépést úgy, hogy a  $0 = c - bt + at^2$  egyenlet másik gyökét eredményezze :

$$3'. \quad ?b > \max(0, c * \overline{(3bx - (bd)^{1/2})})? [tv := -1/(bd)^{1/2}, \rightarrow 5.]$$

$$4'. \quad tv := (b + \sqrt{b^2 - 4ac})/(2a)$$

ill. a (15.338) képletet használjuk  $-t_m$  meghatározására.

8. Ha minden feltétel teljesül és az algoritmus konvergál, akkor (15.343) szerint a sajátértékek legalább lineáris konvergenciájának  $q_m$  tényezője is konvergál, mégpedig

$$q_m \rightarrow q_\infty := 1 - \frac{\gamma_1 \lambda_{k+1} - \lambda_k}{\gamma_2 \lambda_{k+1}}$$

értékhez. Eszerint az aszimptotikus konvergencia sebességet csak a  $\lambda_k$  és  $\lambda_{k+1}$  közötti hézag relatív nagysága, valamint az ügyes prekondicionálás befolyásolja. Feltéve, hogy  $\gamma_1/\gamma_2 = \text{const}$  (függetlenül  $n$ -től, ami a többrácsos módszer segítségével elérhető), a  $k$ -adik sajátérték  $\varepsilon$  pontossággal való kiszámításának összköltsége – (15.341)-et is figyelembe véve – közvetlenül

$$Q(\varepsilon) = n[4(k+1)n + 2(N_A + N_B) + n_p] \log \varepsilon^{-1}, \quad (15.344)$$

ahol  $N_A$ , ill.  $N_B$  a nemnulla elemek száma  $A$ , ill.  $B$  egy tipikus sorában ( $5 \leq N_A, N_B \leq 36$ -ra számíthatunk másodrendű feladatoknál),  $n \cdot n_p$  viszont a (15.339) prekondicionálási egyenlet megoldásának költsége. A (15.344) képlet értelmét röviden foglalhatjuk össze úgy, hogy az algoritmus legköltségesebb része az ortogonalizáció (2. és 7. lépés).

**9.** Abban a gyakori esetben, amikor  $\text{cond}_2 A = \lambda_n/\lambda_1 = O(n^2)$ , a prekondicionálás hiánya (vagyis a  $P = I$  választás) azt eredményezi, hogy  $\gamma_1 = \lambda_1$ ,  $\gamma_2 = \lambda_n$ , és így  $q_m = 1 - O(n^{-2})$ . (Itt és a 8. pont számításában mindig a  $\lambda_i$  „diszkrét” sajátértékeket  $n$ -től függetlennek vettük, mivel az alacsony rendszámú diszkrét sajátértékek elég gyorsan konvergálnak a megfelelő „folytonos” sajátértékekhez.) Amikor a feladat  $n$  dimenziója túllépi kb. az 50-t (bonyolultabb feladatoknál már  $n=10$  esetén), akkor prekondicionálás hiányában érezhetően lassú lesz az algoritmus, sőt az is lehetséges, hogy a magasabb sorszámú sajátértékek megbízható értékeit nem látjuk soha.  $\square$

Ehhez a tételhez ld. a 77. feladatot is.

Az fent megtárgyalt módszerben a többrácsos eljárás hatékonyan segíthet, éspedig kétféle módon is :

1. jó nulladik közelítés kiszámításával;
2. mint olyan prekondicionálási eljárás a

$$Pd_m = f_m, \quad f_m := (A - \rho(x_m)B)x_m \quad (15.345)$$

egyenletek közelítő megoldására, amelynek  $P$  prekondicionálási mátrixa teljesíti a (15.340) feltételt  $n$ -től független  $\gamma_1, \gamma_2$  konstansokkal. Ekkor  $P$  az  $A = A_\ell$  mátrix  $\ell$ -edik diszkrétizációs szintjének megfelelő  $M_\ell$  többrácsos iterációs mátrix segítségével mint  $P = P_\ell = A_\ell(I_\ell - M_\ell^L)^{-1}$  írható fel, ha a többrácsos iterációt  $d_m^{(0)} = 0$  közelítéssel kezdjük meg és ha  $L$  iterációt hajtunk végre :

$$d_m^{(0)} := 0, \quad d_m^{(k)} = M_\ell d_m^{(k-1)} + (I_\ell - M_\ell)A^{-1}f_m, \quad k = 1, \dots, L. \quad (15.346)$$

Emlékeztetünk arra, hogy a többrácsos módszer konstruktív módon juttat olyan szimmetrikus, pozitív definit prekondicionálási  $P_\ell$  mátrixhoz, amelyet explicite nem számítjuk ki, fenti képlete (15.296)-nak felel meg, a prekondicionáló mátrix  $A_\ell$ -vel együtt szimmetrikus, pozitív definit, a (15.340)-ben követelt spektrális ekvivalenciáját  $A_\ell$ -vel a 15.44 lemma biztosítja és  $\gamma_2/\gamma_1$  relatív kondíciószáma  $h$ -től független és közel 1-hez.

A fenti két lehetőséget a többrácsos módszer használatára lehet akkor is igénybe venni, ha mint alapmódszert nem a gradiens, hanem az (eltolásos) inverz iterációt alkalmazzuk. Ezzel, vagy a fenti gradiens módszerrel

a leggyakrabban felmerülő sajátérték feladatot, a spektrum alsó részének a meghatározását, hatékonyan meg tudjuk oldani. Azonban az inverz iterációnak megvannak a maga konvergencia problémái (ld. pl. I. 3.4.4., 191. o.), és nagyon közeli, ill. többszörös sajátértékek kiszámítására nem javasolható – ha a feladat dimenziója nagy ( $n > 50$ -re lehet gondolni).

Ezért megemlítjük, hogy van egy 3. lehetőség : a többrácsos módszeren belül lehet az általánosított sajátérték feladatot – még a spektrum belsejében is – megoldani, ld. Hackbusch 1979-es cikkét, valamint az ő vagy Sajdurov többrácsos könyveit.

Akár ez utóbbi módszert alkalmazzuk, akár a fenti iterációkat, a közelítő sajátvektorok megszerzése után hasznosak a következő pontban megtárgyalt módszerek.

#### 15.9.4 A Rayleigh–Ritz- és a Lánczos–módszer

Most a (*Rayleigh–Ritz módszer*) leírására térünk rá, amelynek elve már I. 2.2. végén szerepelt, valamint az I. 3.9. pont 18. feladatában. Ezen módszer feladata az, hogy a valamilyen módszerrel és kisebb  $m$  számmal kapott  $\{\tilde{v}^{(k)}\}_{k=1}^m$  közelítő sajátvektor rendszer által kifeszített altérben kiválassza a sajátvektorok legjobb közelítéseit, ezeket  $B$ -ortonormalizálja és egyben a hozzátartozó sajátértékeket is állítsa elő. Alkalmazása különösen akkor ajánlott, amikor nagyon közeli sajátértékeket és a hozzátartozó sajátvektorokat kell szétválasztanunk.

Legyenek  $A$  és  $B$  szimmetrikus mátrixok,  $B$  pozitív definit. Induljunk ki abból, hogy a  $\tilde{v}^{(k)}$  vektorok bármely  $w$  lineáris kombinációja felírható mint

$$w := \sum_{k=1}^m t_k \tilde{v}^{(k)} = \tilde{V}_m t, \quad \tilde{V}_m := (\tilde{v}^{(1)}, \dots, \tilde{v}^{(m)}), \quad t := (t_1, \dots, t_m)^T.$$

Összesen  $m$  ilyen lineáris kombináció képezhető úgy, hogy ezeknek rendszere még mindig lineárisan független legyen. A lineáris kombinációkat célszerű a

$$w^{(j)} := \sum_{k=1}^m t_{kj} \tilde{v}^{(k)} = (\tilde{V}_m T_m)_j, \quad j = 1, \dots, m,$$

azaz

$$W_m = (w^{(1)}, \dots, w^{(m)}) = \tilde{V}_m T_m, \quad (15.347)$$

alakban felírni, ahol a  $T_m \in \mathbb{R}^{m \times m}$  mátrix  $j$ -edik oszlopa határozza meg a  $j$ -edik kombinációt. A célunk az, hogy a  $\{w^{(j)}\}_{j=1}^m$  rendszert olyanná válasszuk, hogy az  $B$ -ortonormált legyen,

$$(w^{(i)})^T B w^{(j)} = \delta_{ij}, \quad \text{azaz } W_m^T B W_m = I_m;$$

és hogy a közelítő sajátértékeket a Rayleigh–hányadosokból kapjuk meg:

$$\tilde{\lambda}_j = \frac{(w^{(j)})^T A w^{(j)}}{(w^{(j)})^T B w^{(j)}}. \quad (15.348)$$

A

$\tilde{\lambda}_j \delta_{ij} = \tilde{\lambda}_j (w^{(i)})^T B w^{(j)} = (w^{(i)})^T A w^{(j)}$ , azaz  $\tilde{\Lambda}_m = W_m^T B W_m \tilde{\Lambda}_m = W_m^T A W_m$  relációk mátrixalakja (15.347) szerint

$$\tilde{\Lambda}_m = T_m^T \tilde{B}_m T_m \tilde{\Lambda}_m = T_m^T (\tilde{V}_m^T B \tilde{V}_m) T_m \tilde{\Lambda}_m = T_m^T (\tilde{V}_m^T A \tilde{V}_m) T_m = T_m^T \tilde{A}_m T_m.$$

Itt  $\tilde{\Lambda}_m = \text{diag}(\tilde{\lambda}_i)$  diagonális  $m \times m$ -es mátrix,  $\tilde{B}_m := \tilde{V}_m^T B \tilde{V}_m$ ,  $\tilde{A}_m := \tilde{V}_m^T A \tilde{V}_m$  szimmetrikus  $m \times m$ -es mátrixok, és ezek közül  $\tilde{B}_m$  pozitív definit, ha  $\{\tilde{v}^{(k)}\}_{k=1}^m$  lineárisan független rendszer.

Ahogy látjuk, az  $\tilde{A}_m$ ,  $\tilde{B}_m$  mátrixokkal kapcsolatos egy  $m$ -dimenziós (tehát kisméretű), általánosított sajátérték feladat :

$$\tilde{A}_m t = \tilde{\lambda} \tilde{B}_m t.$$

Ezt pl. Jacobi–módszerrel megoldva kapjuk a  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_m$  sajátértékeket és a  $\{t^{(k)}\}_{k=1}^m$  sajátvektorok rendszerét. A különböző sajátértékekhez tartozó sajátvektorok eleve  $\tilde{B}_m$ -ortogonálisak, ezeket  $\tilde{B}_m$ -normalizáljuk, a megegyező sajátértékekhez tartozókat  $\tilde{B}_m$ -ortonormálisnak választjuk. A sajátvektorokból mint oszlopvektorokból összeállítva a  $T_m$  mátrixot, megkapjuk a Rayleigh–Ritz módszer végeredményét : a  $\{\tilde{\lambda}_k\}_{k=1}^m$  (javított) sajátértékeket és a javított sajátvektor rendszert, a  $W_m = \tilde{V}_m T_m$  mátrix oszlopainak alakjában. Ezek az oszlopok  $B$ -ortonormáltak – konstrukciónk szerint :

$$\delta_{ij} = (t^{(i)})^T \tilde{B}_m t^{(j)} = (t^{(i)})^T \tilde{V}_m^T B \tilde{V}_m t^{(j)} = (w^{(i)})^T B w^{(j)}.$$

Hasonló alapon a  $\tilde{\lambda}_j$  sajátértékek a (15.348) Rayleigh–hányadosokból számíthatók ki.

E pont befejezésül röviden jellemezzük a *Lánosz–módszert*. Ez olyan Rayleigh–Ritz módszer, amely kiindulási  $\{\tilde{v}^{(k)}\}_{k=1}^m$  rendszernek az általánosított sajátérték feladat *Krülöv–sorozatát* használja. Ezt, adott  $\tilde{v}^{(0)}$  vektorral kezdve, a

$$\tilde{v}^{(k)} = A^{-1} B \tilde{v}^{(k-1)}, \quad k = 1, 2, \dots$$

rekurzióval kapjuk meg. Itt  $A$  helyett az  $A - \sigma B$  eltolt mátrix is szerepelhet (és ezeket a mátrixokat nem invertáljuk, hanem a velük felírt egyenletrendszereket oldjuk meg). A Rayleigh–Ritz módszernek megfelelően a  $\tilde{v}^{(0)}, \tilde{v}^{(1)}, \dots$



sorozatot  $B$ -ortogonalizáljuk. Láncczos felfedezése, hogy a Krülov–sorozat mindenkori vektorát csak az utolsó két vektorra kell  $B$ -ortogonalizálni. Ezzel összefüggésben az

$$\tilde{A}_m t = \lambda \tilde{B}_m t$$

redukált sajátérték feladat  $\tilde{B}_m := \tilde{V}_m^T B \tilde{V}_m$  mátrixa diagonális és az  $\tilde{A}_m := \tilde{V}_m^T A \tilde{V}_m$  tridiagonális.

A Láncczos–módszer különösen akkor ajánlott, amikor nagyméretű a sajátérték feladat. Problémái vannak az ortogonalitás elvesztése miatt, és többszörös sajátértékek előállításánál. Ritkán az is előfordul (mint az inverz iterációnál konjugált komplex sajátértékpár esetén), hogy konvergál olyan számhoz, amely nem sajátérték. Ezért itt is a végén érdemes a közelítő sajátvektorok maradékvektorát kiszámítani a hozzátartozó sajátértéken. A módszer programja pl. T. Hughes, ill. Cullum és Willoughby könyveiben található, ld. Demmel könyvét is.

## 15.10 Összefoglalás

Az elliptikus egyenletek peremérték feladatainak numerikus megoldására három módszert ismertettünk, mindhárom esetben a keresett függvény közelítő értékeit egy rács pontjaiban kapjuk meg azáltal, hogy lineáris egyenletrendszerre vezetjük vissza a feladatot. Ez a rendszer ritkamátrixú és legtöbbször nagyméretű.

Az első a *véges differencia módszer*, amelyben a deriváltakat differenciáhozadosokkal helyettesítve lineáris rendszerre jutunk. Ez programozás szempontjából a legegyszerűbb a három módszer közül és a kapott rendszer megoldási algoritmusaként javasolható a *többrácsos módszer*, amely ekkor különösen gyorsan konvergál. Viszont már általánosabb tartomány és nem elsőfajú peremfeltételek olyan problémákat okozhatnak a diszkretizáció során, amelyeket a másodikként tárgyalt *boxmódszerrel*, más néven *véges térfogatmódszerrel* könnyebben meg lehet oldani.

Ez utóbbi nem a differenciálegyenletből indul ki, hanem annak egy lokális tartomány felett vett integráljából, ill. az integráلالakú megmaradási tételből. Ez a módszer kiváló első- és másodrendű olyan differenciálegyenletek esetén, amelyek divergencia-alakban vannak felírva, mint pl. a  $\operatorname{div}(k \operatorname{grad} u) + f = 0$  vagy a  $\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(f(u)) = 0$  egyenletek. Ezzel összefügg a módszer további előnye, a könnyű mérnöki interpretálhatóság: a diszkretizációt abból vezethetjük le, hogy egy adott pont körüli cellára (a boxra) felírjuk a bemenő és kimenő fluxusokat.

Harmadikként a *végeselem módszer* szerepel, amely általános tartományon, önadjungált differenciáloperátor és variációs elv jelenlétében a legjobb

(ami tipikus a mechanikából származó rugalmasságtani feladatokra, ott a minimális energia elvére lehet támaszkodni).

A végeelem módszer megvalósítása nehezebb (de éppen erre épül sok programcsomag, amely a mérnöki munkát segíti), résztartományonként polinomiális bázis segítségével állítjuk össze a közelítő megoldást. Így a megoldás nem csak a rács pontjaiban ismert, hanem az egész megoldási tartományban. A variációs feladatból lineáris egyenletrendszer kapható, ehhez sok integrált kell kiszámítani (alkalmas kvadratúra képletekkel). A lineáris rendszer megoldásvektora tartalmazza a polinomiális bázis együtthatóit.

Negyedrendű feladat esetén, ha egyszerű a tartomány és a peremfeltételek, a differencia módszer még alkalmazható, de általánosabb esetben és különösen még magasabb rendű (önadjungált) differenciálegyenlet esetén szinte csak a végeelem módszerrel lehet megoldáshoz jutni.

Fejezetünkben a kétdimenziós esetre korlátoztuk a tárgyalást, de mind a három módszer háromdimenziós feladatok esetén is működik.

## 15.11 Feladatok

### Feladatok a 15.1–15.3. pontokhoz

1. feladat. Módosítsuk a 15.5. tételt a következőképpen:

**15.13. Tétel** (további, maximumelvvvel rendelkező mátrixosztály). Az  $A$  mátrix legyen irreducibilis, (soronként) domináns főátlójú. Akkor teljesíti a maximumelvet.

[Kommentár. Az irreducibilitás definíciója I. 3.7.3-ban található. A fenti tételben a regularitás feltétele elmaradhat, hiszen az ilyen mátrix – Hadamard tétele szerint – reguláris, aminek bizonyítása elég hasonló a 15.5. és 15.13. tételekére.]

1\*. feladat. Az ötpontos differenciaséma  $\bar{A}_h$  mátrix számára

a) írjuk le az *inhomogén peremfeltételek eliminációját* mint a mátrix sorain és oszlopain, valamint az egyenletrendszer jobboldali vektoron dolgozó folyamatot;

b) lássuk be, hogy  $\bar{A}_h$ -val együtt  $A_h$  is M-mátrix, mert  $A_h = [\bar{A}_h/I_\gamma]$ , ld. (15.34). Ehhez használjuk az I. 1.4. lemmát és az 1.9. tételt.

2. feladat. Tisztázzuk a (15.36) 9-pontos differenciaséma pontosságát!

a) Feltéve, hogy  $u \in C^6(\bar{\Omega})$ , mutassuk meg, hogy a séma képlethibája :

$$\psi_\ell := \begin{cases} (\varphi(x) + u_{\bar{x}_1 x_1} + u_{\bar{x}_2 x_2} + \frac{h_1^2 + h_2^2}{12} u_{\bar{x}_1 x_1 \bar{x}_2 x_2})_\ell, & \text{ha } x^\ell \in \omega_h, \\ 0, & \text{ha } x^\ell \in \gamma_h, \end{cases}$$

negyedrendű, azaz  $O(h_1^4 + h_2^4)$  nagyságrendű! [Ehhez ld. II. 11.4.3-ban a (56)-ra vezető számítást is!]

b) A differenciaséma mátrixa legyen  $A_h^{(4)}$  (a peremfeltételek eliminációja után); bizonyítsuk be, hogy ez M-mátrix, amikor  $\frac{1}{\sqrt{5}} \leq h_1/h_2 \leq \sqrt{5}$ !

c) Az előző feltétel mellett bizonyítsuk be az  $\|(A_h^{(4)})^{-1}\|_{C(\omega_h)} \leq \frac{a^2+b^2}{16}$  becslést!

d) Vezessük le a séma egyenleteinek megoldását Fourier-módszerrel, a (15.48)-(15.54) relációkat megfelelően változtatva!

3. feladat. Vizsgáljuk azt az esetet, amikor (15.13)-ban  $\Omega$  olyan téglalapokból összeállított tartomány, amely összefüggő, és a  $h_1, h_2$  lépéstávolságú  $\omega_h$  ráccsal felosztható elemi téglalapokra! Ekkor vezessük le

a) a (15.16) séma (15.29)-nek megfelelő becslését;

b) a (15.36) sémának megfelelő becslést (ld. a 2.c) feladatot) !

4. feladat. Tekintsük a *háromdimenziós Poisson-egyenlet*

$$\begin{aligned} (\Delta u + f)(x) &= 0, & x &= (x_1, x_2, x_3) \in \{(0, a) \times (0, b) \times (0, c)\} =: \Omega, \\ u(x) &= g(x), & x &\in \Gamma, \end{aligned}$$

elsőfajú peremérték feladatát, és írjuk fel ennek differencia-approximációját a 15.3.1. pont alapján!

a) Feltéve, hogy  $u \in C^4(\overline{\Omega})$ , mutassuk meg, hogy a séma képlethibája :

$$\psi_\ell := \begin{cases} (f(x) + u_{\bar{x}_1 x_1} + u_{\bar{x}_2 x_2} + u_{\bar{x}_3 x_3})_\ell, & \text{ha } x^\ell \in \omega_h, \\ 0, & \text{ha } x^\ell \in \gamma_h, \end{cases}$$

másodrendű (azaz  $O(h_1^2 + h_2^2 + h_3^2)$  nagyságrendű)! Itt  $h_1 = a/N_1$ ,  $h_2 = b/N_2$ ,  $h_3 = c/N_3$  és  $N_j \geq 2$ ,  $j = 1, 2, 3$ .

b) A differenciaséma mátrixa legyen  $\overline{A}_h^{(3)}$ , ill. (a peremfeltételek eliminációja után)  $A_h^{(3)}$ ; rajzoljuk fel ezen mátrixok nulla és nemnulla elemeinek eloszlását! Adjuk meg a nemnulla elemek számát az  $N_1 = N_2 = N_3 = N$  esetben!

c) Bizonyítsuk be, hogy  $\overline{A}_h^{(3)}$  és  $A_h^{(3)}$  M-mátrixok, valamint hogy teljesítik a maximumelvet!

d) Bizonyítsuk be az  $\|(A_h^{(3)})^{-1}\|_{C(\omega_h)} \leq \frac{a^2+b^2+c^2}{24}$  becslést és a séma konvergenciáját!

5. feladat. Legyen  $\omega_h$  a (15.15)-ben definiált rács. A (15.40) sajátérték feladat (15.32)  $A_h$  mátrixát írjuk fel az egydimenziós eset  $\frac{1}{h^2} \text{tridiag}(-1, 2, -1)$  mátrixának a segítségével, és a (15.41)-ben rácsfüggvény alakjában megadott diszkrét sajátérték feladat megoldását írjuk fel blokkvektor formában!

[Útmutatás. Használjuk a Kronecker-szorzatot.]

[Kommentár. A feladat célja az, hogy az ekvivalenciát vektorok és rácsfüggvények között megmutassuk, valamint az is, hogy lássuk : az utóbbiak gyakran áttekinthetőbb felírásmódot nyújtanak, ha diszkretizált többdimenziós feladatokról van szó.]

6. feladat. Az egydimenziós esetben és a  $[0,1]$  intervallumon a második deriváltak megfelelő  $A_h$  mátrix inverzének becsléseként  $1/8$  adódott, akár a sajátértékek alapján akár majoráns függvény segítségével vezettük le. Magyarázzuk a (15.35) és (15.45) között tapasztalható különbség eredetét!

7. feladat. A 15.3.3. pontbeli gondolatmenetet követve állítsuk elő a (15.4) feladat egzakt megoldásának Fourier-sorfejtését! • Ennek alapján számítsuk ki az  $u(\frac{1}{2}, \frac{1}{2})$  értéket 6 számjegyre és hasonlítsuk össze azzal, amit a (15.16) differenciasémával kapunk  $h_1 = h_2 = 1/4, 1/8, 1/16$ -ra, a három közelítésre Richardson-extrapolációt (ld. I. 5.6. pont) alkalmazva!

8. feladat. Írjuk fel a 15.3.4. pont alapján a következő peremérték feladat másodrendű differenciasémáját :

$$\begin{aligned}(\Delta u + f)(x) &= 0, \quad x = (x_1, x_2) \in \{(0, a) \times (0, b)\}, \\ \frac{\partial u}{\partial n} &= g(x), \quad x_1 = 0, \quad 0 \leq x_2 \leq b, \quad \text{és } x_1 = a, \quad 0 \leq x_2 \leq b.\end{aligned}$$

A téglalap másik két oldalán teljesüljön

- a) az  $u(x) = g(x)$  elsőfajú;
- b) a (15.56) harmadfajú peremfeltétel (és ekkor legyen igaz (15.57)).

Vizsgáljuk a séma stabilitását, konvergenciáját, feltéve, hogy az adatok biztosítják, hogy  $u \in C^4(\bar{\Omega})$ !

[Útmutatás a b) részhez. Válasszuk a majoráns függvényt a következőképpen, alkalmas konstansokkal :  $w(x) = c_0 - c_1 x_1(a - x_1) + x_2(b - x_2)$ .]

9. feladat. A Shortley–Weller approximáción kívül tetszőleges tartományban megadott Poisson-egyenlet diszkretizációjára kézenfekvők a következő ötletek :

a) Használjuk csak az ekvidisztáns rácsot és a perem közelében vigyük át a peremértéket a hozzáeső legközelebbi belső rácspontra! Bizonyítsuk be, hogy az ehhez az approximációhoz tartozó diszkrét megoldás pontossága általában nem jobb mint  $O(h_1 + h_2)$  !

b) A peremet  $x^{(0)} = x_\gamma$ -ban metsző egyenes legyen párhuzamos az  $x_1$ -tengellyel. Ezen legyen  $x^{(1)}$  és  $x^{(2)}$  a legközelebbi két (belső) rácspontra. A három pontban felvett  $u(x^{(k)}) =: u_k$  értéket ( $k = 0, 1, 2$ ) interpoláló parabola képletét írjuk fel. Ez egy lineáris reláció  $u_1$ ,  $u_2$  és az adott  $u_0$  peremérték között – amely ugyancsak alkalmas a belső pontokhoz tartozó differencia-egyenletek lezárására. Vizsgáljuk az adódó differenciaséma képlethibáját és

pontosságát, feltéve, hogy a (15.13)-(15.14) feladat megoldása  $u \in C^4(\overline{\Omega})$ . Tisztázzuk ezen approximáció összefüggését a Shortley–Weller sémával!

10. feladat. A kétdimenziós *Poisson-egyenlet polárkoordinátákban* így néz ki :

$$\Delta_{r,\varphi} u + f(r, \varphi) = 0, \quad (r, \varphi) \in \{0 \leq r \leq 1, 0 \leq \varphi \leq 2\pi\},$$

ahol a Laplace-operátort a (15.9) adja meg.

a) Konstruáljuk ezen egyenlet másodrendű differencia-approximációját az ekvidisztáns  $(r, \varphi)$ -rácsra! Egészítsük ki az egyenletet elsőfajú  $u(1, \varphi) = g(\varphi)$  peremfeltétellel, és vizsgáljuk az adódó differenciaséma stabilitását, konvergenciáját!

[Útmutatás. Használjuk fel a II. 11.4.8. pontot, mint majoráns függvény-nyel próbálkozzunk a  $w(r, \varphi) = 1 - r$  függvénnyel!]

b) Az egyenletek megoldására fejlesszük ki a következő módszert :

$\varphi$  irányában a feladat periodikus és a differencia-operátor a másodrendű  $y_{\overline{\varphi}}$  differenciahányados, egy  $\varphi$ -tól független szorzóval. Emiatt ott a Fourier-módszer alkalmazható (ld. 15.3.3.). Az  $r$ -irányban minden Fourier-együtt-hatóra visszamarad egy (15.54)-re hasonló feladat. Ezt a rövidített Gauss-eliminációval oldhatjuk meg.

11. feladat. A pozitív kvadránsban fekvő negyedkörhöz konstruáljunk téglalap alakú rácsot a következő tulajdonságokkal :

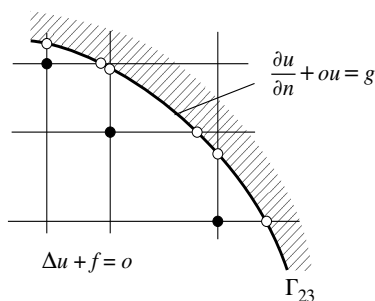
- a) a rács egyenesei párhuzamosak vagy az  $x_1$ -, vagy az  $x_2$ -tengellyel;
- b) minden perempont, amely egy rácsegyenesre illeszkedik, még egy másik (az elsőre merőleges) rácsegyenesre is illeszkedik;
- c) a rács szimmetrikus az  $x_1 = x_2$  egyenesre;
- d) van olyan  $C$  konstans, amellyel a rács  $h_1, h_2$  lépéstávolságai teljesítik a

$$\begin{aligned} |h_{1,i+1/2} - h_{1,i-1/2}| &\leq C \max_k (h_{1,k+1/2})^2, \\ |h_{2,j+1/2} - h_{2,j-1/2}| &\leq C \max_\ell (h_{2,\ell+1/2})^2 \end{aligned}$$

feltételt (amikor rácssereget készítünk).

12. feladat. A 15.10. tétel bizonyítását követve vezessük le a Shortley–Weller séma pontossági becslését, azaz bizonyítsuk be, hogy a séma  $y_h$  megoldására érvényes a következő becslés, ha  $u \in C^4(\overline{\Omega})$  és  $c := (\text{diam}(\Omega))^2$  :

$$\|\vec{u}_h - \bar{y}_h\|_{C(\overline{\omega}_h)} \leq \frac{c}{48} (h_1^2 + h_2^2) (M_4 + 4(h_1 + h_2)M_3 + 2(h_1^2 + h_2^2)M_4).$$



ábra 15.39: Perem harmadfajú peremfeltétellel

13. feladat. Bizonyítsuk be, hogy a (15.83) approximáció képlethibája a peremközeli pontokban általában csak  $O(1)$ -rendű! Ezután alkalmazva a 15.10. tételt lássuk be, hogy a hozzá tartozó differenciaséma pontossága másodrendű.

14. feladat. Tekintsük a Poisson-egyenlet harmadfajú peremérték feladatát az  $\Omega \in \mathbb{R}^2$  tartományban, amely peremének egyenlete kétszer folytonosan differenciálható. Vezessünk le

- a) elsőrendű, M-mátrixra vezető;
- b) másodrendű approximációt – a 15.39. ábrán mutatott helyzetben.

15. feladat. A (15.56) harmadfajú peremfeltétel approximációjára a sarokpontokban egyszerűbbnek tűnik (15.66)-(15.70) helyett a következő: az elsőrendű differenciahányados elsőrendű hibatagját (amely a megoldás második deriváltját tartalmazza) egy szomszédos belső pontban a szokásos másodrendű differenciahányadossal approximáljuk, hiszen ott már minden ehhez szükséges pont  $\bar{\omega}_h$  eleme. Mutassuk meg, hogy az eredő séma mátrixa nem lehet M-mátrix!

#### Feladatok a 15.4. ponthoz

16. a) feladat. Ismételjük meg a 15.4.2-beli számítást, amelynek (15.89) és az  $1/3$ -ados hibacsökkenés volt az eredménye, mégpedig (15.88) helyett azt követelve, hogy  $1 > 1 - \omega \lambda_{n(H)}^H = -(1 - \omega \lambda_{n(h)}^h) > -1$  teljesüljön (vagyis a két rácshoz tartozó maximális sajátértékeket használva).

b) Tekintsük a Poisson-egyenlet elsőfajú peremérték feladatát az egység-négyzetben és annak standard ötponos diszkretizációját négyzetes rácson, amikor  $N = 2n = 1/h = 2/H$ ! Mutassuk meg, hogy az egyszerű iteráció paraméterének kiválasztása az  $1 - \omega \lambda_{1n}^h = \omega \lambda_{N-1, N-1}^h - 1$  követelményből (ahol  $\lambda_{k\ell}^h$  a (15.42) szerinti „diszkrét” sajátértékeket jelöli) azt eredményezi,

hogy

$$|1 - \omega \lambda_{k\ell}^h| \leq \begin{cases} 1, & \text{ha } 1 \leq k, \ell \leq n-1, \\ 0.6, & \text{ha } n \leq k \text{ vagy } n \leq \ell. \end{cases}$$

Ezt a számítást is ismételjük meg az  $1 - \omega \lambda_{\max}^H = \omega \lambda_{\max}^h - 1$  követelménnyel.

17. feladat. A 15.4.3-ban leírt nemrekurzív (sima) többrácsos algoritmusból készítsük a teljes többrácsos módszer algoritmusát!

18. feladat. Adjuk meg a következő restriktós operátorok magterét abban az esetben, amikor a tartomány az egységnégyzet és a rács négyzetes,  $h = 1/N$  lépéstávolsággal! A restrikció legyen

- a) a direkt átvitel;
- b) a 9-pontos restrikció.

19. feladat. • a) Határozzuk meg az egyszerű iteráció optimális paramétereit! (Legyen az „Ellipsz” programban  $\ell = 4$ ,  $\gamma = 2$ ,  $\nu_1 = 1$ ,  $\nu_2 = 0$ ,  $\mu = 3$ , sima módszer, Poisson-egyenlet);

b) határozzuk meg a végső hiba szempontjából legjobb eljárás-kombinációt (tehát az iteráció, restrikció, interpoláció kombinációt) az „Ellipsz” program beépített feladatára, az előző paraméterek mellett!

20. feladat. • a) A (15.16)-(15.17) feladat esetére dolgozzuk ki a következő teljes többrácsos módszer egész programját, beleértve a tárolási struktúráját: V-ciklus, nincs előiteráció, egy Gauss-Seidel lépés mint utóiteráció, direkt átvitel mint restrikció, bilineáris interpoláció. Az „Ellipsz” program itt lehet hasznos a tesztelés során.

b) Figyeljük meg az előbbi feladat megoldásához szükséges gondolkodási és programozási időt, valamint azt a gépidőt is, amely a (15.16)-(15.17) probléma megoldásához,  $h_1 = h_2 = 1/16, 1/32, 1/64$  esetén kell. Hasonlítsuk össze ezeket az időket azzal, hogy mennyi gondolkodási, programozási és futási idő szükséges, ha ahelyett egy formulamanipulációs program segítségével a legegyszerűbb elképzelhető algoritmust valósítjuk meg, a teljes mátrix kiszámítását és a lineáris rendszer megoldását a program által kínált megoldó procedúrával (ha van, akkor ritkamátrixú)!

21. feladat. a) Határozzuk meg az  $N_2/N_1$  viszony minimumát akkor, amikor  $N_1$  a síkbeli téglalap alakú rács összes pontjának a száma és  $N_2$  ezen rács felezésével kapott rács pontszáma! Az így meghatározott rácsokat továbbfelezve kapjuk az  $N_\ell$  számokat is. Melyik  $\ell$ -re igaz  $N_\ell/N_{\ell-1} > 3$ ?

b) Oldjuk meg ezt a feladatot akkor is, amikor a tartomány háromszög, és ennek három csúcsa képezi a kiindulási rácsot!

c) Az a) és b) részfeladatok birtokában ellenőrizzük a 15.4.4. pont becsléseit!

22. feladat. Számítsuk ki a többrácsos módszer tár- és műveletigényét abban az esetben, amikor az egységkockában,  $h_\ell = 1/N_\ell = \frac{1}{2}h_{\ell-1}$  lépéstávolságú rácson megoldandó a Poisson-egyenlet. Konkrétan tekintsük a teljes többrácsos módszert,  $V$ -ciklussal, 2 elősimító iterációval, trilineáris interpolációval, direkt átvitelrel!

22\*. feladat. Legyen  $A \in \mathbb{R}^{n \times n}$  reguláris. Az  $Ay = b$  egyenletrendszer megoldására tekintsük az

$$y^{(0)} \text{ adott, } \quad y^{(m)} = Sy^{(m-1)} + Tb, \quad m = 1, 2, \dots,$$

alakú iterációt. Írjuk fel a  $\nu$  lépés után adódó összefüggést  $y^{(\nu)}$  és  $S, T, b, y^{(0)}$  között! Milyen kapcsolat van  $S, T$  és  $A$  között annak alapján, hogy jogosan elvárhatjuk: tetszőleges  $b$ -re az  $y = A^{-1}b$  pontos megoldás fixpontja az iterációnak? És milyen kapcsolat van az  $e^{(\nu)} := y^{(\nu)} - y$  végső és  $e^{(0)} := y^{(0)} - y$  kiindulási hibák között?

[Megjegyzés. A többrácsos módszer tárgyalásakor szimbolikusan írjuk  $y^{(\nu)} = S^{(\nu)}(A, b)y^{(0)}$ -t, az  $S$  szerepét aláhúzva.]

23. feladat. Terjesszük ki a 15.4.6. pont vizsgálatát arra a kérdésre, hogy mi az optimális paraméter-választás, ha két iterációt veszünk különböző paraméterekkel?

24. feladat. Legyenek  $A_H, A_h, P_H^h$  mint a 15.4.6. pontban, de  $R_h^H$  a *direkt átvitel*, azaz  $(R_h^H y_h)_i = (y_h)_{2i}$ . Vizsgáljuk meg a konvergenciát ebben az esetben! Ehhez

a) mutassuk meg, hogy az  $\omega = \frac{h^2}{4}$  esetben és az  $x := \omega \lambda_j^h, y := \omega \lambda_{n-j}^h$  jelölésekkel (tehát ekkor  $x + y = 1$ , továbbá  $n = 2N = 1/h$ ) az euklideszi normában érvényes

$$\|C_h A_h S_h^\nu (\alpha v^{h,j} + \beta v^{h,2N-j})\|^2 = (\alpha x y^{\nu-1})^2 + (\beta y x^{\nu-1})^2.$$

b) bizonyítsuk be ezután, hogy

$$\|C A_h S_h^\nu\| \leq \begin{cases} \cos \frac{\pi h}{2}, & \text{ha } \nu = 1, \\ \frac{1}{4}, & \text{ha } \nu = 2, \\ \frac{4}{27}, & \text{ha } \nu = 3. \end{cases}$$

c) • Végezzük a következő numerikus kísérletet: az „Ellipsz” program bekapcsolásakor megjelenő módszer-kombinációval oldjuk meg a beépített feladatot, majd ezután ismételjük a számítást a Jacobi-módszerrel ( $\omega = h^2/8$ ), direkt átvitelrel és bilineáris interpolációval!

[Kommentár. Általános érvényű az itt megfigyelhető jelenség, hogy a direkt átvitel mint restrikciónak csak óvatosan használandó.]



25. feladat. a) Mutassuk meg, hogy a 15.14. lemma akkor is igaz, ha  $1 \leq \omega \|A_h\| \leq \frac{5}{3}$ . Pontosabban, ekkor a (15.109) simító tulajdonság teljesül a  $c_s = 20c_0/9$  konstanssal.

b) A (15.89) ill. a 16. feladat szerint meghatározott  $\omega$  paraméterekről állapítsuk meg, vajon benne vannak az  $[\frac{1}{\|A_h\|c_h(\omega_h)}, \frac{5}{3\|A_h\|}]$  intervallumban?

c) A 15.14. lemmában szereplő  $f_\nu(t) := t(1-t)^\nu$  függvényre lássuk be, hogy igaz  $0 \leq f_\nu(t) \leq \frac{1}{2(1+\nu)}$ , amikor  $0 \leq t \leq 1 + \frac{\sqrt{2}-1}{1+\nu}$ .

26. feladat. Lássuk be, hogy két tetszőleges négyzetes  $A, B$  mátrixra teljesül  $\rho(AB) = \rho(BA)$ , az euklideszi normában igaz  $\|A\|^2 = \|A^T A\| = \|AA^T\|$ , de általános normában  $\|AB\| = \|BA\|$  nem érvényes!

27. feladat. • Hasonlítsuk össze a következő módszer-kombinációk hatékonyságát (a végső hiba alapján) a beépített feladat megoldása során, ha a restriktió a direkt átvitel, az interpoláció a bilineáris!

- a) fekete-fehér iteráció;
- b) fehér-fehér iteráció.

Mi a tapasztalható különbség oka? Az a) feladat megoldásakor kísérletezzünk az "Ellipsz" programmal azzal a céllal is, vajon hány iteráció szükséges a konvergenciához?

28. feladat. • Határozzuk meg az optimális módszer-kombinációt (a végső hiba alapján) a következő feladat megoldása során : egzakt megoldás  $u = x^3 + y^2$ ;  $a_1 = a_2 = 1$ ,  $b_1 = c = 0$ ,  $b_2 = -100$ ,  $f(x, y) = 200y - 2 - 6x$ ,  $\ell = itl = 4$ ,  $\gamma = \nu_1 = itk = 2$ ,  $\nu_{21} = 0$ . A Jacobi-iteráció használatakor válasszuk  $\omega$ -t a  $[0.01, 0.1]$  intervallumból.

Ismételjük a kísérleteket a  $b_2 = +100$  esetben!

### Feladatok a 15.5–15.6. pontokhoz

29. feladat. Tekintsük a (15.13)-(15.14) feladatot az egységnyezet esetén ( $a = b = 1$ ,  $h_1 = h_2 = h$ ). Az elemi módon bebizonyítható

$$x^2 \geq \sin^2 x \geq \frac{x^2}{1 + x^2(1 - \frac{4}{\pi^2})}, \quad \text{ha } 0 \leq x \leq \frac{\pi}{2},$$

becslés segítségével javítsuk a 15.19. lemma (15.124) egyenlőtlenségét a következőre :

$$0 \leq G_h^0(x, \xi) \leq \gamma_0 + \gamma_1 \ln \frac{1}{h}, \quad \gamma_0 := \frac{5}{4} + \frac{\ln 2}{\pi} - \frac{4}{\pi^2} = 1.065\dots, \quad \gamma_1 := 2\pi.$$

30. feladat. Legyenek  $y, z$  a  $\{0 \leq x_1 \leq a, 0 \leq x_2 \leq b\}$  téglalap szokásos  $\omega_h$  rácsán definiált diszkrét argumentumú függvények, amelyeket nullával a

diszkrét perem  $\gamma_h$  pontjaira folytatjuk. Továbbá, legyenek  $a_1, a_2$  olyan diszkrét argumentumú függvények, amelyeknek  $a_1(x_1 + \frac{h_1}{2}, x_2)$  ill.  $a_2(x_1, x_2 + \frac{h_2}{2})$  értékei definiáltak az  $x = (x_1, x_2) \in \omega_h^1$  ill. az  $x \in \omega_h^2$  rácspontokon (ld. 15.5.2. pont). Bizonyítsuk be, hogy a 15.3.3. és 15.5.2. pontokban bevezetett skalárszorzatokra érvényesek a következő relációk :

$$\begin{aligned}(y_{\bar{x}_1 x_1}, z)_{(0,h)} &= -(y_{x_1}, z_{x_1}]_1, \\ (y_{\bar{x}_2 x_2}, z)_{(0,h)} &= -(y_{x_2}, z_{x_2}]_2,\end{aligned}$$

valamint

$$\begin{aligned}((a_1 y_{x_1})_{\tilde{x}_1}, z)_{(0,h)} &= -(a_1 y_{x_1}, z_{x_1}]_1, \\ ((a_2 y_{x_2})_{\tilde{x}_2}, z)_{(0,h)} &= -(a_2 y_{x_2}, z_{x_2}]_2.\end{aligned}$$

[Útmutatás. Használjuk fel a II. 11.4.6. pontot, és mindenekelőtt a 11.3. lemmát.]

31. feladat. Tisztázzuk azt, hogy a (15.130)-(15.131) feladat  $k_1, k_2, f$  adatainak milyen simasága esetén másodrendű a (15.133)-(15.134) séma képlethibája!

[Útmutatás. Használjuk a II. 11.4.8. pont anyagát.]

32. feladat. a) A *vegyes deriváltakkal* rendelkező

$$\mathcal{L}u := \sum_{i,j=1}^2 \frac{\partial}{\partial x_i} \left( k_{ij}(x) \frac{\partial u}{\partial x_j} \right), \quad k_{ij} = k_{ji}$$

elliptikus operátorra adjunk olyan differencia-approximációt, amely téglalap alakú rács esetén ( $h_1, h_2$  lépéstávolságokkal)

- 1) szimmetrikus pozitív definit mátrixra;
- 2) konstans együtthatók esetén M-mátrixra vezet!

Itt teljesüljön

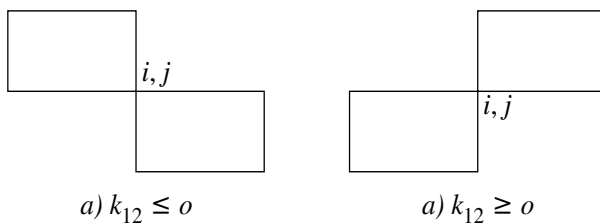
$$(K(x)\xi, \xi) \geq k_0 |\xi|^2, \quad \text{minden } \xi = (\xi_1, \xi_2)^T \text{-ra és minden } x \text{-re,}$$

vagyis, ekvivalensen,

$$k_{11}(x), k_{22}(x) \geq k_0, \quad (k_{11}k_{22} - k_{12}^2)(x) \geq k_0[k_{11} + k_{22} - k_0](x) \quad \text{minden } x \text{-re.}$$

A 2) feladat megoldását keressük a  $|k_{12}| \leq \min(\frac{h_2}{h_1} k_{11}, \frac{h_1}{h_2} k_{22})$  feltétel mellett.

Megjegyezzük, hogy ilyen  $\mathcal{L}$  operátorral  $\mathcal{L}u + f = 0$  alakban a stacionárius (kétdimenziós) rugalmasságtan egyenletrendszere írható fel (ld. 18.1.2-ben \*\*\* a (13) egyenletet).



ábra 15.40: Vegyes deriváltak differencia-approximációja

[Útmutatás 1)-hez: induljunk ki a 15.40. a) ábra szerinti differenciacsillagból, és használjuk a következő differencia operátort :

$$\Lambda y := -\frac{1}{2} \sum_{i,j=1}^2 [(k_{ij}y_{\bar{x}_j})_{x_i} + (k_{ij}y_{x_j})_{\bar{x}_i}].$$

A  $\Lambda$  operátorhoz tartozó  $A_h$  mátrix pozitív definittségét mutassuk meg a (15.43) skalárszorzatra vonatkozólag.

Útmutatás 2)-höz : induljunk ki  $k_{12} \leq 0$  esetén a 15.40. a), ill. ha  $k_{12} \geq 0$ , akkor a 15.40. b) ábra szerinti differenciacsillagból!]

32 b) A harmadfajú peremfeltételnek 15.3.4-beli approximációját és pontossági vizsgálatát általánosítsuk a (15.131) operátor esetére, amikor a peremfeltétel alakja  $k \frac{\partial u}{\partial n} + \sigma u = g$ . A tartomány legyen a téglalap.

33. feladat. Mutassuk meg, hogy a 15.23. lemma becslésében

$$\|\varphi\|_{H_0^{-1}(\omega_h)} \text{ helyett } \frac{ab}{\sqrt{6(a^2 + b^2)}} \|\varphi\|_{L_2(\omega_h)}$$

is állhat, ezután a 31. feladat eredményét felhasználva igazoljuk a (15.133)-(15.134) séma maximumnormabeli konvergenciáját!

[Útmutatás. A 11.5. lemma (II. 185. o.) bizonyításából kiindulva bizonyítsuk be az  $a, b$  oldalhosszúságú téglalap  $\omega_h$  rácán megadott, annak peremén eltűnő  $v$  rácsfüggvényekre teljesülő

$$\|v\|_{L_2(\omega_h)} \leq \frac{ab}{\sqrt{6(a^2 + b^2)}} \|v\|_{H_0^1(\omega_h)}$$

diszkrét beágyazási tételt (a *diszkrét Friedrichs-féle egyenlőtlenséget*)!]

34. feladat. Legyen  $f \in H^\gamma(\Omega)$ ,  $-1 < \gamma \leq 0$ , tehát a Poisson-egyenlet  $u$  általánosított megoldására, amely eleget tesz (15.157)-nek, érvényes  $u \in H^{2+\gamma}(\Omega) \hookrightarrow L_\infty(\Omega)$ . Bizonyítsuk be, hogy ekkor

$$\|y_h - \vec{u}_h\|_{L_2(\omega_h)} \leq M h^{2+\delta} \|u\|_{H^{2+\delta}(\Omega)}$$

ahol  $y_h$  a (15.149), (15.156) séma megoldása és  $\vec{u}_h$  a pontos értékek vektora.

35. feladat. A boxmódszer segítségével vezessük le a (15.163) egyenlet differencia-approximációját az  $x^0 \in \gamma_{23,h}$  pontban, amikor ott a perem belső szöge  $\frac{\pi}{2}$  és az  $x^0$ -ból induló peremszakaszok párhuzamosak a koordináta tengelyekkel. Eredményünket hasonlítsuk össze a (15.66)-(15.70) approximációkkal!

36. feladat. Vezessük le a boxmódszer segítségével a (15.163),  $k \equiv 1$ , egyenletnek a differencia-approximációját akkor, amikor az  $x^0$  rácspont egy peremmenti derékszögű háromszög csúca :

a) Legyen  $x^0 \in \omega_h$  a háromszög derékszögű csúca, a többi  $x^0$  csúcsú elemi tartomány téglalap;

b) az alaprács ekvidisztáns  $h_1, h_2$  lépéstávolságokkal,  $x^0 \in \Gamma_{23}$  és egy téglalap, valamint két olyan (peremmenti) háromszög csúca, amelyeknek átfogója ugyanazon egyenes része.

37. feladat. a) Oldjuk meg a 14. feladat a) részét a boxmódszer segítségével!

b) Vezessük le a boxmódszer (15.163)-hoz tartozó képletét az  $x^0$  pontban abban az esetben, amikor az  $x^0$ -t hat szabályos,  $h$  oldalhosszú háromszög veszi körül.

### Feladatok a 15.7. ponthoz

38. feladat. Legyen  $u|_{\Gamma} = 0$ , teljesüljenek a 15.192 után felsorolt feltételek, beleértve 15.193-at.

a) A Schwarz-egyenlőtlenség segítségével becsüljük az

$$a(u, w) := \int_{\Omega} [(K \operatorname{grad} u) \cdot \operatorname{grad} w + w \vec{v} \cdot \operatorname{grad} u + cuw] dx$$

kifejezést felülről  $\|u\|_1 \|w\|_1$  segítségével!

b) Ugyanezen bilineáris forma  $\int_{\Omega} w \vec{v} \cdot \operatorname{grad} u dx$  tagját becsüljük felülről  $\|u\|_1 \|w\|_0$  segítségével, majd  $w = u$ -t behelyettesítve nyerjük az  $a(u, u)$  alsó becslését

$$a(u, u) \geq k_0 |u|_1^2 - c_2 \|u\|_0^2$$

alakjában (ez egy *Gårding-féle egyenlőtlenség*).

c) Ezután tisztázzuk azt, hogy mekkora lehet  $\vec{v}$  euklideszi normája  $k_0$ -hoz képest (ill.  $k_0$ -hoz és  $c \geq c_0$ -hoz képest), hogy a forma pozitív definit legyen?

[Útmutatás a c) feladathoz. Itt hasznos lesz a Friedrichs-féle egyenlőtlenség.]

d) Tekintsük most a (15.192) egyenletet a (15.203) peremfeltétel mellett, amikor  $\vec{v} \neq 0$  és teljesül (15.201). Adjuk meg a (15.205)-öt és a (15.206)-ot

helyettesítő  $a(u, w)$ - és  $\varphi(w)$ -képleteket! Fogalmazzunk meg alkalmas feltételeket  $\sigma$ -ra és  $(\vec{n}, \vec{v})$ -re, amelyek a bilineáris forma ellipticitását biztosítják!

39. feladat. a) A Babuška–lemma 2. megjegyzésében szereplő Lax–Milgram tétel egzisztencia-részére adjunk közvetlen bizonyítást azáltal, hogy megmutatjuk: az ott definiált  $A$  operátorral felírt  $Au = \psi$  egyenlet megoldása egyben az  $F_\omega(v) := v - \omega(Av - \psi)$  leképezés fixpontja, és a  $v = F_\omega(v)$  egyenlet megoldására alkalmazható a Banach-féle fixpont tétel.

b) Bizonyítsuk be, hogy  $\|A^{-1}\| \leq \frac{1}{\alpha}$ , ahol  $A$  a megjegyzésben szereplő  $V \rightarrow V$ -leképezés.

c) Vizsgáljuk meg, hogy a végelem módszer elméletének melyik helyein használtuk az ellipticitás tulajdonságát!

[Útmutatás a)-hoz. Bizonyítsuk be, hogy teljesül  $\|F_\omega(v) - F_\omega(w)\|_V \leq q\|v - w\|_V$ , ahol  $q \in [0, 1)$ , ha  $0 < \omega < 2m_a/M_a^2$ .]

40. feladat. Adjuk meg azt a lineáris rendszert, amely a  $\Delta$  nem elfajult háromszögön definiált  $u(x) = a + bx_1 + cx_2 + dx_1^2 + ex_1x_2 + fx_2^2$  polinom  $a, \dots, f$  együtthatóit az  $x^j$  csúcsokban ( $j = 1, 2, 3$ ) és az  $x^j$  oldalfelező pontokban ( $j = 4, 5, 6$ ) felvett  $u(x^j)$  értékek segítségével fejezi ki! Számítsuk ki a mátrix determinánsát is!

41. feladat. a) Írjuk fel az  $R_9$  elem formafüggvényeit (amely elem  $R_8$ -ból úgy keletkezik, hogy a téglalap középpontjában felvett függvényértéket 9. szabadsági foknak vesszük) abban az esetben, amikor a téglalap az  $\{-1 \leq x_1, x_2 \leq 1\}$  négyzet az  $x_1, x_2$ -síkon!

b) Oldjuk meg értelemszerűen a 40. feladatot az  $R_9$  elem esetén is!

42. feladat. Mutassuk meg, hogy az  $R_8$  elem másodrendű:  $P_2(\overline{\Omega}_i) \subset P(\overline{\Omega}_i)$ , de bármely  $k$ -adik és  $\ell$ -edik szabadsági fok elhagyása ( $1 \leq k < \ell \leq 8$ ) azt eredményezi, hogy  $P_2(\overline{\Omega}_i) \not\subset P(\overline{\Omega}_i)$ !

[Útmutatás. Az elem (15.213) formafüggvényeivel tekintsük a következő reláció által definiált  $A \in \mathbb{R}^{8 \times 8}$  mátrixot:

$$\sum_{\ell=1}^8 y_\ell w_\ell(x_1, x_2) = zAy,$$

ahol  $z := (1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^2x_2, x_1x_2^2)$  sorvektor és  $y := (y_1, \dots, y_8)^T$  oszlopvektor. Az  $A$  regularitása azt jelenti, hogy tetszőleges  $b \in \mathbb{R}^8$  együttható vektorral a  $z \cdot b$  polinom  $P(\overline{\Omega}_i)$  eleme, és ehhez  $y = A^{-1}b$  választandó. Fogalmazzuk meg  $A$  segítségével, mit jelent két szabadsági fok elhagyása, mit az, hogy  $P_2(\overline{\Omega}_i) \subset P(\overline{\Omega}_i)$ .]

43. feladat. Az  $R_8$  elem 3 dimenzióra való általánosításaképpen tekintsük azt az esetet, hogy  $\Omega_i$  kocka, csomópontjai a  $(\nu_1, \nu_2, \nu_3)$  csúcsokon kívül a kocka élfelező pontjai is! (Itt  $j = 1, 2, 3$ -ra  $\nu_j = \pm 1$ , egymástól független

előjelekkel.) Adjuk meg a formafüggvényeket! Tisztázzuk azt a kérdést is, vajon az adódó elem valóban másodrendű.

[Kommentár. Ezt az elemet  $R_{20}^3$ -szal jelöljük.]

44. feladat. Mutassuk meg, hogy az  $\Omega$  tartomány konform triangulációján definiált, résztartományonként polinomiális  $\{w_\ell\}$  bázis esetén csak akkor igaz  $V_h := \text{span}\{w_\ell\} \in C^1(\bar{\Omega})$ , ha ez a bázis Hermite-féle elemekből áll, vagyis ha minden  $i$ -re az  $\Omega_i$  résztartományra szűkített  $V_h$ , a  $P(\bar{\Omega}_i)$  polinomtér egy-egy elemének egyértelmű meghatározásához deriváltakat is kell megadni.

45. feladat. Bizonyítsuk be a (15.217) és (15.218) feltételek ekvivalenciáját háromszöges elemek esetén! Tehát, konkrétan : legyen  $h$  a háromszög leghosszabb oldala,  $\vartheta$  a legkisebb szöge,  $\rho$  a beírt kör sugara, és érvényes  $h \leq \kappa_1 \rho$ . Akkor igaz  $\sin \vartheta \geq \frac{1}{\kappa_1}$ . Megfordítva, ha  $\sin \vartheta \geq \frac{1}{\kappa_2}$ , akkor következik  $h \leq 4\kappa_2 \rho$ . Mutassuk meg azt is, hogy a háromszög  $|\Delta|$  területe nem kisebb  $h^2/(4\sqrt{\kappa^2 - 1})$ -nél!

46. feladat. Bizonyítsuk be, hogy a Courant–elem esetén érvényes a következő, úgynevezett *inverz egyenlőtlenség* is, ha a trianguláció egyenletesen reguláris :

$$|v_h|_{1,\Omega} \leq \frac{c(\kappa)}{h} \|v_h\|_{0,\Omega}, \text{ minden } v_h \in V_h\text{-ra.}$$

Itt  $h$  a trianguláció maximális oldala. Adjuk meg a  $c$  függőségét a  $\kappa$  konstanstól (ld. (15.217)), ill. a  $\vartheta_i$  belső szögek trianguláció feletti minimumától!

[Útmutatás. Járjunk el háromszögenként, használjuk itt is, mint a 15.31. tételhez fűzött 1. megjegyzésben a transzformációt az egységssimplexre.]

47. feladat. Legyen az  $\Omega$  tartomány  $M$  darab  $h_1 \times h_2$ -méretű téglalap uniója és  $S := \{0 < x_1, x_2 < 1\}$ . A 15.31. tétel bizonyítása részleteit adjuk meg a derékszögű (bilineáris)  $R_4$ -elemek esetére.

[Útmutatás. Ld. a tételhez tartozó 3. megjegyzést.]

48. feladat. a) A 15.7.6-beli eredményekkel magyarázzuk meg, miért nem alkalmas az  $R_4$  elem esetén a középpont szabály? (Az viszont az  $R_4$  használata során keletkező jobboldali  $\varphi$ -integrálokra javasolható.)

b) • A kétdimenziós Poisson–egyenletnek az egység négyzetben megadott elsőfajú peremérték feladatához tartozó összes végeelem képletet számítsuk ki az  $R_4$  elem és a középpont szabály segítségével, programozzuk és vizsgáljuk az adódó eljárás stabilitását és konvergenciáját kísérletileg!

c) A 2-pontos tenzorszorzat (15.235) Gauss–képlet miért nem javasolható az  $R_8$  elemhez?

d) Bizonyítsuk be, hogy a 15.7.6-ban a  $T_6$  elemhez ajánlott (15.237) képlet a másodfokú polinomokat pontosan integrálja!

49. feladat. Milyen kvadratúra képlet adható meg, amely csak az  $R_8$ -elem alappontjait használja és másodfokú polinomokon pontos? Hasonlítsuk össze ezen kvadratúra képlet műveletigényét a 2-pontos Gauss-tenzorsorozat képletével, ha a) egy elem feletti integrál, b) ha az egész tartomány feletti integrál kiszámítandó!

50. feladat. Írjuk fel pszeudokód alakjában a merevségi mátrixot és a terhelési vektort kiszámító 15.7.7-beli algoritmus módosítását arra az esetre, hogy a mátrixot kompakt módon tároljuk!

51. feladat. Bizonyítsuk be, hogy a végeelem megoldás a rácspontokban megegyezik a pontos megoldással, ha az egységnégyzetben megadott  $\Delta u + f = 0$  egyenletet oldjuk meg elsőfajú peremfeltételek mellett,  $f = -30(x_1^4 + x_2^4)$  esetén, ha a diszkretizáció a 15.7.8. pontnak felel meg és az  $(f, w_i)_0$  integrálokat pontosan számítjuk ki!

52. feladat. Tegyük fel, hogy (15.246)-ban  $f \in C^2(\square_1)$ . Bizonyítsuk be, hogy  $(f_h, w_0)_0$  az  $(f, w_0)_0$  másodrendű approximációját adja! Vizsgáljuk ugyanezt a kérdést abban az esetben is, amikor  $f \in H^1(\square_1)$ , a Bramble-Hilbert lemmát alkalmazva!

53. feladat. Amennyiben az  $(1, 0)$ ,  $(\frac{1}{2}, \frac{1}{2})$ ,  $(0, 1)$  pontokhoz tartozó  $(h_1, 0)$ ,  $\frac{1}{2}(\mu h_1, \nu h_2)$ ,  $(0, h_2)$  képpontok az  $x_1, x_2$ -síkon egy egyenesen fekszenek, akkor a  $T_6$ -os izoparametrikus leképezés az  $(1, 0)$ ,  $(0, 1)$  egyenesből mindig egyenest gyárt-e? (Ez a belső háromszögek miatt kívánatos lenne, akkor görbeperemű elem csak  $\Gamma$  mentén fordulna elő.)

54. feladat. Mutassuk meg, hogy (15.255) – parametrikus alakban – egy parabolát definiál alkalmas, ferdeszögű koordinátarendszerben, ha  $\mu \neq 1$  és  $\nu \neq 1$ .

55. feladat. Tekintsük a  $T_6$  elemnek a következő megfeleltetésekkel definiált izoparametrikus leképezését :

$$\begin{aligned} (\xi, \eta) & : (0, 0) (1, 0) (0, 1) (\frac{1}{2}, 0) (0, \frac{1}{2}) (\frac{1}{2}, \frac{1}{2}) \\ (x_1, x_2) & : (0, 0) (1, 0) (0, 1) (\mu, 0) (0, \mu) (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}) \end{aligned}$$

Itt egy  $\mu \in (0, 1)$  paraméter. Mutassuk meg, hogy

a) a  $\xi + \eta = 1$  peremvonal nem megy át az egységkörbe (függetlenül  $\mu$ -tól), hanem érvényes  $x^2(\xi, 1 - \xi) + y^2(\xi, 1 - \xi) = 1 + g(\xi)$ , ahol  $|g(\xi)| \leq (3 - 2\sqrt{2})/8$ ,  $0 \leq \xi \leq 1$ , és ez a görbe nem lép ki az egységkörből;

b) a  $\xi = 0$ ,  $0 \leq \eta \leq 1$  (ill. a  $\eta = 0$ ,  $0 \leq \xi \leq 1$ ) peremszakasz csak akkor megy át az  $x = 0$ ,  $0 \leq y \leq 1$  (ill.  $y = 0$ ,  $0 \leq x \leq 1$ ) peremszakaszba, ha  $\frac{1}{4} \leq \mu \leq \frac{3}{4}$ .

56. feladat. Tekintsük az  $R_8$  elem alsó oldalának a következő megfeleltetésekkel definiált izoparametrikus leképezését :

$$\begin{aligned}(\xi, \eta) &: (0, 0) (1, 0) \left(\frac{1}{2}, 0\right) \\(x_1, x_2) &: (0, 0) (1, 0) \left(\frac{1}{2}, \frac{1}{2}\right)\end{aligned}$$

Határozzuk meg az adódó görbe távolságát az  $(x - \frac{1}{2})^2 + y^2 = \frac{1}{4}$  körtől!

57. feladat. Tekintsük az  $R_8$  elem azon (elfajult) izoparametrikus leképezését az  $x_1, x_2$ -sík egységsimplexére, amely alatt a  $\xi = 0$ ,  $0 \leq \eta \leq 1$  oldal átmege a  $(0,0)$  pontba, a  $\xi = 1$ ,  $0 \leq \eta \leq 1$  oldal viszont az  $x_1 + x_2 = 1$  oldalba! Mutassuk meg, hogy a  $(g$  folytonos függvénnyel képzett)  $f(x_1, x_2) = g(x_1, x_2)/\sqrt{x_1^2 + x_2^2}$  integrálja az egységsimplex felett átmege olyan integrálba az  $R_8$  elem felett, amelynek integrandusa folytonos!

58. feladat. Az  $R_8$  elem esetére írjuk ki a Newton-módszer minden képletét, amellyel a (15.253) egyenletrendszert meg lehet oldani  $(\xi, \eta)$  szerint, adott  $(x_1, x_2)$  mellett!

59. feladat. a) Téglalapos felosztás és az  $R_4$  elem esetén tisztázzuk azt, vajon folytonos a végeelem interpoláció abban az esetben, ha 3 – koordináta-irányokkal párhuzamos – téglalapnak olyan közös (perem)pontja van, amely kettőnek a csúcpontja, a harmadiknak az oldalfelező pontja.

b) Oldjuk meg az előző feladatot arra az esetre, hogy az  $R_8$  elemet használjuk.

60. feladat. Legyen  $\Omega$  az egységnégyzet,  $\Gamma$  a pereme. Bizonyítsuk be a (15.207) egyenlőtlenség diszkrét megfelelőjét,

$$\|v\|_{L_2(\gamma_h)} \leq C_1 \|v\|_{H^1(\bar{\omega}_h)}, \quad (C_1 = \sqrt{5})$$

kiindulva abból, hogy

$$v_{0j} = v_{ij} - \sum_1^i v_{\bar{x},kj} h, \quad v_{Nj} = v_{ij} + \sum_{i+1}^N v_{\bar{x},kj} h$$

stb., a Cauchy- és az  $\varepsilon$ -egyenlőtlenséget használva. Itt  $\bar{\omega}_h$  az  $\bar{\Omega}$ -beli,  $h = 1/N$  lépéstávolságú rácspontok halmaza,  $\gamma_h := \Gamma \cap \bar{\omega}_h$  a diszkrét perem, és a normákra érvényesek a következő definíciók :

$$\begin{aligned}\|v\|_{L_2(\gamma_h)}^2 &:= \sum_{i=0}^N (v_{i0}^2 + v_{iN}^2) \bar{h}_i + \sum_{j=0}^N (v_{0j}^2 + v_{Nj}^2) \bar{h}_j, \\ \bar{h}_k &:= \begin{cases} h, & \text{ha } 1 \leq k \leq N-1, \\ h/2, & \text{ha } k = 0, N, \end{cases} \\ \|v\|_{H^1(\omega_h)}^2 &:= \sum_{i=0}^N \sum_{j=0}^N v_{ij}^2 \bar{h}_i \bar{h}_j + \sum_{i=1}^N \sum_{j=0}^N (v_{\bar{x},ij})^2 h \bar{h}_j + \sum_{i=0}^N \sum_{j=1}^N (v_{\bar{y},ij})^2 h \bar{h}_i.\end{aligned}$$



Mindkét esetben tehát a trapézsabályról van szó.

**Feladatok a 15.8. ponthoz**

61. feladat. Courant–elemek esetén írjuk fel a (15.258) összefüggés  $\beta_{nm}$  együtthatóit!

62. feladat. Mutassuk meg, hogy az  $L_2(0, 1)$  tér  $\|\cdot\|_0$  normája és a  $h^{1/2}$ -del súlyozott euklideszi norma ekvivalensek a szakaszonként lineáris függvények (ld. II. 11.6.2)  $V_h$  terén a következő értelemben. Ha

$$u_h := \sum_{j=1}^N y_j w_j \in V_h, \quad y := (y_1, \dots, y_N)^T \in \mathbb{R}^N, \quad h := 1/(N+1),$$

ahol  $V_h$  a kalapfüggvények által kifizített tér, akkor  $h$ -tól független  $0 < c_0 \leq c_1$  konstansokkal érvényes

$$c_0 \left( \sum_{j=1}^N y_j^2 h \right)^{1/2} \leq \|u_h\|_0 \leq c_1 \left( \sum_{j=1}^N y_j^2 h \right)^{1/2}.$$

[Útmutatás. Használjuk a II. 11.6.2-beli összefüggéseket, pl. a  $\lambda_{\min}(G)$  becslését.]

63. feladat. Tekintsük a 15.7.8-ban szereplő (15.249) egyenlet elsőfajú peremérték feladata Lagrange-féle végelem diszkretizációjának  $A$  mátrixát,

$$A = (\alpha_{ij}), \quad \alpha_{ij} := \int_{\Omega} \sum_{\ell=1}^n k(x) \frac{\partial w_j}{\partial x_{\ell}} \frac{\partial w_i}{\partial x_{\ell}} dx,$$

ahol  $\Omega \in \mathbb{R}^d$  véges tartomány. Bizonyítsuk be, hogy  $0 \leq k(x) \leq K_0$ ,  $x \in \Omega$  esetén érvényes

$$\lambda_{\max}(A) \leq dK_0\kappa^2 h^{d-2} m_1^2 m_2,$$

ahol  $h := \max h_{\max}(\Omega_m)$  a legnagyobb elem átmérője,

$h_{\max}(\Omega_m)$  az  $m$ -edik elem átmérője,

$h_{\min}(\Omega_m)$  viszont az  $m$ -edik elem legkisebb távolsága két csomópont között és

$\kappa \geq h_{\max}(\Omega_m)/h_{\min}(\Omega_m)$  minden  $m$ -re;

$m_1$  az egy csomópontoz tartozó elemek maximális száma és

$m_2$  az egy elemhez tartozó szabadsági fokok maximális száma.

64. feladat. Folytatva a 62. feladatot mutassuk meg általánosabb poligonális  $\Omega \in \mathbb{R}^2$  tartomány és  $V_h(\Omega) := \text{span}\{w_j\}_{j=1}^N$  végelem tér esetén, hogy az  $u_h$   $L_2(\Omega)$ -beli és az  $y$  euklideszi normák ekvivalensek, megengedve viszont megfelelő  $h$ -hatványok megjelenését  $c_0$  és  $c_1$  mellett.

[Útmutatás. Vizsgáljuk a kérdést a  $\{w_j\}_{j=1}^N$  rendszer Gram-mátrixa szélső sajátértékeinek segítségével.]

65. feladat. • Futassuk az „Ellipsz” programot a következő esetben : a beépített feladathoz állítsuk be az  $\ell = 4$  szintszámot, kérjük a sima többrácsos módszert. A megoldáshoz használjuk a fekete-fehér Gauss–Seidel iterációt, valamint a végeelem interpolációt / végeelem restrikiót, jegyezzük fel a kapott eredményeket (4 teljes többrácsos iteráció utáni maximum és  $L_2$  normábeli hibáját). Ezeket hasonlítsuk össze a 9-pontos (azaz harmadrendű) interpoláció / 9-pontos restrikió kombinációval kapható eredménnyel!

66. feladat. A 15.41. lemma segítségével vezessünk le olyan becslést a  $\gamma = 2$  esetben, amely élesebb a 15.42. tételénél!

67. feladat. Bizonyítsuk be, hogy az euklideszi normában és tetszőleges  $A \in \mathbb{R}^{n \times n}$  mátrix esetén  $\|A\|^2 = \|AA^T\| = \|A^T A\|$ .

68. feladat. Lássuk be, hogy a (15.300) sajátérték feladat minden sajátértéke nemnegatív!

### Feladatok a 15.9. ponthoz

69. feladat. A (15.317) feladatnak  $325\pi^2$  többszörös sajátértéke. Adjuk meg az összes hozzátartozó sajátfüggvényt!

70. feladat. A  $\lambda^* = 50\pi^2$  a (15.317) feladat háromszoros sajátértéke. Mutassuk meg, hogy  $\lambda_{1,7}^h \rightarrow \lambda^*$ ,  $\lambda_{5,5}^h \rightarrow \lambda^*$ ,  $0 < \lambda_{1,7}^h - \lambda_{5,5}^h \leq 4\pi^4 h^2 \{-24 + 65\pi^2 h^2\}$ , amikor  $h \leq \frac{1}{8}$ .

71. feladat. a) Bizonyítsuk be, hogy tetszőleges  $\Omega$  tartomány esetén az (elsőfajú peremfeltételekkel kiegészített) Laplace-operátor sajátfüggvénye nem lehet polinom.

b) Konstruáljunk olyan másodrendű differenciáloperátort, amelynek (alkalmas tartományon, elsőfajú peremfeltételek mellett) polinomiális sajátfüggvénye van!

72. feladat. Bizonyítsuk be, hogy a (15.318) sajátértékek között van  $(2N)^2$ , és ennek multiplicitása pontosan  $N - 1$ .

72\*. feladat. A  $T_3$ - (avagy Courant-) elemes approximációt alkalmazzuk a (15.317) sajátérték feladatra, amikor az  $\omega_h$  rács a következő (angolul ezt *criss-cross grid*-nek hívják) : Az egységnégyzetet osszuk fel egyenletesen kis,  $2h$  oldalhosszú négyzetekre, ahol  $h = h_1 = h_2 = 1/N$ , páros  $N$ -nel. Ezután a kis négyzeteket tovább bontjuk fel 8-8 háromszögre úgy, hogy mind-egyikbe a két átlót és az oldalfelező pontokat összekötő két merőleges vonalat behúzzuk. A diszkrét sajátfüggvényeket  $v^{(k,\ell)}(x) = \sin k\pi x_1 \sin \ell\pi x_2$  alakban tételezzük fel,  $x = (x_1, x_2) \in \omega_h = \{(ih, jh), 1 \leq i, j \leq N - 1\}$ . Számítsuk ki analitikusan az ehhez tartozó  $\lambda_{k\ell}^h$  sajátértékeket!

73. feladat. Tekintsük a (15.331) feladatot abban az esetben, amikor az  $A, B$  mátrixok bár szimmetrikusak, de egyik sem pozitív definit. Adjunk meg olyan  $2 \times 2$  mátrixpárt, amelyhez komplex sajátértékek tartoznak!

[Kommentár. A sajátértékek pl. akkor valósak, amikor az  $A$  és  $B$  valamelyik lineáris kombinációja pozitív definit.]

74. feladat. Kiindulva az  $x = \sum_1^n c_i v^{(i)}$  felbontásból és a  $(Bx, x) = \sum_{i=1}^n c_i^2$  azonosságból, bizonyítsuk be a Rayleigh-hányados 1. és 2. tulajdonságát!

75. feladat. Parciális deriválással vezessük le a Rayleigh-hányados gradiensének és Hesse-mátrixának képleteit!

76. feladat. Mutassuk meg, hogy (15.335) és (15.336)-(15.337) ekvivalensek!

[Útmutatás. Használjuk a Rayleigh-hányados és  $d_m$  definícióját, valamint  $x_m$  normáltságát.]

77. feladat. Lássuk be, hogy  $P$  pozitív definitisége miatt nem lehet  $\|w\|_B = 0$  a prekondicionált gradiens módszer algoritmus 7. lépése után, ha előtte  $(A - \rho_m B)x_m \neq 0$ .

78. feladat. Az egységnyezeten és a  $h = 1/N$  lépéstávolságú négyzetes rácson vizsgáljuk a diszkrét biharmonikus operátor,

$$\Delta_h^2 y := y_{\bar{x}_1 x_1 \bar{x}_1 x_1} + 2y_{\bar{x}_1 \bar{x}_2 x_1 x_2} + y_{\bar{x}_2 x_2 \bar{x}_2 x_2},$$

sajátértékeit elsőfajú peremfeltételek esetén. \*\*\*

79. feladat. a) Számítsuk ki az  $A = \text{tridiag}(a, b, c) \in \mathbb{R}^{n \times n}$  mátrix sajátértékeit (ahol  $a, b, c$  tetszőleges valós számok)!

[Útmutatás. Ha  $ac \neq 0$ , akkor szimmetrikus a  $DAD^{-1}$  mátrix, ahol

$$D := \text{diag}_{1 \leq i \leq n}(d^i), \quad d := (c/a)^{1/2}.]$$

b) Adjunk kritériumot arra, hogy mikor valós az  $A$  összes sajátértéke!

c) Alkalmazzuk az eredményeket az  $au'' + bu' + \lambda u = 0$ ,  $0 \leq x \leq 1$ ;  $u(0) = u(1) = 0$  feladat upwind és központi differencia approximációjára (ld. \*\*\* 17.4.0.)!



# Fejezet 16

## Parabolikus egyenletek

### 16.1 Parabolikus egyenletek eredete és néhány alkalmazása

#### 16.1.1 Környezetvédelmi balesetek

Sajnos elég gyakran fordulnak elő az olyan helyzetek, mint a következő. Egy kamion felborult az országúton, a szállított tartály igen mérges anyagot tartalmaz, és az elkezd kipárologni. A polgári védelemnek ekkor lehetőleg gyorsan fel kellene mérnie azt a kérdést, vajon evakuálni kell a legközelebbi falut vagy sem? Ehhez olyan információra van szüksége, mint a kamion és a falu pozíciója, a (konstansnak feltételezett) szélirány és -sebesség ( $v$ ), a diffúziós együttható ( $D$ ), a kiszabaduló anyag mennyisége per időegység ( $Q$ ) (ez utóbbi leginkább problematikus), és az adott anyagnak olyan  $c_{krit}$  koncentrációja, amely már veszélyes.

Ezután a folyamat matematikailag a következő egyenletekkel írható le. Legyen  $c(x, t)$  a koncentráció egy adott  $x = (x_1, x_2, x_3)$  pontban és  $t$  pillanatban, az  $x_1$  koordináta mutasson a szél irányába, a kamion pozíciója legyen az origóban, a falué pedig  $x^0$ -ban (és  $L$  legyen a távolsága a kamiontól);  $t = 0$  legyen a baleset időpontja. Akkor a  $c$  koncentrációt  $\Omega := \{-L \leq x_i \leq L, i = 1, 2, 0 \leq x_3 \leq H\}$ -ban (alkalmas  $H$  magassággal) és  $t > 0$ -ra a következő egyenletekből határozzuk meg :

a) a konvekció-diffúzió egyenletből (ez az anyag szétterülését időben és térben írja le)

$$\frac{\partial c}{\partial t} + v \frac{\partial c}{\partial x_1} = D \sum_{i=1}^3 \frac{\partial^2 c}{\partial x_i^2} + Q \delta(x, 0), \quad x \in \Omega, \quad t > 0, \quad (16.1)$$

Itt  $\delta(x, 0)$  az origóban koncentrált Dirac-féle delta-függvény.

b) a  $c(x, 0) = 0$  kezdeti feltételből, minden  $x \in \Omega$ -ra,

c) a tartomány szélén is meg kell adnunk a koncentrációt vagy annak gradiensét. Konkrét esetünkben az a célszerű és bevált, hogy minden  $t > 0$ -ra a következőket követeljük :

$$\begin{aligned} x_1 = -L : \quad c = 0, & \quad x_1 = L : \quad \frac{\partial c}{\partial x_1} = 0; \\ x_2 = -L : \quad -\frac{\partial c}{\partial x_2} = 0, & \quad x_2 = L : \quad \frac{\partial c}{\partial x_2} = 0; \\ x_3 = 0 : \quad -\frac{\partial c}{\partial x_3} = 0, & \quad x_3 = H : \quad \frac{\partial c}{\partial x_3} = 0. \end{aligned}$$

Ezen feladat megoldására kell, hogy rendelkezésre álljon gyors és megbízható program, amelynek már csak a konkrét adatokat kell megadni, és keresett  $c(x^0, t)$ . A fenti kérdés most úgy hangzik, vajon  $c(x^0, t) \geq c_{krit}$ , amikor  $t > 0$ ? Ennek megválaszolása segít dönteni az evakuálásról.

A parabolikus egyenletekre jellemző az, hogy az egyik független változó szerinti derivált csak elsőrendű ((16.1)-ben ez a  $t$ ), míg a többi változó (fent  $x_1, x_2, x_3$ ) szerint minden másodrendű derivált fordul elő. (A parabolikus egyenlet pontos definícióját 16.2-ben adjuk meg.)

Az a változó, amely szerint csak az elsőrendű derivált fordul elő, gyakran – de nem mindig – az időt jelöli.

Erre a következő példát tudjuk adni. A

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x_1^2} - v \frac{\partial u}{\partial x_1} + \frac{\partial}{\partial x_2} \left( d(x_2) \frac{\partial u}{\partial x_2} \right) + f(x_1, x_2, t)$$

térben kétdimenziós parabolikus egyenlet segítségével pl. a folyók vízszennyeződését modellezhetjük :

$u$  a mélység felett átlagolt koncentráció,  $D = \text{const} > 0$  a hosszirányú és  $d > 0$  a keresztirányú diffúziós együttható,  $v > 0$  az áramlási sebesség és  $f$  a szennyező forrás intenzitása.

Ha stacionárius esetet vizsgálunk és a  $D$ -t tartalmazó tagot elhanyagoljuk (mivel  $D$  kicsi), akkor ugyancsak parabolikus egyenlethez jutunk, amelyben ekkor viszont az  $x_1$  hosszirányú térkoordináta átveszi az idő szerepét :

$$v \frac{\partial u}{\partial x_1} = \frac{\partial}{\partial x_2} \left( d(x_2) \frac{\partial u}{\partial x_2} \right) + f.$$

### 16.1.2 A Black–Scholes egyenlet

A tőzsdén igen nyereségesek az opciók eladásával foglalkozó üzletek, de kockázatuk nagy. Ennek felmérésére alkalmasnak találták (bizonyos, az uralkodó

gazdasági viszonyokra vonatkozó feltételek mellett) a következő egyenletet :

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0, \quad (16.2)$$

ahol  $V = V(S, t)$  az opció eladási ára,  $0 \leq t \leq T$  az idő,  $T$  az opció esedékességi időpontja,  $S$  azon áru (vagy részvény) ára, amelyre az opció vonatkozik,  $\sigma$  az áru kockázati tényezője,  $r$  a mindenkori kamatláb bankkölcsönökre. Adott az opció eladási ára  $t = T$ -nél, valamint peremfeltételek  $S = 0$ -nál (pl.  $V = 0$ ) és  $S \rightarrow \infty$ -nél (pl.  $V/S \rightarrow 1$ ). A kérdés az, hogy mekkora  $V(S, 0)$ , mert ekkor eldönthető, vajon érdemes megvenni az opciót (ha a jelenlegi,  $t = 0$  időpontban  $V(S, 0)$ -nál kisebb árára kínálják).

A következő transzformációval az egyenletet átalakítjuk a később tárgyalásra kerülő formára :

$$V(S, t) = v(x, \tau), \quad S = e^x, \quad t = T - \tau/(\sigma^2/2).$$

A transzformáció eredménye az alábbi egyenlet :

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} + (k - 1) \frac{\partial v}{\partial x} - kv, \quad k := \frac{2r}{\sigma^2}. \quad (16.3)$$

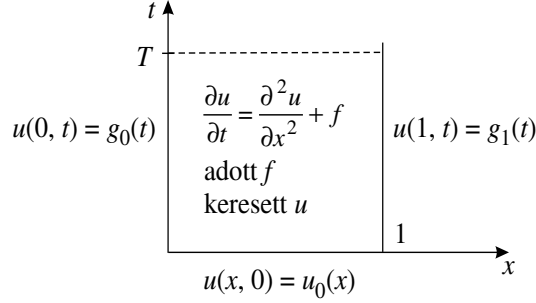
Ehhez most  $\tau = 0$ -nál adott a kezdeti feltétel, valamint  $x \rightarrow \pm\infty$  esetén a  $v$  viselkedése, keresett  $v(x, \frac{1}{2}\sigma^2 T)$ . Nem nehéz az egyenletet még tovább egyszerűsíteni, úgy hogy alakja

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} \quad (16.4)$$

lesz, ld. az 1. feladatot. (Hasonló transzformációval „eltüntethető” (16.1)-ből a szél hatásával kapcsolatos  $v \frac{\partial c}{\partial x_1}$  tag. (16.1) gyakorlati megoldásának szempontjából ilyen transzformáció viszont nem javasolható. \*\*\* 17.4.1-ben visszatérünk az (16.1), ill. (16.3) alakú egyenletek numerikus problémáira.) Ha közvetlenül a (16.2) egyenletet akarjuk diszkretizálni, akkor egyrészt az  $S$  független változó végtelen intervalluma gond (emiatt Laguerre-függvények bevetése célszerű), másrészt a  $\frac{\partial^2 V}{\partial S^2}$  legmagasabb derivált előtt álló  $\frac{1}{2}\sigma^2 S^2$  szorzó jelent problémát, mert  $S = 0$ -nál eltűnik (ami miatt a Laguerre-függvények súlyozott formája megfelelő, ld. pl. Ben-yu és Xiao-yong dolgozatát).

### 16.1.3 A hővezetési egyenlet

Tekintsünk olyan nyugvó, hővezető közeget, amelyben hő fejlődik ki  $f$  sűrűséggel időegységenként és egységterfogatonként. A közeg hővezetési tényezője



ábra 16.1: *Parabolikus vegyes feladat*

legyen  $k > 0$ , sűrűsége  $\rho > 0$ , hőmérséklete  $T$ , belső energiája  $E = c_p T$ , ahol  $c_p$  a hőkapacitás.

Az energiamegmaradási tételből azt lehet ekkor levezetni, hogy érvényes

$$\rho c_p \frac{\partial T}{\partial t} = \sum_{i=1}^3 \frac{\partial}{\partial x_i} \left( k \frac{\partial T}{\partial x_i} \right) + f = \operatorname{div}(k \operatorname{grad} T) + f. \quad (16.5)$$

Általában itt a  $\rho$ ,  $c_p$ ,  $k$  és  $f$  együtthatók mind a koordináták és az idő függvényei, nagy hőmérsékleti intervallumon viszont a  $T$ -től is függhetnek (amire 16.6-ban visszatérünk).

A legegyszerűbb esetben (homogén, izotróp közeg) a  $\rho$ ,  $c_p$  és  $k$  pozitív konstansok. Továbbá, ha egy koordináta irányban a vizsgált test vagy nagyon hosszú (rúd) vagy rövid (lemez), akkor eltekinthetünk a másik kettő iránytól (ld. a 16.3. pontot is). Ilyen speciális esetben (16.5) az egydimenziós, konstans együtthatós hővezetési egyenletre egyszerűsödik :

$$\frac{\partial T}{\partial t} = \lambda \frac{\partial^2 T}{\partial x^2} + f, \quad \text{ahol } \lambda := \frac{k}{\rho c_p}$$

a hőfokvezetési tényező.

Ha az  $x$  megfelelő transzformációjával az eredeti (mondjuk  $[0, L]$ ) intervallumot  $[0, 1]$ -re normáltuk, akkor új időegység bevezetésével a  $\lambda$ -t is el tudjuk tüntetni. Ezután a következő egyenlettel állunk szemben (amely  $f$  kivételével (16.4)-gyel egyezik meg), amelyben a keresett függvény most a szokásosabb  $u$ , és  $T$ -vel inkább az időintervallumot jelöljük :

$$\frac{\partial u}{\partial t} = \mathcal{L}u + f, \quad 0 < x < 1, \quad 0 < t \leq T, \quad \mathcal{L}u := \frac{\partial^2 u}{\partial x^2}. \quad (16.6)$$

Ennek egyértelmű megoldásához  $Q_T := (0, 1) \times (0, T]$ -ben az alábbi kezdeti-



és peremértékeket szükséges megadni :

$$t = 0, \quad 0 \leq x \leq 1 : \quad u(x, t) = u_0(x); \quad (16.7)$$

$$t > 0, \quad x = 0 : \quad u(x, t) = g_0(t); \quad (16.8)$$

$$t > 0, \quad x = 1 : \quad u(x, t) = g_1(t). \quad (16.9)$$

Az itt előírt (16.8)-(16.9) elsőfajú peremfeltételek helyett másod- és harmad-fajúak is kitűzhetők, v.ö. II. 11.2-vel. A (16.7)-(16.9) feltételek által érintett peremszakaszokat a (16.6)-(16.9) feladat *Dirichlet-peremének* nevezzük és  $\Gamma_1 = \Gamma_1(T)$ -vel jelöljük,

$$\Gamma_1 := \{t = 0, \quad 0 \leq x \leq 1\} \cup \{x = 0, \quad 0 < t \leq T\} \cup \{x = 1, \quad 0 < t \leq T\}. \quad (16.10)$$

magát a (16.6)-(16.9) feladatot a hővezetési egyenlet *vegyes feladatának* hívjuk, arra célozva, hogy itt kezdeti- és peremértékeket írtunk elő.

## 16.2 Elméleti tudnivalók a parabolikus egyenletek vegyes feladatairól

Elsőnek azt határozzuk meg, hogy mit értünk az (16.6)-(16.9) feladat megoldása alatt.

**Definíció.** Ha  $u$  olyan  $\overline{Q}_T$ -ben definiált folytonos függvény, amely folytonosan deriválható  $Q_T$ -ben (egyszer  $t$  és kétszer  $x$  szerint) :  $u \in C^{2,1}(Q_T) \cap C(\overline{Q}_T)$ , és amely a (16.6) egyenletet, valamint a (16.7)-(16.9) mellékfeltételeket kielégíti, akkor  $u$ -t az (16.6)-(16.9) feladat *klasszikus megoldásának* hívjuk.  $\square$

Ezután emlékeztetünk a következő eredményre (részletesebben ehhez a ponthoz ld. Simon és Baderko, vagy Ladüzenszkaja, Szolonnyikov és Uralceva vagy Friedman könyvét).

**16.1. Tétel** (parabolikus egyenlet maximumelve). Az (16.6)-(16.9) feladat minden klasszikus  $u$  megoldása rendelkezik a következő tulajdonsággal.

Ha  $f \leq 0$  (ill.  $f \geq 0$ ), akkor  $u$  maximális (minimális) értékét a  $\Gamma_1$  peremen veszi fel.

**Megjegyzések. 1.** Ha  $f \equiv 0$ , akkor eszerint

$$\min_{\Gamma_1} u \leq u(x, t) \leq \max_{\Gamma_1} u \quad \text{minden } (x, t) \in Q_T\text{-re.}$$

**2.** A tételből következik az unicitás : az (16.6)-(16.9) feladatnak legfeljebb egy klasszikus megoldása van.  $\square$

Az unicitást az alábbi eredmény is mutatja (v.ö. (15.11)-gyel 15.2-ben).

**16.2. Következmény** (összehasonlítási tétel). Legyen  $u_1$  és  $u_2$  a (16.6)-(16.9) feladat két klasszikus megoldása,  $f_i$  jobboldallal,  $g_{0i}$ ,  $g_{1i}$  peremértékekkel,  $u_{0i}$  kezdeti értékkel,  $i = 1, 2$ .

Teljesüljön

$$\begin{aligned} f_1(x, t) &\geq f_2(x, t), & (x, t) \in Q_T, \\ u_{01}(x) &\geq u_{02}(x), & 0 \leq x \leq 1, \\ g_{01}(t) &\geq g_{02}(t), & g_{11}(t) \geq g_{12}(t), & t > 0. \end{aligned}$$

Ekkor érvényes, hogy  $u_1(x, t) \geq u_2(x, t)$ ,  $(x, t) \in \overline{Q}_T$ .

**Bizonyítás.** Ld. a 2. feladatot!

**Megjegyzések. 1.** Legyen  $u \in C^{2,1}(Q_T) \cap C(\overline{Q}_T)$  a (16.6)-(16.9) feladat megoldása abban az esetben, amikor  $u_0(x)$ ,  $g_1(t)$ ,  $g_2(t)$ ,  $f(x, t) \geq 0$ . Ekkor  $u \geq 0$ .

**2.** A 16.1. tétel a (16.6)-(16.9) feladat klasszikus megoldásának következő becslését is vonja maga után :

$$|u(x, t)| \leq \max_{\Gamma_1} |u| + \frac{1}{2}x(1-x)\|f\|_{C(\overline{Q}_T)} \text{ minden } (x, t) \in \overline{Q}_T\text{-ra.} \quad (16.11)$$

Ennek belátásához bontsuk fel  $u$ -t mint  $u = v + w$ , ahol  $v$  a homogén egyenlet megoldása,

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2}, \quad (x, t) \in Q_T; \quad v = u, \quad (x, t) \in \Gamma_1,$$

és

$$\frac{\partial w}{\partial t} = \frac{\partial^2 w}{\partial x^2} + f, \quad (x, t) \in Q_T; \quad w = 0, \quad (x, t) \in \Gamma_1.$$

A 16.1. tétel szerint egyrészt

$$\min_{\Gamma_1} u \leq v \leq \max_{\Gamma_1} u.$$

Másrészt, legyen  $z := \frac{1}{2}x(1-x)\|f\|_{C(\overline{Q}_T)}$ , akkor  $z|_{\Gamma_1} \geq 0 = \pm w|_{\Gamma_1}$  és

$$\frac{\partial z}{\partial t} - \frac{\partial^2 z}{\partial x^2} = \|f\|_C \geq \pm f = \pm \left( \frac{\partial w}{\partial t} - \frac{\partial^2 w}{\partial x^2} \right).$$

Emiatt  $z \geq \pm w$ ,  $(x, t) \in \overline{Q}_T$ . Ekkor (16.11) következik a háromszög egyenlőtlenség segítségével :  $|u(x, t)| \leq |v(x, t)| + |w(x, t)| \leq \max_{\Gamma_1} |u| + z(x, t)$ .  $\square$

Tekintsünk most egy jóval általánosabb feladatot! Legyen  $\Omega \in \mathbb{R}^d$  véges tartomány, és  $Q_T := \Omega \times (0, T)$  adott  $T > 0$ -val. A parabolikus egyenlet általános alakja :

$$\frac{\partial u}{\partial t} = \mathcal{L}u + f(x, t), \quad (x, t) = (x_1, \dots, x_d, t) \in Q_T, \quad (16.12)$$

$$\mathcal{L}u := \sum_{i,j=1}^d \frac{\partial}{\partial x_i} \left( a_{ij}(x, t) \frac{\partial u}{\partial x_j} \right) + \sum_{i=1}^d a_i(x, t) \frac{\partial u}{\partial x_i} + a_0(x, t)u. \quad (16.13)$$

Ezzel az egyenlettel kapcsolatos a következő *vegyes feladat*. Oldjuk meg a (16.12) egyenletet az alábbi,  $t = 0$ -nál, ill. az  $S_T := \Gamma \times [0, T]$  peremen megadott kezdeti- és peremfeltételek mellett (ahol  $\Gamma := \partial\Omega$ ) :

$$(x, t) \in S_T : \quad u = 0, \quad t = 0, \quad x \in \Omega : \quad u = u_0(x). \quad (16.14)$$

A (16.12) egyenletet akkor hívjuk parabolikusnak, ha a (16.13) operátorának főrésze elliptikus :

$$0 \leq c_0 \sum_{i=1}^d \xi_i^2 \leq \sum_{i,j=1}^d a_{ij}(x, t) \xi_i \xi_j \leq c_1 \sum_{i=1}^d \xi_i^2 \text{ minden } (x, t) \in Q_T\text{-re.}$$

Ha itt  $c_0$  pozitív konstans, akkor *egyenletesen parabolikus egyenletről* beszélünk.

A (16.12)-(16.14) kezdetiérték-peremérték feladat klasszikus megoldása hasonlóan definiálható, mint az elliptikus esetben (ld. 15.2), és hasonló feltételek mellett a klasszikus megoldás létezik, egyértelmű és stabil módon függ a feladat adataitól (az  $S_T$  peremen inhomogén peremértékek is lehetnek :  $u|_{S_T} = g$ ). A (16.12)-(16.14) feladat tehát *korrekt kitűzésű* bizonyos (simasági) feltételek mellett, ld. Zeidler, vagy Friedman vagy Ladüzenszkaja, Szolonnikov és Uralceva könyveit.

A következő különbségek adódnak viszont. Egyrészt szükséges egy *kompatibilitási feltétel*, mert az  $u_0$  és  $g = u|_{S_T}$  függvényeket ((16.14)-ben  $g = 0$ ) csak egyeztetetten lehet megadni, ha azt kívánjuk, hogy az  $\{x \in \bar{\Omega}, t = 0\}$  és  $\bar{S}_T$  halmazok közös pontjaiban is folytonos legyen az  $u$  megoldás :

$$\frac{\partial g}{\partial t} \Big|_{t=0} = \mathcal{L}u_0 + f(x, 0).$$

Másrészt, az  $a_0$  együttható előjelére nem kell feltétel. Amennyiben ugyanis az  $\mathcal{L}$  operátornak éppen nulla sajátértéke volna, és így a vele megfogalmazott elliptikus feladatnak vagy nem lenne megoldása vagy az nem lenne

egyértelmű, akkor a parabolikus feladat esetén van még a kezdeti feltétel is és biztosítja az unicitást (ld. ehhez a 3. feladatot : az a tény, hogy egyszerű transzformációval nullává tehető  $a_0$  ha konstans, is ebbe az irányba mutat, v.ö. az 1. feladattal).

A parabolikus feladat része lehetnek másodfajú vagy harmadfajú peremfeltételek is, és ekkor is korrekt kitűzésű a vegyes feladat. Itt viszont, megint az elliptikus egyenlet megfelelő peremérték feladatával összehasonlítva, az tűnik fel, hogy a  $\frac{\partial u}{\partial n} + \sigma u = g$  peremfeltétel  $\sigma$  együtthatója negatív is lehet (hasonló okból mint ezelőtt az  $a_0$  pozitív).

**Definíció.** A (16.12)-(16.14) feladat *gyenge megoldásának* hívjuk az olyan  $u \in H^1[(0, T) \rightarrow H_0^1(\Omega)]$  függvényt, amelyre, minden  $\overline{Q_T}$ -ben definiált és ott végtelenül sokszor differenciálható, és az  $S_T$  perem környezetében eltűnő  $\varphi$  függvénnyel teljesül

$$\begin{aligned} \int_{Q_T} \left[ \frac{\partial u}{\partial t} \varphi + \sum_{i,j=1}^d a_{ij} \frac{\partial u}{\partial x_j} \frac{\partial \varphi}{\partial x_i} + \sum_{i=1}^d a_i \frac{\partial u}{\partial x_i} \varphi + a_0 u \varphi \right] dx dt & \quad (16.15) \\ & = \int_{Q_T} f(x, t) \varphi(x, t) dx dt. \quad \square \end{aligned}$$

**16.3. Tétel** (Ladüzsenszkaja, Uralceva; gyenge megoldás tulajdonságai). Amennyiben a (16.12) egyenlet egyenletesen parabolikus, együtthatói mérhetőek és korlátosak, azonkívül érvényesek a következő feltételek :

$$\int_0^T \left\| \frac{\partial a_{ij}(\cdot, \tau)}{\partial t} \right\|_{C(\overline{\Omega})} d\tau \leq M,$$

minden  $i, j$  indexre, valamint

$$f \in L_2(Q_T), \quad u_0 \in L_2(\Omega),$$

akkor a vegyes parabolikus feladatnak pontosan egy gyenge megoldása van, ezenkívül érvényes a következő becslés  $0 \leq t \leq T$ -re :

$$\|u(\cdot, t)\|_{L_2(\Omega)}^2 + 2c_0 \int_0^t \|u(\cdot, \tau)\|_{H_0^1(\Omega)}^2 d\tau \leq (1 + te^t) \{ \|u_0\|_{L_2(\Omega)}^2 + \|f\|_{L_2(Q_t)} \}. \quad (16.16)$$

Ha viszont  $u_0 \in H_0^1(\Omega)$ , akkor

$$c_0 \|u_x(\cdot, t)\|_{(L_2(\Omega))^d}^2 + \left\| \frac{\partial u}{\partial t} \right\|_{L_2(Q_T)} \leq \mu(t) \left\{ c_1 \|u_0\|_{H_0^1(\Omega)}^2 + \|f\|_{L_2(Q_t)} \right\}. \quad (16.17)$$

Itt a  $\mu(t) \geq 1$  folytonos  $t$ -ben és  $\int_0^t \left\| \frac{\partial a_{ij}}{\partial t} \right\|_{C(\Omega)} d\tau$ -tól függ.  $\square$

Az előző (16.16)-(16.17) becslések azt mutatják meg, hogy a gyenge megoldás folytonosan függ a vegyes feladat adataitól és ebben az értelemben stabil.

Mint  $H^1[(0, T) \rightarrow H_0^1(\Omega)]$  eleme a gyenge megoldás  $L_2(\Omega)$ -normája folytonos  $t$ -ben, a beágyazási tételek szerint. Ebben az értelemben a kezdeti feltételt is kielégíti :

$$\|u(\cdot, t) - u_0(\cdot)\|_{L_2(\Omega)} \rightarrow 0, \quad t \rightarrow 0.$$

A (16.16) egyenlőtlenség levezetéséhez a (16.15) relációba  $v = u$ -t helyettesítjük be, a (16.17) levezetéséhez pedig  $v = \frac{\partial u}{\partial t}$ -t, majd az eredményt alakítjuk át. Ugyanezt a technikát a parabolikus egyenletet approximáló differenciasémák stabilitási bizonyításánál fogjuk alkalmazni.

Szemléltetésnek a két módszert akkor mutatjuk be, amikor (16.12)-(16.14) a legegyszerűbb parabolikus vegyes peremérték-kezdetiérték feladat, azaz  $d = 1$ , (16.6),  $f = 0$ , (16.7)-(16.9),  $g_0 = g_1 = 0$ .

Legyen  $u$  ezen feladat  $H^1[(0, T) \rightarrow H^2(0, 1)]$ -beli megoldása. Szorozzuk meg a (16.6) egyenletet  $u$ -val és integráljunk  $(0, 1)$  felett (vagyis : helyettesítsük be  $\varphi = u$ -t (16.15)-be) !

$$\begin{aligned} \int_0^1 \frac{\partial u}{\partial t} u \, dx &= \int_0^1 \frac{\partial (\frac{1}{2}u^2)}{\partial t} \, dx \\ &= \int_0^1 \frac{\partial^2 u}{\partial x^2} u \, dx = \left[ \frac{\partial u}{\partial x} u \right]_0^1 - \int_0^1 \left( \frac{\partial u}{\partial x} \right)^2 \, dx. \end{aligned}$$

A homogén peremfeltételek miatt innen következik, hogy

$$\frac{1}{2} \frac{\partial}{\partial t} \int_0^1 u^2(x, t) \, dx = - \int_0^1 \left( \frac{\partial u}{\partial x} \right)^2 \, dx \leq 0,$$

vagyis, (16.16)-nak megfelelően

$$\|u(\cdot, t)\|_0^2 := \int_0^1 u^2(x, t) \, dx \leq \int_0^1 u^2(x, 0) \, dx = \|u_0\|_0^2.$$

A megoldás  $L_2$ -normája ezek szerint nem növekedhet.

A második út stabilitási becsléshez a (16.17) becsléssel kapcsolatban elmondottaknak felel meg : szorozzuk meg a (16.6) egyenletet  $\frac{\partial u}{\partial t}$ -vel és integráljunk  $(0, 1)$  felett (vagyis : helyettesítsük be  $\varphi = \frac{\partial u}{\partial t}$ -t (16.15)-be) ! Ekkor

$$\begin{aligned} 0 \leq \int_0^1 \left( \frac{\partial u}{\partial t} \right)^2 \, dx &= \int_0^1 \frac{\partial^2 u}{\partial x^2} \frac{\partial u}{\partial t} \, dx = \left[ \frac{\partial u}{\partial x} \frac{\partial u}{\partial t} \right]_0^1 - \int_0^1 \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial t \partial x} \, dx \\ &= - \int_0^1 \frac{1}{2} \frac{\partial}{\partial t} \left( \frac{\partial u}{\partial x} \right)^2 \, dx \end{aligned}$$

ahol megint a homogén peremfeltételeket használtuk, amelyekből

$$\frac{\partial u}{\partial t} \Big|_{x=0} = \frac{\partial u}{\partial t} \Big|_{x=1} = 0$$

következik. Ez azt jelenti, hogy

$$\frac{\partial}{\partial t} \int_0^1 \left( \frac{\partial u(x, t)}{\partial x} \right)^2 dx \leq 0,$$

vagyis

$$|u(\cdot, t)|_1^2 := \int_0^1 \left( \frac{\partial u(x, t)}{\partial x} \right)^2 dx \leq \int_0^1 \left( \frac{\partial u(x, 0)}{\partial x} \right)^2 dx = |u_0|_1^2.$$

Ezek szerint a megoldás  $H_0^1$ -normája sem növekedhet. Mindkét becslés azt mutatja, hogy a (16.6) hővezetési egyenlet megoldása, homogén (16.8)-(16.9) peremfeltételek mellett, folytonosan függ az  $u_0$  kezdetiértéktől.

### 16.3 Néhány hasznos fogás

Adott parabolikus egyenlet (mint pl. a (16.1)) elemzésének első lépéseként azt tanácsoljuk, hogy dimenziótlan alakra hozzuk az egyenletet. Ez azt jelenti, hogy minden mennyiséget tipikus értékére normalizáljuk. Ennek eredményeként ugyanis kiderül, hogy mitől függ valóban a probléma, milyen paraméter-kombinációkra érdemes figyelni. Szemléltessük a tanácsot a (16.1) példáján! Mint tipikus koncentrációt a kritikust vesszük :  $u := c/c_{krit}$ , új térkoordinátákként az  $y_i = x_i/L$ -t vezetjük be ( $i = 1, 2, 3$ ), a  $\tau$  új időt úgy vezetjük be, hogy a legmagasabb deriváltak előtt (az idő- és a második  $y$ -deriváltak előtt) 1 legyen az együttható :  $\tau = \frac{L^2}{D}t$ . Ezután az egyenlet új alakja, a  $D/L^2$ -tel való egyszerűsítés után :

$$\frac{\partial u}{\partial \tau} + \frac{vL}{D} \frac{\partial u}{\partial y_1} = \sum_{i=1}^3 \frac{\partial^2 c}{\partial y_i^2} + \frac{QL^2}{Dc_{krit}} \delta(x, 0).$$

Itt tehát két paraméter maradt :  $\frac{vL}{D}$  és  $\frac{QL^2}{Dc_{krit}}$ . Eszerint pl. az mondható el, hogy nem a  $D$  diffúzió együttható kis értéke miatt hanyagolhatók el esetleg a második deriváltak (16.1)-ben (amitől az egyenlet matematikailag gyökerezen változna meg, és nem kellene a tartomány összes peremén megadni az információt a koncentrációról), hanem a  $\frac{vL}{D} = P$  úgynevezett Pécelet-szám nagy értéke miatt.

Következően néhány olyan függvényt írunk fel, amely megoldja a (16.6) egyenletet  $f \equiv 0$  esetén :

a)  $u_0(x, t) = 1$ ,  $u_1(x, t) = x$ ,  $u_2(x, t) = x^2 + 2t$  (és további hasonló, magasabb fokszámú polinomok, ld. lent);

b)  $u = e^{\pm 2ax + 4a^2t}$ ;

c)  $u = \frac{1}{\sqrt{t}} e^{-\frac{x^2}{4t}}$  (ahol  $t > 0$ ).

Az utóbbi függvény segítségével a homogén (16.6) egyenlet Cauchy-feladatának megoldása írható fel, tehát az egész  $x$ -tengelyen kitűzött (16.6)-(16.7) feladat, ha  $u_0$  nem növekszik  $x$ -szel gyorsabban, mint  $c_1 e^{c_2 x^2}$  :

$$u(x, t) = \frac{1}{2\sqrt{\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-\xi)^2}{4t}} u_0(\xi) d\xi.$$

Az a) alatt felsorolt úgynevezett ( $k$ -adfokú) *hő-polinomok* általános alakja

$$u_k(x, t) = k! \sum_{j=0}^{[k/2]} \frac{x^{k-2j} t^j}{(k-2j)! j!},$$

ahol  $[r]$  az  $r$  egész része. Ezután, az egyenlet linearitása miatt, a

$$p(x, t) = \sum_{k=0}^n a_k u_k(x, t)$$

polinomokat is hő-polinomoknak nevezhetjük.

Ezeknek felhasználásával az (16.6)-(16.9),  $f \equiv 0$  probléma vegyes feladatának a közelítő megoldását Gauss-féle képletek segítségével lehet megadni rögzített  $(x^*, t^*) \in Q_T$ -re :

$$u(x^*, t^*) \approx \sum_{i=1}^N a_i^* u(x_i^*, t_i^*), \quad (16.18)$$

Az  $(x_i^*, t_i^*) \in \Gamma_1$  és  $a_i^*$  paraméterek meghatározására ( $i = 1, \dots, N$ ) követeljük, hogy (16.18) helyett legyen érvényes szigorú egyenlőség, amikor  $u$   $k$ -adfokú hő-polinom ( $0 \leq k < 2N$ ) :

$$u_k(x^*, t^*) = \sum_{i=1}^N a_i^* u_k(x_i^*, t_i^*), \quad k = 0, 1, \dots, 2N - 1.$$

Ez egy  $2N$ -dimenziós nemlineáris egyenletrendszer a (16.18) képlet  $2N$  paraméterének meghatározására. (Az  $(a_i^*, x_i^*, t_i^*)$  számhármakban hol a  $t_i^*$ , és

hol az  $x_i^*$  értéke következik abból, hogy  $(x_i^*, t_i^*) \in \Gamma_1$ . A másik – és az  $a_i^*$  – az egyenletrendszerből meghatározandó.) Ha az  $(x_i^*, t_i^*) \in \Gamma_1$  pontokat írjuk elő és csak az  $a_i^*$  együtthatókat keressük, akkor elég egy lineáris egyenletrendszert megoldani, de az eredmény nem Gauss-képlet.

Az így meghatározott konstansok csak a kiválasztott  $(x^*, t^*)$  párra jók.

Ez azt jelenti, hogy ilyen eljárással csak akkor célszerű dolgozni, amikor különböző peremértékekre (ill. kezdeti értékre) keresett a megoldás ugyanabban a pontban. A  $\Gamma_1$  nem kell, hogy éppen a (16.10) legyen, hanem  $(x(s), t(s))$  folytonos, nem önmagát metsző görbe lehet,  $(x(0), t(0)) = (0, t^*)$ ,  $(x(1), t(1)) = (1, t^*)$ ,  $t(s) < t^*$  egyébként. Az ötlet természetesen továbbfejleszthető úgy, hogy több pontban kapjuk az eredményeket.

A tapasztalatok szerint  $N = 5$  esetén már 4–5 pontos decimális jegyre számíthatunk a kapott eredményben.

A fejezet további részében csak olyan eljárásokkal foglalkozunk, amelyek a megoldás egész mezejét adják, tehát sok  $x$ - és  $t$ -értékre.

Mielőtt annak részleteibe merülünk, nem árt megjegyezni, hogy ezen az úton olyan részletes numerikus eredményekhez jutunk, amelyekhez esetleg nincs meg a megfelelő háttérünk : pl. vagy nincs olyan mérési adatunk, amely a numerikus eredmény helyességét ilyen részletességgel igazolná, vagy a kitűzött feladat valójában nem az egyenlet megoldása, hanem a megoldás valamelyik funkcionáljának (pl. egy pontbeli értékének vagy integráljának) a kiszámítása.

Ilyenkor meggondolandó, hogy nem célszerűbb-e csak néhány paraméterrel leírt közelítő megoldással dolgozni (mint fent (16.18)-ban)? Ennek elősegítésére gyakran hasznosak a *dimenziócsökkentő eljárások*.

Ha pl. megoldandó a

$$\frac{\partial u}{\partial t} = k_0 \Delta_3 u + f = k_0 \sum_{i=1}^3 \frac{\partial^2 u}{\partial x_i^2} + f \quad (16.19)$$

egyenlet az  $x = (x_1, x_2) \in \Omega$ ,  $x_3 \in (0, 1)$  tartományban és a megoldás  $x_3$ -irányú változása várhatóan nem jelentős vagy nem érdekes, vagy pedig nincs is adatunk az  $f$  forrástág  $x_3$ -irányú változásáról, csak  $x_3$  szerinti átlagértékével rendelkezünk, akkor célszerű  $u(x_1, x_2, x_3, t)$  helyett

$$\bar{u}(x_1, x_2, t) := \int_0^1 u(x_1, x_2, x_3, t) dx_3$$

kiszámítását célba venni. Ehhez, ha (16.19) mellett az  $x_3$ -irányú peremfeltételek

$$x_3 = 0 : -k_0 \frac{\partial u}{\partial x_3} + \sigma_0 u = g_0(t), \quad (16.20)$$

$$x_3 = 1 : k_0 \frac{\partial u}{\partial x_3} + \sigma_1 u = g_1(t), \quad (16.21)$$



a (16.19) egyenletet  $x_3$  felett integrálva kapjuk, hogy

$$\frac{\partial \bar{u}}{\partial t} = k_0 \Delta_2 \bar{u} + \varphi_2 = k_0 \sum_{i=1}^2 \frac{\partial^2 \bar{u}}{\partial x_i^2} + \varphi_2, \quad (16.22)$$

ahol

$$\begin{aligned} \varphi_2(x_1, x_2, t) &:= \int_0^1 f(x_1, x_2, x_3, t) dx_3 + g_0(t) + g_1(t) \\ &- \sigma_0 u(x_1, x_2, 0, t) - \sigma_1 u(x_1, x_2, 1, t). \end{aligned}$$

A (16.22) egyenlet lezárásának érdekében a következő közelítéseket használhatjuk, ha a kétdimenziós  $\Omega$  tartomány átmérője 1-hez képest nagy :

$$u(x_1, x_2, 0, t), u(x_1, x_2, 1, t) \approx \bar{u}(x_1, x_2, t). \quad (16.23)$$

Ezután a

$$\frac{\partial \bar{u}}{\partial t} = k_0 \Delta_2 \bar{u} - (\sigma_0 + \sigma_1) \bar{u} + \psi_2 \quad (16.24)$$

egyenlet megoldását keressük, amelyben

$$\psi_2(x_1, x_2, t) := \int_0^1 f(x_1, x_2, x_3, t) dx_3 + g_0(t) + g_1(t),$$

megfelelő kezdeti- és peremfeltételek mellett.

Vegyük észre, hogy másodfajú (16.20)-(16.21) peremfeltételek esetén (tehát amikor  $\sigma_0 = \sigma_1 = 0$ ) a (16.23) közelítésre nincs szükségünk, (16.22) következik (16.19)-ből és (16.20)-(16.21)-ből. Ekkor  $\bar{u}$  semmi más mint az  $u$  első Fourier-együtthatója, ha  $u$ -t sorbafejtjük a homogén másodfajú peremfeltételek mellett vizsgált  $-\frac{\partial^2}{\partial x_3^2}$  operátor sajátfüggvényei szerint. Így  $x_3$  irányában egydimenziós altérben dolgozunk, a konstans függvénynek megfelelően. Amennyiben pl.  $x_2$  irányban hasonló átlagolással élhetünk, akkor majd még kevesebb paraméterrel közelíthetjük a megoldást.

Ehhez a ponthoz ld. a 4\*. és 4. feladatokat is!

## 16.4 Parabolikus egyenletek megoldása súlyozott sémával

### 16.4.1 A súlyozott differenciaséma

Tekintsük a (16.6)-(16.9) vegyes feladatot az egydimenziós hővezetési (más néven : diffúziós) egyenlet számára!

A szokásos,  $h = 1/N$  lépéstávolságú  $\omega_h := \{x_i, i = 1, \dots, N-1\}$  rácson (ld. II. 11.4.1.) az  $x$  szerinti második deriváltat differencia-hányadossal helyettesítve, és  $u(x_i, t)$ -t  $y_i(t)$ -vel közelítve azt kapjuk, hogy

$$\dot{y}_i(t) = (y(t))_{\bar{x}x,i} + f_i(t), \quad 0 < t \leq T, \quad (16.25)$$

$$y_i(0) = u_0(x_i), \quad x_i := ih, \quad 1 \leq i \leq N-1, \quad (16.26)$$

ahol

$$y_h(t) = (y_1(t), \dots, y_{N-1}(t))^T, \quad y_0(t) = g_0(t), \quad y_N(t) = g_1(t),$$

$$f_i(t) = f(x_i, t),$$

és  $\dot{y}_i$  jelöli az  $y_i$  komponens deriváltját  $t$  szerint.

A (16.6)-(16.9) egyes feladat ezen visszavezetését a (16.25)-(16.26) közönséges differenciálegyenlet-rendszer kezdetiérték feladatára *egyenesek módszerének* hívjuk, mert az  $x = x_i = ih$  egyenesek mentén kell a rendszert megoldanunk. Azt is mondhatjuk, hogy (16.25)-(16.26) a (16.6)-(16.9) feladat *szemidiszkrétizációja*, mivel csak az egyik független változót diszkrétizáltuk. A (16.25)-(16.26) közönséges differenciálegyenlet-rendszert a következő alakba írhatjuk át (emlékeztetünk a 15.3.1-beli jelölésünkre, miszerint ugyanazt a diszkrét megoldást rácsfüggvényként  $y$ -nal jelöljük – pl. differenciakifejezésekben, vektorként viszont  $y_h$ -val) :

$$\dot{y}_h = -A_h^{(0)} y_h + q_h(t), \quad t > 0, \quad y_h(0) \text{ adott}, \quad (16.27)$$

ahol

$$A_h^{(0)} = \frac{1}{h^2} \text{tridiag}(-1, 2, -1)$$

a már II. 11.4.1-ből ismert  $M$ -mátrix és

$$\begin{aligned} q_h(t) &:= (q_1, \dots, q_{N-1})^T, & q_i &:= f_i, \quad 2 \leq i \leq N-2; \\ q_1 &:= f_1(t) + \frac{1}{h^2} g_0(t), & q_{N-1} &:= f_{N-1}(t) + \frac{1}{h^2} g_1(t). \end{aligned}$$

A rendszer  $-A_h^{(0)}$  mátrixának sajátértékei mind negatívak és kondíciószáma nagy :  $\text{cond}(A_h) \approx 1/(2h^2)$ , hiszen tudjuk, hogy  $8 \leq \lambda(A_h) \lesssim \frac{4}{h^2}$  (ld. I. 3.1.1., II. 11.4.6.).

Eszerint rendszerünk egy tipikus merev egyenletrendszer. Megoldását a II. 10.5-ben részletesen megtárgyalt módszerekkel próbálhatjuk meg, pl. a Gear- vagy implicit Runge–Kutta vagy Rosenbrock–módszerekkel. Ezek a módszerek jó szolgálatot tesznek, különösen akkor, amikor  $f = f(u, x, t)$  a

megoldástól is függ, viszont eléggé művelet- és tárigenyesek. Ez döntő szempont akkor, ha  $N \gg 1$ , ami többdimenziójú feladatok – pl. egy országrész feletti légszennyeződési szimuláció esetén – több százezer egyenletet jelenthet! Ezért most egyszerűbb (de megbízható) módszercsaláddal foglalkozunk, amelynek tár- és műveletigénye kisebb az előző módszerekénél. (A szemi-diszkretizáció vizsgálatára 16.4.6-ban és 16.8-ban térünk vissza, tanácsokat a többdimenziós feladatok megoldása során használandó módszerekhez 16.7.3-ban is adunk.)

A következő módszercsaládot tekintjük a (16.6)-(16.9) feladat approximációjára, amelyben (16.27)-hez képest az időderiváltat is diszkretizáltuk, tehát *teljes diszkretizációt* hajtottunk végre :

$$y_{t,i} = \sigma y_{\bar{x},i}^{j+1} + (1 - \sigma) y_{\bar{x},i}^j + \varphi_i^j, \quad (16.28)$$

$$\varphi_i^j := f_i^{j+1/2}, \quad i = 1, 2, \dots, N - 1,$$

$$y_0^{j+1} = g_0(t_{j+1}), \quad y_N^{j+1} = g_1(t_{j+1}), \quad j = 0, 1, \dots, m - 1, \quad (16.29)$$

$$y_i^0 = u_0(x_i), \quad 0 \leq i \leq N. \quad (16.30)$$

Itt  $0 \leq \sigma \leq 1$  a *stabilitási paraméter*, vagy más néven a (16.28)-(16.30) differenciaséma *súlya*, maga (16.28)-(16.30) a *súlyozott differenciaséma*. (16.28)-ban használtuk az

$$y_t = y_{t,i} = y_{t,i}^j := (y_i^{j+1} - y_i^j) / \tau$$

jelölést, ahol  $\tau := T/m$  az időlépés,  $m \geq 1$ , és

$$f_i^{j+1/2} = f(x_i, t_j + \frac{\tau}{2}), \quad y_i^j := y(x_i, t_j) \approx u(x_i, t_j), \quad t_j := j\tau.$$

A térbeli  $\omega_h$  rács mintájára bevezethetjük az  $\omega_\tau$  időbeli rácsot is,

$$\omega_\tau := \{t_j, j = 1, \dots, m\}, \quad \bar{\omega}_\tau := \{t_j, j = 0, \dots, m\}. \quad (16.31)$$

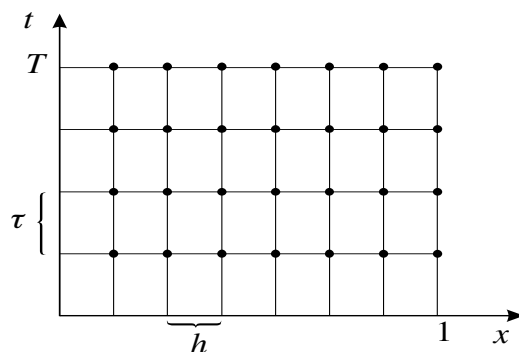
Ezután az összes téridő-rács pont halmazát  $\bar{\omega}_{h,\tau} := \bar{\omega}_h \times \bar{\omega}_\tau$ -val jelöljük, ahol  $\bar{\omega}_h = \omega_h \cup \gamma_h$ , a perempontok  $\gamma_h := \{x_0, x_N\}$  halmazával.

A matematikai lényegét gyakran jobban mutatja az indexmentes felírás-mód, amelyhez az

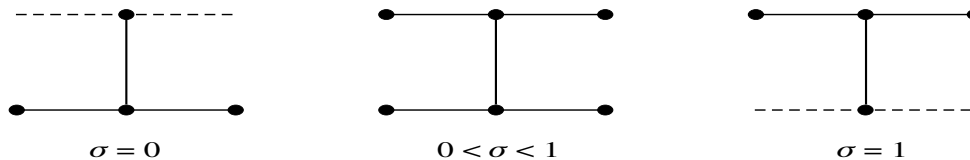
$$y := y_i^j, \quad \hat{y} := y_i^{j+1}, \quad \sigma y_i^{j+1} + (1 - \sigma) y_i^j = \sigma \hat{y} + (1 - \sigma) y =: y^{(\sigma)} \quad (16.32)$$

képletek segítségével megyünk át. Ekkor (16.28)-(16.30) az alábbi alakot veszi fel (ahol  $g(x, t) = g_0(t)$ , ha  $x = 0$  és  $g(x, t) = g_1(t)$ , ha  $x = 1$ ) :

$$\begin{aligned} y_t &= (y^{(\sigma)})_{\bar{x}} + f, & x \in \omega_h, \\ y &= g, & x \in \gamma_h, \quad j = 0, 1, \dots, m - 1; \\ y^0 &= u_0, & x \in \bar{\omega}_h. \end{aligned}$$



ábra 16.2:  $\omega_{h,\tau}$ -rács



ábra 16.3: Súlyozott séma differencia-csillagjai

A 10. fejezetben már vizsgált módszerek közül (16.28)-ban felismerjük a fenti (16.27) rendszerre alkalmazott explicit ( $\sigma = 0$ ) és implicit ( $\sigma = 1$ ) Euler-módszert, valamint a trapézszabályt ( $\sigma = \frac{1}{2}$ ). (16.28)-cal szinte azonos módszerrel már 10.5.2-ben is találkoztunk.

Itt, parabolikus egyenletek numerikus megoldásával kapcsolatban, a következő hagyományos elnevezéseket használjuk :

$$\sigma = \begin{cases} 0, & \text{explicit séma,} \\ \frac{1}{2}, & \text{Crank-Nicolson séma,} \\ 1, & \text{(tiszván) implicit séma.} \end{cases}$$

Míg  $\sigma = 0$  esetén közvetlenül az

$$y_i^{j+1} = y_i^j + \tau \left( y_{xx,i}^j + f_i^{j+1/2} \right)$$

képletből számítható ki a megoldás,  $\sigma \neq 0$  esetén tridiagonális egyenletrendszer megoldása szükséges (ami a rövidített Gauss-eliminációval  $O(N)$  művelettel kivitelezhető, ld. I. 1.3.9.). Ezért az utóbbi esetben implicitnek hívjuk a (16.28)-(16.30) sémát. • A súlyozott differenciaséma programozásra kész képleteit a „Hipy” program segédszövegeiben találhatjuk.

### 16.4.2 A súlyozott differenciaséma képlethibája

Taylor–sorfejtéssel tisztázzuk a súlyozott differenciaséma approximációjának rendjét  $h$ -ban és  $\tau$ -ban :

Egyrészt (ld. II. 11.4.1., (30)), ha az  $u$  függvény  $x$  szerint négyszer folytonosan differenciálható, akkor

$$u_{\bar{x},i} = \frac{\partial^2 u}{\partial x^2}(x_i, t) + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(x_i + \vartheta_i h, t), \quad |\vartheta_i| < 1. \quad (16.33)$$

Másrészt (ld. II. 11.4.2., (46)), ha  $u$  a  $t$  szerint háromszor folytonosan differenciálható, akkor

$$u_t = \frac{\partial u}{\partial t}(x, t_{j+1/2}) + \frac{\tau^2}{24} \frac{\partial^3 u}{\partial t^3}(x, t_{j+1/2} + \vartheta_j \tau), \quad |\vartheta_j| < \frac{1}{2},$$

és ekkor az is következik, hogy

$$\begin{aligned} u^{(\sigma)} &= u(x, t_{j+1/2}) + \tau(\sigma - \frac{1}{2}) \frac{\partial u}{\partial t}(x, t_{j+1/2}) \\ &\quad + \frac{\tau^2}{8} \frac{\partial^2 u}{\partial t^2}(x, t_{j+1/2} + \eta \tau), \quad |\eta| < \frac{1}{2}. \end{aligned} \quad (16.34)$$

Ezekből a sorfejtésekből a súlyozott differenciaséma

$$\psi_i^j := \begin{cases} u_{t,i} - (u^{(\sigma)})_{\bar{x},i} - \varphi_i^j, & \text{ha } (x_i, t_j) \in \omega_{h,\tau}, \\ 0, & \text{ha } (x_i, t_j) \in \gamma_h \times \omega_\tau, \\ 0, & \text{ha } (x_i, t_j) \in \bar{\omega}_h \times \{t = 0\} \end{cases} \quad (16.35)$$

képlethibájára következik, hogy

$$\begin{aligned} \psi_i^j &= \left( \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - f \right)_i^{j+1/2} + \left( f_i^{j+1/2} - \varphi_i^j \right) \\ &\quad - \left[ \tau(\sigma - \frac{1}{2}) \frac{\partial^3 u}{\partial x^2 \partial t}(x_i, t_{j+1/2}) + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(x_i + \vartheta_i h, t_{j+1/2}) \right] \\ &\quad + O(\tau^2 + h^4), \quad (x_i, t_j) \in \omega_{h,\tau}. \end{aligned} \quad (16.36)$$

Ez utóbbi eredményt annak figyelembevételével kapjuk, hogy  $2\tau h^2 \leq \tau^2 + h^4$ .

Ha most  $u$  a (16.6)-(16.9) feladat megoldása, akkor (16.36)-ból azt olvaszuk le, hogy a képlethiba  $\psi_i^j = O(\tau|\sigma - \frac{1}{2}| + \tau^2 + h^2)$ , ha  $u \in C^{4,3}(\bar{Q}_T)$ , azaz ha a megoldás  $x$  szerint négyszer és  $t$  szerint háromszor folytonosan differenciálható a  $[0, 1] \times [0, T] =: Q_T$  megoldási tartományban, és ha  $\varphi_i^j = f_i^{j+1/2}$

vagy  $\varphi_i^j = f_i^{j+\sigma}$ . De  $\psi_i^j = O(\tau + h^2)$  már akkor is igaz, ha  $u \in C^{4,2}$ , és ehhez  $\varphi_i^j = f_i^j$  is megfelelő.

Végül az  $u \in C^{6,3}$  esetben (16.33) helyett érvényes

$$u_{\bar{x},i} = \left( \frac{\partial^2 u}{\partial x^2} + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4} \right)_i + \frac{h^4}{360} \frac{\partial^6 u}{\partial x^6}(x_i + \eta_i h),$$

és ekkor a (16.36) szögletes zárójelében álló mennyiség

$$\left[ \tau(\sigma - \frac{1}{2}) \frac{\partial^3 u}{\partial x^2 \partial t} + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4} \right]_i^{j+1/2} =: R_h$$

-ra változik, amit a (16.6) egyenlet segítségével tovább alakíthatjuk :

$$R_h = \left( \left[ \tau(\sigma - \frac{1}{2}) + \frac{h^2}{12} \right] \frac{\partial^3 u}{\partial x^2 \partial t} - \frac{h^2}{12} \frac{\partial^2 f}{\partial x^2} \right)_i^{j+1/2}.$$

Ez azt jelenti, hogy ekkor  $\psi_i^j = O(\tau^2 + h^4)$ , amennyiben

$$\begin{aligned} \sigma = \sigma_* &:= \frac{1}{2} - \frac{h^2}{12\tau}, \\ \varphi_i^j = (\varphi_*)_i^j &:= \left( f + \frac{h^2}{12} f_{\bar{x}\bar{x}} \right)_i^{j+1/2} \\ &= \frac{1}{12} (f_{i+1} + 10f_i + f_{i-1})^{j+1/2}. \end{aligned} \quad (16.37)$$

Ha ugyanis  $u \in C^{6,3}$ , akkor (16.6) szerint  $f \in C^{4,2}$ , és ekkor (16.33) mintájára  $\left( \frac{\partial^2 f}{\partial x^2} \right)_i = f_{\bar{x}\bar{x},i} + O(h^2)$ .

Foglaljuk össze eredményeinket! A súlyozott differenciaséma képlethibája

$$\psi_i^j = \begin{cases} O(\tau + h^2), & \text{ha } u \in C^{4,2} \text{ és } \varphi_i^j = f_i^j, \\ O(\tau|\sigma - \frac{1}{2}| + \tau^2 + h^2), & \begin{cases} \text{ha } u \in C^{4,3} \text{ és } \varphi_i^j = f_i^{j+1/2} \\ \text{vagy } \varphi_i^j = f_i^{j+\sigma}, \end{cases} \\ O(\tau^2 + h^4), & \begin{cases} \text{ha } u \in C^{6,3}, \sigma = \sigma_* \\ \text{és } \varphi_i^j = (\varphi_*)_i^j. \end{cases} \end{cases} \quad (16.38)$$

Ennek speciális esete az, hogy a Crank–Nicolson séma képlethibája  $O(\tau^2 + h^2)$ , ha  $u \in C^{4,3}$ . Egy további speciális eset kidolgozása az 5. feladat tárgya. A  $\sigma = \sigma_*$ -gal jellemzett sémát *magasabbrendű sémának* hívjuk. Ekkor a  $\sigma$  negatív is lehet.

Ehhez a ponthoz ld. a 6. feladatot is.

### 16.4.3 A súlyozott differenciaséma stabilitása a szimmetrikus esetben

Vizsgáljuk most a súlyozott séma stabilitását! Ugyanis egyrészt a tapasztalat (7. feladat) megmutatja, hogy vannak gondok, másrészt II. 10.5.3. szerint (ld. a 45. feladatot II. 151. o.-n, valamint a 8.a feladatot is), amikor  $\sigma < \frac{1}{2}$ , számítanunk kell egy  $\tau L \leq \text{const}$  alakú stabilitási feltétellel, ahol  $L$  a feladat Lipschitz-állandója, amely (16.27) esetén  $4/h^2$ .

Ha pl.  $\sigma = 0$ , akkor

$$\begin{aligned}\hat{y}_i &= y_i + \frac{\tau}{h^2}(y_{i+1} - 2y_i + y_{i-1}) + \tau\varphi_i \\ &= \frac{\tau}{h^2}(y_{i+1} + y_{i-1}) + \left(1 - \frac{2\tau}{h^2}\right)y_i + \tau\varphi_i, \quad 1 \leq i \leq N-1.\end{aligned}\quad (16.39)$$

Itt többször szerepel a  $\gamma := \frac{\tau}{h^2}$  (néha *Courant-számnak* elnevezett) paraméter (amely  $\frac{\tau k}{\rho c_p h^2}$ -re változik, ha a (16.5) egyenletből indulunk ki és annak együtt-hatói konstansok). Ettől a paramétertől döntően függhet a számítási eredmény, amit az alábbi lemmával a maximum normára vonatkozólag megmutatunk. Ezen norma definíciója

$$\|y_h\|_{C(\bar{\omega}_h)} := \max_{0 \leq i \leq N} |y_i|, \quad \text{ill.} \quad \|y_h\|_{C(\omega_h)} := \max_{1 \leq i \leq N-1} |y_i|. \quad (16.40)$$

Analóg módon  $\|y_h\|_{C(\gamma_h)} := \max(|y_0|, |y_N|)$ . Ezenkívül (16.10) mintájára még a  $\Gamma_{h,t_j} = \bar{\omega}_{h,\tau} \cap \partial Q_{t_j}$  „diszkrét” Dirichlet-peremet” is vezetjük be, azaz

$$\Gamma_{h,t_j} := \{(x_i, 0), \quad 0 \leq i \leq N\} \cup \{(0, t_k), (1, t_k), \quad 1 \leq k \leq j\}.$$

Ezzel a jelöléssel tehát  $\Gamma_{h,t_m} = \Gamma_{h,T}$ , és ha  $y_h$  a (16.28)-(16.30) séma megoldása, akkor

$$\|y_h\|_{C(\Gamma_{h,t_j})} = \max \left( \max_{x \in \bar{\omega}_h} |u_0(x)|, \max_{1 \leq k \leq j} |g_0(t_k)|, \max_{1 \leq k \leq j} |g_1(t_k)| \right). \quad (16.41)$$

**16.4. Lemma** (az explicit séma maximumnormabeli stabilitása). Ha

$$\gamma = \frac{\tau}{h^2} \in \left(0, \frac{1}{2}\right], \quad (16.42)$$

akkor a (16.28)-(16.30),  $\sigma = 0$  séma stabil : érvényes az

$$\|y_h^j\|_{C(\bar{\omega}_h)} \leq \|y_h\|_{C(\Gamma_{h,t_j})} + \sum_{k=0}^{j-1} \tau \|\varphi_h^k\|_{C(\omega_h)} \quad (16.43)$$

becslés. Ha viszont  $\gamma > \frac{1}{2}$  rögzített, akkor a séma nem lehet stabil a maximum normában.

**Bizonyítás.** (16.42) teljesülése esetén (16.39)-ből kapjuk közvetlenül

$$\begin{aligned} |\hat{y}_i| &\leq \gamma(|y_{i+1}| + |y_{i-1}|) + (1 - 2\gamma)|y_i| + \tau|\varphi_i| \\ &\leq \max(|y_{i+1}|, |y_i|, |y_{i-1}|) + \tau|\varphi_i|, \end{aligned}$$

és így

$$\begin{aligned} \|\hat{y}_h\|_{C(\omega_h)} &\leq \|y_h\|_{C(\bar{\omega}_h)} + \tau\|\varphi_h\|_{C(\omega_h)} \\ &= \max(\|y_h\|_{C(\gamma_h)}, \|y_h\|_{C(\omega_h)}) + \tau\|\varphi_h\|_{C(\omega_h)}. \end{aligned}$$

Ezért, most a peremet is figyelembe véve (és a felső indexet is jelölve),

$$\begin{aligned} \|y_h^{j+1}\|_{C(\bar{\omega}_h)} &\leq \max(\|y_h^{j+1}\|_{C(\gamma_h)}, \|y_h^j\|_{C(\gamma_h)}, \|y_h^j\|_{C(\omega_h)}) + \tau\|\varphi_h^j\|_{C(\omega_h)} \\ &\leq \max(\|y_h^{j+1}\|_{C(\gamma_h)}, \|y_h^j\|_{C(\gamma_h)}, \|y_h^{j-1}\|_{C(\bar{\omega}_h)}) \\ &\quad + \tau\|\varphi_h^{j-1}\|_{C(\omega_h)} + \tau\|\varphi_h^j\|_{C(\omega_h)} \\ &\leq \dots \leq \max(\|y_h^{j+1}\|_{C(\gamma_h)}, \dots, \|y_h^1\|_{C(\gamma_h)}, \|y_h^0\|_{C(\bar{\omega}_h)}) + \sum_{k=0}^j \tau\|\varphi_h^k\|_{C(\omega_h)}. \end{aligned}$$

Innen következik a (16.43) becslés. A lemma második részének bizonyításához elég, ha speciális esetet vizsgálunk. Legyen tehát  $\gamma > \frac{1}{2}$  rögzített. Továbbá, legyen

$$\begin{aligned} y_i^0 &= (-1)^i, \quad i = 0, 1, \dots, N, \\ g_0^j &= (-1)^j, \quad g_1^j = (-1)^{N+j}, \quad j = 0, 1, \dots, \end{aligned}$$

és  $\varphi \equiv 0$ . Az  $y_i^j$  megoldási értékeket csak  $0 \leq i+j \leq N$ -re fogjuk kiszámítani. (Így a peremértékekre valójában nincs is szükségünk.) (16.39)-ből kapjuk  $j = 1$ -re

$$y_i^1 = (1 - 4\gamma)(-1)^i = (4\gamma - 1)(-1)^{i+1}, \quad 1 \leq i \leq N - 1.$$

Ilymódon folytatva következik

$$y_i^j = (4\gamma - 1)^j (-1)^{i+j}, \quad j \leq i \leq N - j,$$

tehát amikor  $\tau, h \rightarrow 0$ ,  $j = \lfloor \frac{N}{2} \rfloor \rightarrow \infty$  ( $\gamma = \text{const} > \frac{1}{2}$  mellett), akkor  $|y_j^j| = (4\gamma - 1)^j \rightarrow \infty$ . Itt az  $y_j^j$  érték tartozik a rács  $(jh, j\tau) = (jh, jh^2\gamma) = (x_j, x_j h\gamma)$  pontjához, ami azt jelenti, hogy egyre közelebb az  $x$ -tengelyhez kapunk tetszőlegesen nagy eredményeket, annak ellenére, hogy  $\|y_h\|_{C(\Gamma_{h,t_j})} = 1$ .  $\square$



Tehát ahhoz, hogy a diszkrét  $C$ -normában stabil legyen az explicit séma, kemény árat kell fizetni : a (16.42) feltétel azt jelenti, hogy az időlépések száma  $m \geq 2T/h^2 = 2TN^2$ , azaz összesen  $O(N^3)$  művelet szükséges a (16.28)-(16.30),  $\sigma = 0$  egyenletek stabil megoldásához. Ehhez ld. a 9. feladatot is.

Ezután a speciális eset megtárgyalása után lényegesen általánosabb stabilitási vizsgálattal foglalkozunk.

Ehhez a már II. 11.4.6-ban bevezetett

$$(u_h, v_h)_{(0,h)} := \sum_{i=1}^{N-1} u_i v_i h, \quad (u_h, v_h)_{(1,h)} := \sum_{i=1}^N u_{\bar{x},i} v_{\bar{x},i} h \quad (16.44)$$

skaláris szorzatokat használjuk. A hozzátartozó normákat  $\|\cdot\|_{(0,h)}$ -val és  $\|\cdot\|_{(1,h)}$ -val jelöljük és mint  $L_2(\omega_h)$ -normát, ill.  $H_0^1(\omega_h)$ -normát is említjük.

A (16.28) súlyozott sémát a következő, általánosabb alakba írhatjuk át abban az esetben, amikor  $y_h := (y_1, \dots, y_{N-1})^T$ , a peremfeltételek homogének (akár csak formálisan, mert előzetesen elimináltuk az inhomogén peremfeltételeket) :

$$B_h y_t + A_h y_h^j = \varphi_h^j, \quad j = 0, 1, \dots, m-1, \quad y_h^0 \text{ adott.} \quad (16.45)$$

A (16.45) egyenletet a *kétréteges séma normálalakjának* hívjuk. Konkrét példánkban  $A_h$  a már többször említett  $A_h^{(0)}$  mátrix :

$$(A_h y_h)_i = (A_h^{(0)} y_h)_i = -y_{\bar{x},i}, \quad (16.46)$$

ahol a differenciahányados homogén peremfeltételeknek megfelelően definiált, és  $\varphi_i^j = f_i^{j+1/2}$ ,  $2 \leq i \leq N-2$ ,  $\varphi_1^j = f_1^{j+1/2} + g_0^{(\sigma)}/h^2$ ,  $\varphi_{N-1}^j = f_{N-1}^{j+1/2} + g_1^{(\sigma)}/h^2$ .

A továbbiakban  $A_h$ -ról csupán annyit teszünk fel ((16.46)-nak megfelelően), hogy

$$A_h = A_h^T > 0, \quad (16.47)$$

tehát  $A_h$  legyen szimmetrikus és pozitív definit a  $(\cdot, \cdot)_{(0,h)}$  skaláris szorzatban. Továbbá, példánkban

$$B_h = I_h + \tau \sigma A_h, \quad (16.48)$$

ugyanis (16.28) jobboldala felírható mint  $-(A_h^{(0)} y_h^{(\sigma)} - \varphi_h)_i$ , ahol  $y_h^{(\sigma)}$ -t (16.32) definiálja, de az egyenlő a következő kifejezéssel :

$$y_h^{(\sigma)} = y_h^j + \sigma \tau y_t.$$

Egyelőre  $B_h$ -ről is csak azt tételezzük fel, hogy

$$B_h = B_h^T > 0, \quad (16.49)$$

ami konkrét példánkban nyilván teljesül és a (16.45) egyenletek egyértelmű megoldhatóságát biztosítja. Ugyanis  $B_h$  az a mátrix, amellyel minden lépésben kell egyenletrendszer megoldanunk :

$$B_h y_h^{j+1} = (B_h - \tau A_h) y_h^j + \tau \varphi_h^j.$$

Most vezethetjük be az alapvető stabilitási definíciókat, amelyekhez a (16.43) becslés jó szemléltetés.

**Definíció.** Legyen (16.45)-ben  $B_h$  invertálható és  $y_h^j$  a kétréteges séma megoldása  $\varphi_h^j = 0$  esetén,  $j = 0, \dots, m$ . Ekkor azt mondjuk, hogy a (16.45) séma *stabil a kezdetiértékekre nézve*, ha tetszőleges  $m = T/\tau$ -ra,  $N = 1/h$ -ra és  $y_h^0$ -ra igaz az

$$\|y_h^j\|_1 \leq M_1 \|y_h^0\|_1$$

alakú becslés alkalmas  $\|\cdot\|_1$  normával, és olyan  $M_1$  konstanssal, amely nem függ  $\tau$ -tól vagy  $h$ -tól.  $\square$

Az „alkalmas” normákhoz ld. II. 10.2.2. A norma (16.43)-ban a maximum norma (amely alkalmas) és  $M_1 = 1$ .

**Definíció.** Legyen (16.45)-ben  $B_h$  invertálható és  $y_h^j$ ,  $j = 0, \dots, m$ , a kétréteges séma megoldása  $y_h^0 = 0$  esetén.

Ekkor azt mondjuk, hogy a (16.45) séma *stabil a jobboldalra nézve*, ha tetszőleges  $m$ -re,  $N$ -re és  $\varphi_h^j$ -re ( $j = 0, \dots, m-1$ ) igaz az

$$\|y_h^j\|_2 \leq M_2 \|\varphi_h^{(j)}\|_3$$

alakú becslés alkalmas  $\|\cdot\|_2$  normával, a

$$\varphi_h^{(j)} := (\varphi_h^0, \dots, \varphi_h^{j-1})$$

blokkvektor alkalmas  $\|\cdot\|_3$  normájával, és olyan  $M_2$  konstanssal, amely nem függ a diszkretizációtól.  $\square$

(16.43)-ban  $\|\cdot\|_1 = \|\cdot\|_2 = \|\cdot\|_{C(\overline{\omega}_h)}$ ,  $M_2 = 1$  és

$$\|\varphi_h^{(j)}\|_3 = \sum_{k=0}^{j-1} \tau \|\varphi_h^k\|_{C(\omega_h)}.$$

A stabilitási vizsgálat előkészítéséhez a (16.45) egyenletünket skalárisan megszorozzuk  $2\tau y_t = 2(y_h^{j+1} - y_h^j)$ -vel (v.ö. a 16.2. pontban a (16.17) egyenlőség kapcsán elmondottakkal) :

$$2\tau (B_h y_t, y_t)_{(0,h)} + 2(A_h y_h^j, y_h^{j+1} - y_h^j)_{(0,h)} = 2\tau (\varphi_h, y_t)_{(0,h)}. \quad (16.50)$$

Itt a második tagot az  $2a(b-a) = -(b-a)^2 + b^2 - a^2$  azonosság alapján alakítjuk át, kihasználva, hogy  $A_h$  szimmetrikus :

$$\begin{aligned} 2(A_h y_h^j, y_h^{j+1} - y_h^j)_{(0,h)} &= -(A_h(y_h^{j+1} - y_h^j), y_h^{j+1} - y_h^j)_{(0,h)} \\ &\quad + (A_h y_h^{j+1}, y_h^{j+1})_{(0,h)} - (A_h y_h^j, y_h^j)_{(0,h)} \\ &= -\tau^2 (A_h y_t, y_t)_{(0,h)} \\ &\quad + (A_h y_h^{j+1}, y_h^{j+1})_{(0,h)} - (A_h y_h^j, y_h^j)_{(0,h)}. \end{aligned}$$

(16.50)-ből tehát az következik, hogy

$$2\tau \left( \left( B_h - \frac{\tau}{2} A_h \right) y_t, y_t \right)_{(0,h)} + \|y_h^{j+1}\|_A^2 - \|y_h^j\|_A^2 = 2\tau (\varphi_h, y_t)_{(0,h)}, \quad (16.51)$$

ahol

$$\|y_h\|_A = (A_h y_h, y_h)_{(0,h)}^{1/2} = \|A_h^{1/2} y_h\|_{(0,h)}$$

az  $A_h$ -hoz tartozó *energetikai norma*.  $A_h = A_h^{(0)}$  esetén ld. ehhez II-ben a 11.3. lemmát. Ez a norma az alábbi stabilitási eredményben most mint a stabilitás definíciójának 1-es normája szerepel, és az  $M_1$  konstans értéke 1.

**16.5. Tétel** (Szamarszkij; kétréteges séma kezdetiértékek szerinti stabilitása). Teljesüljön (16.47) és (16.49), legyen  $\varphi_h \equiv 0$ , az  $A_h$  és  $B_h$  mátrixok ne függjenek  $j$ -től. Ekkor

$$B_h \geq \frac{\tau}{2} A_h \quad (16.52)$$

szükséges és elégséges ahhoz, hogy érvényes legyen a

$$\|y_h^j\|_A \leq \|y_h^0\|_A \quad (16.53)$$

becslés. Ekkor tehát a kezdetiértékekre nézve stabil a (16.45) séma az energetikai normában.

[Kommentár. A (16.52) feltétel azt jelenti, hogy

$$(B_h v_h, v_h)_{(0,h)} \geq \frac{\tau}{2} (A_h v_h, v_h)_{(0,h)}$$

minden  $v_h$ -re.]

**Bizonyítás.** Tekintsük tehát a homogén (16.45) egyenletet. Ekkor az következik (16.51)-ből, hogy

$$\|y_h^{j+1}\|_A^2 - \|y_h^j\|_A^2 = -2\tau \left( \left( B_h - \frac{\tau}{2} A_h \right) y_t, y_t \right)_{(0,h)} \leq 0,$$

azaz

$$\|y_h^{j+1}\|_A \leq \|y_h^j\|_A,$$

ahonnan (16.53) azonnal adódik.

Ha most megfordítva (16.53) igaz minden  $j$ -re és  $y_h^0$ -ra, akkor

$$\left( \left( B_h - \frac{\tau}{2} A_h \right) v_h, v_h \right) = \frac{1}{2\tau} (\|y_h^0\|_A^2 - \|y_h^1\|_A^2) \geq 0,$$

ahol  $v_h := y_t$ . Csupán azt kell még belátnunk, hogy itt  $v_h$  tetszőleges vektor lehet. De ez (16.45)-ből következik:  $v_h = -B_h^{-1} A_h y_h^0$ , mivel  $A_h$  és  $B_h$  regulárisak (16.47) és (16.49) miatt. Tehát a (16.52) stabilitási feltétel szükséges is.  $\square$

**Megjegyzések. 1.**  $B_h$ -ről elég lett volna feltenni, hogy reguláris. Azután, ha teljesül (16.47), a (16.52) feltétel szükséges és elégséges a kezdetiértékek szerinti stabilitáshoz.

**2.** A

$$by_t + ay^j = 0$$

egyenletet (ahol  $y$  nem rácsfüggvény vagy vektor, hanem skalár és  $b, a > 0$ ) az

$$y^{j+1} = \left( 1 - \frac{a}{b}\tau \right) y^j$$

alakba írhatjuk át, és így következik

$$|y^{j+1}| \leq |y^j| \leq |y^0|$$

pontosan akkor, amikor  $-1 \leq 1 - \frac{a}{b}\tau \leq 1$ , azaz  $b \geq \frac{\tau}{2}a$ .

**3.** A  $B_h \geq \frac{\tau}{2}A_h$  feltétel azért előnyös, mert változó együtthatójú – sőt többdimenziós – egyenletekre is alkalmazható, és a mátrixok sajátértékeinek kiszámítását nem követeli.

**4.** A végeselem módszer automatikusan biztosítja az  $A_h = A_h^T > 0$ ,  $B_h = B_h^T > 0$  feltételeket, ha a  $\frac{\partial u}{\partial t} = \mathcal{L}u$  parabolikus egyenlet elliptikus  $\mathcal{L}$  operátora az adott peremfeltételek mellett önadjungált, ld. 15.7.1. és lejjebb 16.4.6. A változó együtthatójú  $\mathcal{L}$  megfelelő differencia approximációját ld. 16.5.1-ben és 15.5.2-ben.  $\square$

Feltéve, hogy (16.47) teljesül, eredményünk alkalmazását először az általános (16.45), (16.48) súlyozott differenciasémára mutatjuk meg. Ekkor

$$B_h = I_h + \tau\sigma A_h \geq \left( \frac{1}{\|A_h\|} + \tau\sigma \right) A_h, \quad (16.54)$$

hiszen  $\mathbb{R}^n$ -ben  $(Ax, x) \leq \|A\|\|x\|^2$  vagyis  $A \leq \|A\|I$ . A (16.52) feltétel tehát akkor igaz, amikor

$$\sigma \geq \frac{1}{2} - \frac{1}{\tau\|A_h\|}. \quad (16.55)$$

Tekintsük most speciálisan a (16.46) esetet, azaz a súlyozott differenciáséma legegyszerűbb (16.28)-(16.30) kiadását! Az  $\|A_h\| = \|A_h^{(0)}\|_{(0,h)}$  normát ismerjük II. 11.4.6-ból :

$$\|A_h^{(0)}\|_{(0,h)} = \lambda_{\max}(A_h^{(0)}) = \left(\frac{2}{h} \cos \frac{\pi h}{2}\right)^2 < \frac{4}{h^2} = \|A_h^{(0)}\|_{C(\omega_h)}.$$

Így  $\frac{h^2}{4} + \tau\sigma \geq \frac{\tau}{2}$  esetén (16.52) teljesül, ami  $\sigma$ -ra nézve a

$$\sigma \geq \frac{1}{2} - \frac{h^2}{4\tau} \quad (16.56)$$

feltételt jelenti. Vegyük észre, hogy ez általában már nem szükséges a stabilitáshoz! (16.56)-ból kapjuk  $\sigma \geq \frac{1}{2}$  esetében, hogy a séma stabil további feltétel nélkül is (ez tehát a Crank–Nicolson sémára és a tisztán implicit sémára vonatkozik). De a (16.28), (16.37) magasabbrendű séma is stabil, hiszen  $\sigma_* > \frac{1}{2} - \frac{h^2}{4\tau}$ .

Viszont  $\sigma = 0$  esetén éppen akkor következik a stabilitás (16.56)-ból, ha (16.42) teljesül.

Ez a stabilitási feltétel erős korlátozást jelent és hátrányossá teszi az explicit séma használatát, mert pl. a Crank–Nicolson sémánál kézenfekvő a  $\tau = O(h)$  időlépés, ami azt jelenti, hogy  $O(N^2)$  a teljes műveletigény a  $Q_T$  megoldási tartományban (ld. a 10. feladatot is).

Most a (16.45) kétréteges sémához térjünk vissza, és foglalkozzunk az inhomogén egyenlet becslésével, kiindulva a (16.51) azonosságból!

**16.6. Tétel** (kezdetiérték szerinti stabilitás következménye a jobboldal szerinti stabilitás). Legyen  $\varphi_h \not\equiv 0$  és  $y_h^0 = 0$ , egyébként teljesüljenek a 16.5. tétel feltételei. Ekkor a kezdetiérték szerinti stabilitásból (a (16.53) becslés értelmében) következik a jobboldal szerinti stabilitás : érvényes ugyanis

$$\|y_h^j\|_A \leq \sum_{k=0}^{j-1} \tau \|B_h^{-1} \varphi_h^k\|_A.$$

**Bizonyítás.** Mivel  $y_h^0 = 0$  és (a  $\delta_{jk}$  Kronecker-szimbolumokkal)

$$B_h \frac{1}{\tau} (y_h^j - y_h^{j-1}) + A_h y_h^{j-1} = \varphi_h^{j-1} = \sum_{k=1}^m \delta_{jk} \varphi_h^{k-1}, \quad j = 1, \dots, m,$$

így a (16.45) egyenletek megoldását a linearitás miatt az

$$y_h^j = \sum_{k=1}^m y_{(k)}^j, \quad j = 1, \dots, m, \quad (16.57)$$

formában írhatjuk fel, beleértve a  $j = 0$  esetet is. Itt az  $y_{(k)}$  vektor (ahol  $k$  rögzített és  $\{1, 2, \dots, m\}$ -ből való) a következő feladat megoldása :

$$y_{(k)}^0 = 0, \quad B_h \frac{1}{\tau} (y_{(k)}^j - y_{(k)}^{j-1}) + A_h y_{(k)}^{j-1} = \delta_{jk} \varphi_h^{k-1}, \quad j = 1, \dots, m.$$

Ezek az  $y_{(k)}$  segédvektorok kis  $j$ -re (míg  $j < k$ ) még nullvektorok, mert addig a kezdeti feltétel és a jobboldal is nulla. Amikor viszont  $j = k$ , akkor

$$B_h y_{(k)}^k = \tau \varphi_h^{k-1}, \quad \text{tehát } y_{(k)}^k = \tau B^{-1} \varphi^{k-1}.$$

Ezután,  $j > k$ -ra,  $y_{(k)}^j$  már a következő egyenlet megoldása :

$$B_h \frac{1}{\tau} (y_{(k)}^j - y_{(k)}^{j-1}) + A_h y_{(k)}^{j-1} = 0.$$

Ezért, (16.52)-nek köszönhetően,  $j \geq k$ -ra

$$\|y_{(k)}^j\|_A \leq \|y_{(k)}^{j-1}\|_A \leq \dots \leq \|y_{(k)}^k\|_A = \tau \|B^{-1} \varphi^{k-1}\|_A.$$

Most (16.57)-ből a háromszög egyenlőtlenség segítségével következik az állítás.  $\square$

Úgy tűnik, a (16.52) feltétel mindent biztosít – a probléma viszont az, hogy most a jobboldalt elég bonyolult normában mérjük. Ezért tovább foglalkozunk a témával és olyan feltételt adunk meg, amely erősebb ugyan, mint (16.52), de kényelmesebb normára vezet.

**16.7. Tétel** (Szamarszkij; kétréteges séma jobboldal szerinti stabilitása). Legyen  $\varphi_h \not\equiv 0$  és  $y_h^0 = 0$ , egyébként teljesüljenek a 16.5. tétel feltételei. Ekkor

$$B_h - \frac{\tau}{2} A_h \geq \varepsilon I_h, \quad \varepsilon = \text{const} > 0, \quad (16.58)$$

biztosítja a (16.45) séma jobboldal szerinti stabilitását a következő becslés értelmében :

$$\|y_h^j\|_A \leq \left( \frac{1}{2\varepsilon} \sum_{k=0}^{j-1} \tau \|\varphi_h^k\|_{(0,h)}^2 \right)^{1/2}. \quad (16.59)$$

**Bizonyítás.** (16.51) jobboldalát becsülhetjük az  $\varepsilon$ -egyenlőtlenséggel (ld. II. 189. o.) :

$$2\tau |(\varphi_h, y_t)_{(0,h)}| \leq \tau \left( \frac{1}{2\varepsilon} \|\varphi_h^j\|_{(0,h)}^2 + 2\varepsilon \|y_t\|_{(0,h)}^2 \right).$$

Innen és (16.51)-ből következik, hogy

$$\|y_h^{j+1}\|_A^2 - \|y_h^j\|_A^2 \leq \frac{\tau}{2\varepsilon} \|\varphi_h^j\|_{(0,h)}^2.$$

Ezt összegezve  $j = 0$ -tól  $j = k - 1$ -ig azt kapjuk, hogy

$$\|y_h^k\|_A^2 \leq \|y_h^0\|_A^2 + \frac{\tau}{2\varepsilon} \sum_{j=0}^{k-1} \|\varphi_h^j\|_{(0,h)}^2,$$

vagyis,  $\|y_h^j\|_A$ -ra felírva, a (16.59) becslés.  $\square$

**Megjegyzések. 1.** Hasonlítsuk össze a tétel (16.59) eredményét a jobb-  
oldal szerinti stabilitás definíciójával!

Ekkor azt látjuk, hogy  $\|y_h^j\|_2 := \|y_h^j\|_A$ , tehát ez megint az energetikai  
norma, de

$$\|\varphi_h^{(j)}\|_3 = \left( \sum_{k=0}^{j-1} \tau \|\varphi_h^k\|_{(0,h)}^2 \right)^{1/2} \quad (16.60)$$

és  $M_2 = 1/\sqrt{2\varepsilon}$ .

**2.** A 16.7. tétel feltételei mellett kapjuk (16.53)-ból és (16.59)-ből, hogy  
a (16.45) feladat  $y_h$  megoldása eleget tesz a következő becslésnek :

$$\|y_h^j\|_A \leq \|y_h^0\|_A + \left( \frac{1}{2\varepsilon} \sum_{k=0}^{j-1} \tau \|\varphi_h^k\|_{(0,h)}^2 \right)^{1/2}. \quad (16.61)$$

Ennek bizonyításához a (16.45) megoldását  $y_h = y_0 + y_\varphi$  alakban bont-  
juk fel. Itt  $y_0$  a (16.45) séma megoldása, amikor  $\varphi_h \equiv 0$ ,  $y_\varphi$  viszont a  
(16.45) megoldása, ha ott  $y_h^0 = 0$ . Ezután a 16.5. ill. 16.7. tétel segítségével  
kapjuk az  $y_0$  ill.  $y_\varphi$  becslését, ahonnan a háromszög egyenlőtlenség segít-  
ségével következik (16.61).

**3.** Amennyiben a séma  $\varphi$  jobboldala (a peremfeltételek eliminációja mi-  
att) a  $g^{(\sigma)}/h^2$  tagot is tartalmazza, a tétel egyfajta peremértékek szerinti  
stabilitási eredményt is tartalmaz, csak hogy ez nem ad igazi benyomást a  
peremfeltételek kihatásáról a megoldásra, v.ö (16.43)-mal. Ez a probléma  
nem merül fel, amikor eredményeinket a konvergencia vizsgálatra alkalmaz-  
zuk, mert a hibaegyenlet peremértékei homogének, ld. a következő pontot.  
 $\square$

Nézzük most meg, hogy (16.58)  $\sigma$ -ra vonatkozólag milyen feltételt jelent,  
amikor (16.45)-ben  $B_h$ -t a (16.48) reláció határozza meg! (16.54)-re hason-  
lóan kapjuk  $0 < \varepsilon \leq 1$  esetén

$$B_h = I_h + \sigma\tau A_h \geq \varepsilon I_h + \left( \frac{1 - \varepsilon}{\|A_h\|} + \sigma\tau \right) A_h \geq \frac{\tau}{2} A_h + \varepsilon I_h,$$

ha  $\sigma \geq \frac{1}{2} - \frac{1-\varepsilon}{\|A_h\|^\tau}$ , vagyis,  $A_h = A_h^{(0)}$  esetén, ha

$$\sigma \geq \frac{1}{2} - \frac{(1-\varepsilon)h^2}{4\tau}. \quad (16.62)$$

Tehát az általános (16.45) súlyozott differenciaséma stabil, ha pl.  $\sigma \geq \frac{1}{2}$ , ami  $\varepsilon = 1$ -nek felel meg. A speciális (16.28), (16.37) séma esetén elegendő  $\varepsilon = \frac{2}{3}$ -at választani ahhoz, hogy stabilitását belássuk.

Továbbá jegyezzük meg, hogy az  $(A_h y_h)_i = -y_{\bar{x},i}$ ,  $y_0 = y_N = 0$  esetben

$$\|y_h\|_A = \left( \sum_{i=1}^N (y_{\bar{x},i})^2 h \right)^{1/2} = |y_h|_{(1,h)} \geq 2 \|y_h\|_{C(\omega_h)},$$

ld. a 11.5. lemmát II-ben, amiből (16.60)-nal és (16.61)-gyel együtt az következik, hogy ilyenkor a maximum normában is stabil a súlyozott séma :

$$\|y_h^j\|_{C(\omega_h)} \leq \frac{1}{2} \left( \|y_h^0\|_A + \frac{1}{\sqrt{2\varepsilon}} \|\varphi_h^{(j)}\|_3 \right).$$

#### 16.4.4 A súlyozott differenciaséma konvergenciája

A stabilitásból és az approximációból következik a konvergencia. Ennek belátásához legyen  $z_h^j$  a hibavektor, ahol

$$z_h^j := \vec{u}_h^j - y_h^j, \quad \text{és ahol az } \vec{u}_h^j := (u(x_1, t_j), \dots, u(x_{N-1}, t_j))^T$$

vektor tartalmazza a (16.6)-(16.9) feladat  $u$  pontos megoldásának értékeit a rácspontokban és  $y_h$  a (16.28)-(16.30) séma megoldása.

**16.8. Tétel** (súlyozott differenciaséma konvergenciája). Amennyiben  $u \in C^{4,3}$  és a (16.62) feltétel teljesül, akkor a (16.28)-(16.32) séma konvergens. Igaz a

$$\|z_h^j\|_{C(\bar{\omega}_h)} = O\left(\tau \left|\sigma - \frac{1}{2}\right| + \tau^2 + h^2\right)$$

becslés minden véges  $[0, T]$  intervallumon.

**Bizonyítás.** (16.28) és (16.35) összevetése azt mutatja meg, hogy a  $z_h$  hibavektorok eleget tesznek a következő egyenleteknek :

$$z_h^0 = 0, \quad j > 0 : (B_h z_t + A_h z_h)_i = \psi_i = O\left(\tau \left|\sigma - \frac{1}{2}\right| + \tau^2 + h^2\right),$$

és a peremfeltételek homogének.

(16.62) miatt a (16.59) becslés alkalmazható, amely szerint tehát

$$\begin{aligned} 2 \|z_h^j\|_{C(\bar{\omega}_h)} &\leq \|z_h^j\|_A \leq \frac{j\tau}{2\varepsilon} O\left(\tau \left|\sigma - \frac{1}{2}\right| + \tau^2 + h^2\right) \\ &\leq \frac{T}{2\varepsilon} O\left(\tau \left|\sigma - \frac{1}{2}\right| + \tau^2 + h^2\right), \quad \text{ha } 0 \leq j\tau = t_j \leq T. \quad \square \end{aligned}$$



**Megjegyzés.** Hasonlóan, ha  $u \in C^{6,3}$ , akkor a (16.28)-(16.32), (16.37) séma konvergens, és igaz, hogy  $\|z_h^j\|_{C(\bar{\omega}_h)} = O(\tau^2 + h^4)$ .  $\square$

Ehhez a ponthoz ld. a 6., valamint a 39. feladatot is.

### 16.4.5 Más peremfeltételek

Tekintsük most a hővezetési egyenletet másod- vagy harmadfajú peremfeltételek mellett! Az ilyen feltételek fizikai hátterére nézve ld. II. 11.2. A vegyes feladat képletei :

$$t > 0, \quad 0 < x < 1 : \quad \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f, \quad (16.63)$$

$$t = 0, \quad 0 \leq x \leq 1 : \quad u(x, t) = u_0(x), \quad (16.64)$$

$$t > 0, \quad x = 0 : \quad -\frac{\partial u}{\partial x} + \sigma_0 u = g_0(t), \quad (16.65)$$

$$t > 0, \quad x = 1 : \quad \frac{\partial u}{\partial x} + \sigma_1 u = g_1(t). \quad (16.66)$$

Itt  $\sigma_0, \sigma_1 \geq 0$  adott konstansok és  $g_0, g_1$  adott függvények.

A térbeli operátor approximációjával részletesen már II. 11.4.5-ben és 11.4.6-ban foglalkoztunk. Onnan tudjuk, hogy másodrendű approximációt akkor kapunk, ha a korábbi  $A_h$  mátrix helyére a következőt veszünk :

$$(A_h y_h)_i = (A_h(\sigma_0, \sigma_1) y_h)_i = \begin{cases} -\frac{2}{h}(y_{x,0} - \sigma_0 y_0), & \text{ha } i = 0, \\ -y_{\bar{x},i}, & \text{ha } 1 \leq i \leq N-1, \\ \frac{2}{h}(y_{\bar{x},N} + \sigma_1 y_N), & \text{ha } i = N. \end{cases} \quad (16.67)$$

Ezután a súlyozott differenciaséma képletei a szokásos alakban felírhatóak. A jobboldali függvényeket most a  $t_{j+\sigma} = (j+\sigma)\tau$  pillanatban értékeljük ki (erre a lehetőségre már 16.4.2-ben mutattunk rá), és feltesszük, hogy  $0 \leq \sigma \leq 1$  :

$$y_{t,i} + (A_h(\sigma_0, \sigma_1) y_h^{(\sigma)})_i = \varphi_i^j, \quad j = 0, 1, \dots, m-1, \quad (16.68)$$

$$\varphi_i^j := f_i^{j+\sigma}, \quad 1 \leq i \leq N-1, \quad (16.69)$$

$$\varphi_0^j := f_0^{j+\sigma} + \frac{2}{h} g_0^{j+\sigma}, \quad \varphi_N^j := f_N^{j+\sigma} + \frac{2}{h} g_1^{j+\sigma}, \quad (16.70)$$

$$y_i^0 = u_0(x_i), \quad 0 \leq i \leq N. \quad (16.71)$$

Példaként vezessünk le az  $i = 0$ -nak megfelelő differencia-egyenletet és képlethibáját!

Legyen  $u$  a (16.63)-(16.66) vegyes feladat elég sima megoldása :  $u \in C^{4,3}(\bar{Q}_T)$ . Ekkor, határérték értelmében, a (16.63) egyenlet a peremen is

teljesül. Ezért, a (16.65) peremfeltételt is figyelembe véve, azt kapjuk, hogy minden  $0 \leq t \leq T$ -re érvényes

$$\begin{aligned} u_{x,0} &= \left( \frac{\partial u}{\partial x} + \frac{h}{2} \frac{\partial^2 u}{\partial x^2} \right) \Big|_{x=0} + \frac{h^2}{6} \frac{\partial^3 u}{\partial x^3}(\vartheta h, t) \quad (0 < \vartheta < 1) \\ &= -g_0(t) + \left[ \sigma_0 u + \frac{h}{2} \left( \frac{\partial u}{\partial t} - f \right) \right] (0, t) + \frac{h^2}{6} \frac{\partial^3 u}{\partial x^3}(\vartheta h, t). \end{aligned}$$

Ebbe a relációba  $t = t_{j+\sigma}$ -t helyettesítjük be, és (16.34) helyett felhasználva, hogy  $v \in C^2[0, T]$  esetén

$$v^{(\sigma)} = v^{j+\sigma} + \frac{\tau^2}{2} \sigma(1-\sigma) \left( \frac{\partial v^2}{\partial t^2} \right)^{j+\eta}, \quad 0 < \eta < 1,$$

azt kapjuk  $v(t) = (u_{x,0} - \sigma_0 u_0)(t)$ -ra – a korábbi  $u \in C^{4,3}$  feltételen túl követelve  $\frac{\partial^3 u}{\partial t^2 \partial x}$  folytonosságát, hogy igaz

$$\begin{aligned} (u_x - \sigma_0 u)_0^{j+\sigma} &= - \left( \frac{h}{2} f_0 + g_0 \right)^{j+\sigma} + \left( \frac{h}{2} \frac{\partial u}{\partial t} \right)_0^{j+\sigma} + \frac{h^2}{6} \frac{\partial^3 u}{\partial x^3}(\vartheta h, t_{j+\sigma}) \\ &= (u_x - \sigma_0 u)_0^{(\sigma)} - \frac{\tau^2}{2} \sigma(1-\sigma) \left( \frac{\partial^2 (u_x - \sigma_0 u)}{\partial t^2} \right)_0^{j+\eta}. \end{aligned}$$

Itt a  $\left( \frac{h}{2} \frac{\partial u}{\partial t} \right)_0^{j+\sigma}$  tag deriváltját még a differencia-képlettel helyettesítjük :

$$\left( \frac{\partial u}{\partial t} \right)_0^{j+\sigma} = u_{t,0} + \tau \left( \sigma - \frac{1}{2} \right) \left( \frac{\partial^2 u}{\partial t^2} \right)_0^{j+\sigma} + O(\tau^2),$$

ahol a  $O(\tau^2)$  tag  $\frac{\partial^3 u}{\partial t^3}$  értékeivel kapcsolatos. Ekkor összefoglalva és átrendezve :

$$\begin{aligned} -(u_x - \sigma_0 u)_0^{(\sigma)} &+ \frac{h}{2} u_{t,0} - \left( \frac{h}{2} f_0 + g_0 \right)^{j+\sigma} \\ &= \frac{h\tau}{2} \left( \frac{1}{2} - \sigma \right) \left( \frac{\partial^2 u}{\partial t^2} \right)_0^{j+\sigma} + O(h\tau^2) \\ &\quad - \frac{h^2}{6} \frac{\partial^3 u}{\partial x^3}(\vartheta h, t_{j+\sigma}) - \frac{\tau^2}{2} \sigma(1-\sigma) \left( \frac{\partial^2 (u_x - \sigma_0 u)}{\partial t^2} \right)_0^{j+\eta} \\ &= O(h\tau |\sigma - \frac{1}{2}|) + O(\tau^2 + h^2). \end{aligned}$$

Ez azt jelenti, hogy a (16.67), (16.68), (16.70) egyenleteknek megfelelő képlethiba

$$\begin{aligned} \psi_0^j &:= u_{t,0} - \frac{2}{h} (u_{x,0} - \sigma_0 u_0)^{(\sigma)} - \left( f_0 + \frac{2}{h} g_0 \right)^{j+\sigma} \\ &= O(\tau |\sigma - \frac{1}{2}|) + \frac{1}{h} O(\tau^2 + h^2). \end{aligned} \tag{16.72}$$

$i = N$ -re analóg módon következik

$$\begin{aligned}\psi_N^j &:= u_{t,N} + \frac{2}{h}(u_{\bar{x},N} + \sigma_1 u_N)^{(\sigma)} - \left(f_N + \frac{2}{h}g_1\right)^{j+\sigma} \\ &= O(\tau|\sigma - \frac{1}{2}|) + \frac{1}{h}O(\tau^2 + h^2).\end{aligned}\quad (16.73)$$

Ezután a  $z_h := \vec{u}_h - y_h$  hibavektor (azaz az  $x \in \bar{\omega}_h$  racson definiált  $z(x)$  függvény) eleget tesz a következő egyenleteknek :

$$z_t + A_h(\sigma_0, \sigma_1)z_h^{(\sigma)} = \psi_h^j, \quad j = 0, 1, \dots, m-1, \quad (16.74)$$

$$z_h^0 = 0, \quad (16.75)$$

ahol  $i = 0$ -ra (16.72),  $i = N$ -re (16.73) és  $\omega_h$ -n (16.38) adja a képlethiba komponenseit :

$$\psi_i^j = O(\tau|\sigma - \frac{1}{2}| + \tau^2 + h^2), \quad 1 \leq i \leq N-1. \quad (16.76)$$

Olyan differenciaséma szerkesztése, amelynek képlethibája ennél magasabbrendű, a 12. feladat tárgya.

A (16.72)-(16.76) feladat elemzését a (16.67)-ben definiált  $A_h(\sigma_0, \sigma_1) \in \mathbb{R}^{(N+1) \times (N+1)}$  mátrixnál kezdjük, feltéve, hogy paraméterei nemnegatívak, mint (16.63)-(16.66)-ban.

**16.9. Tétel** (másodfajú és harmadfajú peremfeltételekhez tartozó mátrix tulajdonságai). A (16.67) mátrix  $\sigma_0, \sigma_1$  paraméterei legyenek nemnegatívak. Ekkor :

1. Az  $A_h(0, 0)$  mátrix szinguláris, rangja  $N$ , és  $A_h(0, 0)v_h^{(0)} = 0$ , ahol  $v_h^{(0)} := (1, \dots, 1)^T \in \bar{\omega}_h$ ;

2.  $A_h(\sigma_0, \sigma_1)$  pozitív szemidefinit a trapézsabállyal definiált  $(y_h, v_h)_{(0,h)}$  skalárszorzatban (ld. II 11.4.6-ban (101)-et is), és érvényes

$$\begin{aligned}0 &\leq (A_h(\sigma_0, \sigma_1)v_h, v_h)_{(0,h)} \leq M_1 \|v_h\|_{(0,h)}^2 \text{ minden } v_h\text{-ra,} \quad (16.77) \\ M_1 &:= \left(\frac{2}{h}(\sigma_0 + \sigma_1) + \frac{4}{h^2}\right);\end{aligned}$$

3. Amennyiben  $\sigma_0 + \sigma_1 > 0$ , akkor  $A_h(\sigma_0, \sigma_1)$  pozitív definit :

$$M_0 \|v_h\|_{(0,h)}^2 \leq (A_h(\sigma_0, \sigma_1)v_h, v_h)_{(0,h)} \text{ minden } v_h\text{-ra,} \quad M_0 := \frac{\sigma_0 + \sigma_1}{1 + \sigma_0 + \sigma_1}. \quad (16.78)$$

**Bizonyítás.** 1. Világos, hogy  $A_h(0, 0)v_h^{(0)} = 0$  és így a mátrix szinguláris. Rangja  $N$ , mert az  $N \times N$ -es felső főminor-mátrix reguláris, hiszen

annak felel meg, hogy a jobboldali peremfeltétel homogén elsőfajú. Arra is hivatkozhatunk, hogy a II 11.4.11-beli (177) becslés szerint a homogén  $A_h(0,0)y_h = 0$  feladatnak csak egy megoldása lehet, ha az  $(y_h, v_h^{(0)})_{(0,h)} = 0$  feltételt is előírjuk.

2. Tetszőleges  $v_h \in \bar{\omega}_h$  vektorra igaz (ld. a (102) relációt II. 11.4.6-ban, ahonnan  $A_h$  szimmetriája is következik), hogy

$$\begin{aligned} (A_h(\sigma_0, \sigma_1)v_h, v_h)_{(0,h)} &= \sigma_0 v_0^2 + \sigma_1 v_N^2 + \sum_{i=1}^{N-1} (v_{\bar{x},i})^2 h & (16.79) \\ &\leq \left(\sigma_0 + \frac{2}{h}\right) v_0^2 + \left(\sigma_1 + \frac{2}{h}\right) v_N^2 + \frac{4}{h} \sum_{i=1}^{N-1} v_i^2 \\ &\leq \left(\frac{2}{h}\sigma_0 + \frac{4}{h^2}\right) \frac{h}{2} v_0^2 + \left(\frac{2}{h}\sigma_1 + \frac{4}{h^2}\right) \frac{h}{2} v_N^2 + \frac{4}{h^2} \sum_{i=1}^{N-1} v_i^2 h \\ &\leq \left(\frac{2}{h}(\sigma_0 + \sigma_1) + \frac{4}{h^2}\right) \|v_h\|_{(0,h)}^2, \end{aligned}$$

ezenkívül (16.79) mutatja, hogy  $0 \leq (A_h(\sigma_0, \sigma_1)v_h, v_h)_{(0,h)}$ . Így (16.77) teljesül.

3. Ld. a 13. feladatot.  $\square$

Most fordulunk a (16.72)-(16.76) feladat megoldásának, a  $z_h$  hiba becsléséhez. Ennek során használjuk a (16.44)-beli  $\|\cdot\|_{(0,h)}$  avagy  $L_2(\bar{\omega}_h)$ -norma mellett a (16.79) képlettel definiált

$$(A_h(\sigma_0, \sigma_1)v_h, v_h)_{(0,h)}^{1/2} =: \|v_h\|_{(1,h)}$$

félnormát is, amelyet  $H^1(\bar{\omega}_h)$ -félnormának is nevezzük.

**16.7'. Tétel** (Kétréteges séma jobboldal szerinti  $L_2(\bar{\omega}_h)$ -stabilitása, ha  $\sigma \geq \frac{1}{2}$ ). Legyen (16.68)-(16.71)-ben  $\varphi_h \neq 0$ ,  $y_h^0 = 0$ ,  $\sigma_0 + \sigma_1 > 0$  és  $\sigma \geq \frac{1}{2}$ . Ekkor a séma stabil a jobboldal szerint a következő becslés értelmében :

$$\|y_h^j\|_{(0,h)} \leq \left( \frac{1}{2M_0} \sum_{k=0}^{j-1} \tau(|\varphi_h^k|, 1)_{(0,h)}^2 \right)^{1/2}, \quad (16.80)$$

ahol  $M_0$  a (16.78)-beli konstans.

**Bizonyítás.** Egy stabilitási becslés bizonyításához a (16.45) egyenletből is tudnánk kiindulni, mert (16.68) ilyen alakú. De célünk a 16.7. tételhez képest a jobboldalt a (16.80)-beli gyengébb normában mérni. Ehhez (16.68)-at nem  $2\tau y_t$ -vel szorozzuk meg skalárisan, hanem  $2\tau y_h^{(\sigma)}$ -val :

$$2\tau(y_t, y_h^{(\sigma)})_{(0,h)} + 2\tau\|y_h^{(\sigma)}\|_{(1,h)}^2 = 2\tau(\varphi_h, y_h^{(\sigma)})_{(0,h)}. \quad (16.81)$$

Itt az első tagot alakítjuk át, felhasználva, hogy  $y_h^{(\sigma)} = \frac{1}{2}(\hat{y}_h + y_h) + \tau(\sigma - \frac{1}{2})y_t$  :

$$\begin{aligned} 2\tau(y_t, y_h^{(\sigma)})_{(0,h)} &= 2(\hat{y}_h - y_h, \frac{1}{2}(\hat{y}_h + y_h) + \tau(\sigma - \frac{1}{2})y_t)_{(0,h)} \\ &= \|\hat{y}_h\|_{(0,h)}^2 - \|y_h\|_{(0,h)}^2 + 2\tau^2(\sigma - \frac{1}{2})\|y_t\|_{(0,h)}^2. \end{aligned}$$

Ezért (16.81)-ből következik ((16.51) helyett) az alábbi azonosság :

$$\|\hat{y}_h\|_{(0,h)}^2 - \|y_h\|_{(0,h)}^2 + 2\tau^2(\sigma - \frac{1}{2})\|y_t\|_{(0,h)}^2 + 2\tau\|y_h^{(\sigma)}\|_{(1,h)}^2 = 2\tau(\varphi_h, y_h^{(\sigma)})_{(0,h)}. \quad (16.82)$$

Ennek jobboldalát becsüljük (feltételeink mellett teljesül (16.77), és  $\|\cdot\|_{(1,h)}$  norma) az

$$M_0\|v_h\|_{C(\bar{\omega}_h)}^2 \leq \|v_h\|_{(1,h)}^2 \text{ minden } v_h\text{-ra,} \quad M_0 := \frac{\sigma_0 + \sigma_1}{1 + \sigma_0 + \sigma_1}, \quad (16.83)$$

diszkrét beágyazási tétellel (ld. a 13. feladatot) a következőképpen :

$$\begin{aligned} 2\tau(\varphi_h, y_h^{(\sigma)})_{(0,h)} &\leq 2\tau(|\varphi_h|, |y_h^{(\sigma)}|)_{(0,h)} \leq 2\tau(|\varphi_h|, 1)_{(0,h)} \frac{1}{\sqrt{M_0}} \|y_h^{(\sigma)}\|_{(1,h)} \\ &\leq 2\tau\|y_h^{(\sigma)}\|_{(1,h)}^2 + \frac{\tau}{2M_0} (|\varphi_h|, 1)_{(0,h)}^2, \end{aligned}$$

ahol a végén az  $\varepsilon$ -egyenlőtlenséget alkalmaztuk (ld. II. 189. o., (103)). Innen és (16.82)-ből (ahonnan elhagyjuk a nemnegatív  $2\tau^2(\sigma - \frac{1}{2})\|y_t\|_{(0,h)}^2$  tagot) következik, hogy

$$\|y_h^{j+1}\|_{(0,h)}^2 - \|y_h^j\|_{(0,h)}^2 + 2\tau\|y_h^{(\sigma)}\|_{(1,h)}^2 \leq 2\tau\|y_h^{(\sigma)}\|_{(1,h)}^2 + \frac{\tau}{2M_0} (|\varphi_h^j|, 1)_{(0,h)}^2.$$

Ezt összegezve  $j = 0$ -tól  $j = k - 1$ -ig azt kapjuk, hogy

$$\|y_h^k\|_{(0,h)}^2 \leq \|y_h^0\|_{(0,h)}^2 + \frac{\tau}{2M_0} \sum_{j=0}^{k-1} (|\varphi_h^j|, 1)_{(0,h)}^2,$$

ami  $y_h^0 = 0$  miatt és  $\|y_h^j\|_{(0,h)}$ -ra felírva adja a (16.80) becslést.  $\square$

**16.9'. Következmény.** (konvergencia  $L_2(\bar{\omega}_h)$ -normában másod- és harmadfajú peremfeltételek esetén). Legyen  $\sigma_0, \sigma_1 \geq 0$  és  $\sigma_0 + \sigma_1 > 0$ , a (16.63)-(16.66) feladat megoldása  $u \in C^{4,3}$ , és még  $\frac{\partial^3 u}{\partial t^2 \partial x}$  is legyen folytonos. Ekkor, ha  $\sigma \geq \frac{1}{2}$ , a (16.68)-(16.71) séma hibájára igaz a következő becslés :

$$\|(\vec{u}_h - y_h)^j\|_{(0,h)} \leq O(\tau|\sigma - \frac{1}{2}| + \tau^2 + h^2). \quad (16.84)$$

[Kommentár. Feltételeink megengedik, hogy az egyik peremfeltétel másodfajú legyen, pl.  $\sigma_0 = 0, \sigma_1 > 0$ .]

**Bizonyítás.** Elegendő a 16.7'. tételt a (16.72)-(16.76) feladatra alkalmazni. Ehhez a (16.80) egyenlőtlenség  $(|\varphi_h^k|, 1)_{(0,h)}$  tagját kell becsülni  $\varphi_h^k = \psi_h^k$  esetén. A  $\psi_i^j$  mindenütt érvényes  $O(\tau|\sigma - \frac{1}{2}| + \tau^2 + h^2)$  része persze összegzéskor is ennyit ad, viszont a csak a peremen szereplő (16.72), (16.73) kifejezések (akár a durvább  $\frac{1}{h}O(\tau|\sigma - \frac{1}{2}| + \tau^2 + h^2)$  formában) csak  $i = 0$ -ra és  $i = N$ -re kerülnek be a

$$(|\psi_h^k|, 1)_{(0,h)} = \frac{h}{2}|\psi_0^k| + \frac{h}{2}|\psi_N^k| + \sum_{i=1}^{N-1} |\psi_i^k| h$$

trapézösszegbe, és akkor még  $h$ -val is szorozódnak. Így (16.84) következik.  $\square$

**Megjegyzés.** Ha az eredeti 16.7. tételt alkalmazzuk (16.72)-(16.76)-ra, akkor  $\varphi_h^k = \psi_h^k = \psi_h^k(1) + \psi_h^k(2)$ -val kiértékelve a  $\|\varphi_h^k\|_{(0,h)}$  normát (ahol  $\psi_h^k(1) = O(\tau|\sigma - \frac{1}{2}| + \tau^2 + h^2)$  és  $\psi_h^k(2) = \frac{1}{h}O(\tau^2 + h^2)(1, 0, \dots, 0, 1)^T$ ) egyrészt azt kapjuk, hogy  $\|\varphi_h^k\|_{(0,h)} \leq \|\psi_h^k(1)\|_{(0,h)} + \|\psi_h^k(2)\|_{(0,h)}$ , másrészt itt  $\|\psi_h^k(1)\|_{(0,h)} = O(\tau|\sigma - \frac{1}{2}| + \tau^2 + h^2)$ , és végül

$$\|\psi_h^k(2)\|_{(0,h)}^2 = \frac{h}{2}(\psi_0^k(2))^2 + \frac{h}{2}(\psi_N^k(2))^2 = \frac{1}{h}O((\tau^2 + h^2)^2).$$

Tehát összesítve igaz a következő becslés :

$$\|(\vec{u}_h - y_h)^j\|_{(1,h)} \leq O(\tau|\sigma - \frac{1}{2}|) + h^{-1/2}O(\tau^2 + h^2), \quad (16.85)$$

ha  $1 \geq \varepsilon = \text{const} > 0$ -val a súly a

$$\sigma \geq \frac{1}{2} - \frac{1 - \varepsilon}{4(1 + \frac{h}{2}(\sigma_0 + \sigma_1))} \frac{h^2}{\tau} \quad (16.86)$$

egyenlőtlenségnek tesz eleget, ld. a 16.7. tételt, (16.62)-t és a (16.77) becslést. (16.85) előnye, hogy (16.83) alapján még a maximumnormában is ugyanolyan rendű konvergencia következik (ehhez ld. a 14. feladatot is).  $\square$

Amikor  $x = 0$ -nál és  $x = 1$ -nél másodfajú peremfeltétel adott, a fent definiált  $A_h$  mátrix szinguláris (16.9. tétel), az  $\|\cdot\|_{(1,h)}$  csak félnorma, és a 16.7. tétel csak a  $v_h^{(0)} = (1, \dots, 1)^T$  sajátvektorra ortogonális altérben alkalmazható. De az egész  $L_2(\bar{\omega}_h)$ -térben és minden véges  $t$ -intervallumra is levezethetünk egy pontossági becslést.

**16.10. Tétel** (konvergencia  $L_2(\bar{\omega}_h)$ -normában másodfajú peremfeltételek esetén). A (16.63)-(16.66),  $\sigma_0 = \sigma_1 = 0$ , feladat megoldására teljesüljenek a 16.9'. következmény feltételei. Ekkor a (16.68)-(16.71) séma az

$L_2(\bar{\omega}_h)$ -normában konvergál, ha a súly  $\sigma \geq \frac{1}{2}$  és ha  $t_j \leq T$ . Igaz a következő becslés :

$$\|(\bar{u}_h - y_h)^j\|_{(0,h)} \leq O(\tau|\sigma - \frac{1}{2}|) + h^{-1/2}O(\tau^2 + h^2).$$

**Bizonyítás.** Induljunk ki a 16.7'. tétel (16.82) azonosságából, amelynek (16.72)-(16.76)-ra való alkalmazásához  $y_h$  helyettesítendő  $z_h$ -val és  $\varphi_h$  a  $\psi_h$  képlethibával :

$$\|\hat{z}_h\|_{(0,h)}^2 - \|z_h\|_{(0,h)}^2 + 2\tau^2(\sigma - \frac{1}{2})\|z_t\|_{(0,h)}^2 + 2\tau\|z_h^{(\sigma)}\|_{(1,h)}^2 = 2\tau(\psi_h, z_h^{(\sigma)})_{(0,h)}. \quad (16.87)$$

De mivel a (16.83) diszkrét beágyazási tételt most nem tudjuk használni, ennek jobboldalát másképpen becsüljük, figyelembe véve, hogy  $z_h^{(\sigma)} = \frac{1}{2}(\hat{z}_h + z_h) + \tau(\sigma - \frac{1}{2})z_t$ , a Cauchy- és az  $\varepsilon$ -egyenlőtlenség segítségével :

$$\begin{aligned} 2\tau(\psi_h, z_h^{(\sigma)})_{(0,h)} &= \tau(\psi_h, \hat{z}_h + z_h)_{(0,h)} + 2\tau^2(\sigma - \frac{1}{2})(\psi_h, z_t)_{(0,h)} \\ &\leq \tau\|\psi_h^j\|_{(0,h)}\|\hat{z}_h + z_h\|_{(0,h)} + 2\tau^2(\sigma - \frac{1}{2})\|\psi_h\|_{(0,h)}\|z_t\|_{(0,h)} \\ &\leq \frac{\tau\varepsilon}{2}\|\hat{z}_h + z_h\|_{(0,h)}^2 + \frac{\tau}{2\varepsilon}\|\psi_h^j\|_{(0,h)}^2 \\ &\quad + 2\tau^2(\sigma - \frac{1}{2})(\|z_t\|_{(0,h)}^2 + \frac{1}{4}\|\psi_h\|_{(0,h)}^2), \end{aligned}$$

és itt  $\frac{\tau\varepsilon}{2}\|\hat{z}_h + z_h\|_{(0,h)}^2 \leq \tau\varepsilon(\|\hat{z}_h\|_{(0,h)}^2 + \|z_h\|_{(0,h)}^2)$ . Összegezve (16.87)-tel kapjuk

$$(1 - \varepsilon\tau)\|z_h^{j+1}\|_{(0,h)}^2 \leq (1 + \varepsilon\tau)\|z_h^j\|_{(0,h)}^2 + \frac{\tau}{2\varepsilon}(1 + \tau\varepsilon(\sigma - \frac{1}{2}))\|\psi_h^j\|_{(0,h)}^2,$$

ahonnan  $\varepsilon < 1/T \leq 1/\tau$  esetén következik, hogy

$$\begin{aligned} \|z_h^{j+1}\|_{(0,h)}^2 &\leq \frac{1 + \varepsilon\tau}{1 - \varepsilon\tau} \left\{ \|z_h^j\|_{(0,h)}^2 + \frac{\tau(\sigma + \frac{1}{2})}{2\varepsilon}\|\psi_h^j\|_{(0,h)}^2 \right\} \\ &\leq \left( \frac{1 + \varepsilon\tau}{1 - \varepsilon\tau} \right)^{j+1} \left\{ \|z_h^0\|_{(0,h)}^2 + \frac{\sigma + \frac{1}{2}}{2\varepsilon} \sum_{k=0}^j \tau\|\psi_h^k\|_{(0,h)}^2 \right\}. \end{aligned}$$

Itt  $z_h^0 = 0$ , ezenkívül azt használjuk fel, hogy  $\varepsilon := 1/(3T) \leq 1/(3\tau)$  választásával

$$\frac{1 + \varepsilon\tau}{1 - \varepsilon\tau} \leq 1 + 3\varepsilon\tau \leq e^{\tau/T},$$

és ezért

$$\|z_h^{j+1}\|_{(0,h)}^2 \leq \frac{3T}{2}(\sigma + \frac{1}{2})e \sum_{k=0}^j \tau\|\psi_h^k\|_{(0,h)}^2.$$

Mivel a 16.9'. következményhez tartozó megjegyzésben közölt  $\|\psi_h^k\|_{(0,h)}$  norma becslése most is érvényes, a tétel következik.  $\square$

A véges időintervallumra való korlátozás a 16.10. tételben nem véletlen, hiszen már abban a speciális esetben, amikor (16.63)-(16.66)-ban  $g_0, g_1$  és  $f$  nem függenek  $t$ -től, a (16.68)-(16.71) sémának  $v_h^{(0)}$ -val kapcsolatos komponens a megoldásában  $\tau(\varphi_h, v_h^{(0)})_{(0,h)} = \tau c$ -vel növekszik minden lépésben, ahol (a  $\varphi_h$  vektor trapézösszegét  $I_N(\varphi_h)$ -val jelölve)

$$\begin{aligned} (\varphi_h, v_h^{(0)})_{(0,h)} &= \frac{h}{2}\varphi_0 + \frac{h}{2}\varphi_N + \sum_{i=1}^{N-1} \varphi_i h =: I_N(\varphi_h) \\ &= \frac{h}{2}(f_0 + \frac{2}{h}g_0) + \frac{h}{2}(f_N + \frac{2}{h}g_1) + \sum_{i=1}^{N-1} f_i h \\ &= g_0 + g_1 + I_N(\vec{f}_h), \end{aligned}$$

és ha a  $c$  konstáns nem nulla, akkor az a komponens úgy nő  $t$ -vel, mint  $\tau c$ . Ilyenkor  $t \rightarrow \infty$  esetén nincs stacionárius megoldás.

Ha  $c = 0$ , akkor viszont van, ami szoros kapcsolatban áll a (16.68)-(16.71) séma *megmaradási tételével*. Ennek levezetéséhez a séma egyenleteit megszorozzuk  $\frac{h}{2}$ -vel, ha  $i = 0$  vagy  $i = N$ , ill.  $h$ -val egyébként, majd az eredményeket összeadjuk. Ekkor azt kapjuk, hogy

$$\begin{aligned} I_N(y_t) &= \frac{1}{\tau} (I_N(y_h^{j+1}) - I_N(y_h^j)) \\ &= -\sigma_0 y_0^{(\sigma)} - \sigma_1 y_N^{(\sigma)} + g_0^{j+\sigma} + g_1^{j+\sigma} + I_N(\vec{f}_h^{j+\sigma}), \end{aligned} \quad (16.88)$$

mert az  $x$ -irányú differenciahányadosok az összegzésnél mind kiesnek. A (16.88) *diszkrét megmaradási tétel* approximálja folytonos analogonját, amely a (16.63)-(16.66) egyenletek integrálásával előállítható :

$$I\left(\frac{\partial u}{\partial t}\right) = \frac{\partial}{\partial t} I(u) = -\sigma_0 u(0, t) - \sigma_1 u(1, t) + g_0(t) + g_1(t) + I(f(\cdot, t)).$$

Itt  $I(w) := \int_0^1 w(x) dx$ . Az approximáció másodrendű  $t = t_{j+\sigma}$  esetén. Ha másodfajúak a peremfeltételek és igaz  $g_0^{j+\sigma} + g_1^{j+\sigma} + I_N(\vec{f}_h^{j+\sigma}) = 0$  minden  $j$ -re, ill.  $g_0(t) + g_1(t) + I(f(\cdot, t)) = 0$  minden  $t$ -re, akkor teljesül a diszkrét, ill. „folytonos” megmaradási tétel szűkebb értelemben :

$$I_N(y_h^{j+1}) = I_N(y_h^j) = I_N(y_h^0), \text{ ill. } I(u(\cdot, t)) = I(u(\cdot, 0)).$$

Ezzel kapcsolatban ld. a II. 10.7. pontot is. Amennyiben a (16.68)-(16.69) egyenleteket csak  $1 \leq i \leq N-1$ -re használjuk, és  $i = 0$ -nál ill.  $i = N$ -nél csupán elsőrendű approximációkat veszünk :

$$-(y_{x,0} - \sigma_0 y_0)^{(\sigma)} = g_0^{j+\sigma}, \quad (y_{x,N} + \sigma_1 y_N)^{(\sigma)} = g_1^{j+\sigma},$$



akkor a diszkrét megmaradási tétel (az  $\tilde{I}_N(v_h) := \sum_{i=1}^{N-1} v_i h$  jelöléssel)

$$\tilde{I}_N(y_t) = -\sigma_0 y_0^{(\sigma)} - \sigma_1 y_N^{(\sigma)} + g_0^{j+\sigma} + g_1^{j+\sigma} + \tilde{I}_N(\tilde{f}_h^{j+\sigma}),$$

és ez csak elsőrendben approximálja a folytonos tételt.

### 16.4.6 A végeelem séma

Tekintsük most a végeelem módszer alkalmazását a hővezetési egyenletre, kezdeti és homogén elsőfajú peremfeltételek mellett :

$$\begin{aligned} t > 0, \quad 0 < x < 1 : \quad \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + f(x, t); & (16.89) \\ t = 0, \quad 0 \leq x \leq 1 : \quad u(x, t) &= u_0(x); \\ t \geq 0, \quad x = 0, x = 1 : \quad u(x, t) &= 0. \end{aligned}$$

A gyenge megoldás definíciójához a következő variációs feladatból indulunk ki (v.ö. (16.15)-tel 16.2-ben) :

Keressünk olyan  $u \in H^1\{(0, T) \rightarrow H_0^1(0, 1)\}$  függvényt, hogy

$$\left( \frac{\partial u_h}{\partial t}, v \right)_0 + (u'_h, v')_0 = (f, v)_0, \quad (16.90)$$

minden  $v \in H_0^1(0, 1)$ -re. Itt  $(u, v)_0$  az  $L_2(0, 1)$ -norma.

A végeelem módszernek megfelelően először a  $H_0^1(0, 1)$  lineáris alterét definiáljuk a  $\{w_i\}_{i=1}^{N-1}$  rendszer segítségével :

$$V_h := \text{span}\{w_i\}_{i=1}^{N-1} \subset H_0^1(0, 1),$$

ahol – mint legegyszerűbb példa – a  $w_i$ -k a kalapfüggvények (amelyeket már II. 11.6.2-ben használtunk) :

$$w_i(x) := \begin{cases} 0, & \text{ha } x \notin (x_{i-1}, x_{i+1}), \\ 1 - \left| \frac{x-x_i}{h} \right|, & \text{egyébként.} \end{cases}$$

Itt  $x_i := ih$ ,  $i = 0, \dots, N$ , a csomópontok,  $N - 1$  a belső pontok száma és  $h := 1/N$  az elemek hossza.

Ezután a (16.90) feladat megoldásának közelítését

$$u(x, t) \approx u_h(x, t) := \sum_{j=1}^{N-1} y_j(t) w_j(x) \quad (16.91)$$

alakjában keressük, ahol  $u_h \in H^1\{(0, T) \rightarrow V_h\}$ .

$u_h$  meghatározásához a következő variációs egyenletet tekintjük (16.90) helyett (ehhez ld. részletesebben a 16.8. pontot) :

$$\left( \frac{\partial u_h}{\partial t}, v_h \right)_0 + (u'_h, v'_h)_0 = (f, v_h)_0,$$

ahol  $v_h \in V_h$  tetszőleges.  $v_h$  gyanánt a  $w_i$  függvényeket elegendő ide behelyettesíteni :

$$\sum_{j=1}^{N-1} [\dot{y}_j(w_j, w_i)_0 + y_j(w'_j, w'_i)_0] = (f, w_i)_0, \quad i = 1, \dots, N-1. \quad (16.92)$$

Itt  $\dot{y}_j$  jelöli a keresett (16.91) megoldás  $j$ -edik együtthatójának deriváltját  $t$  szerint. Így az  $y_j$ -kből összeállított  $y$  vektor számára (16.92) egy közönséges differenciálegyenletekből álló rendszer. Annak jobboldala a (16.89) egyenlet jobboldalának Galjorkin-projekciója (v.ö. 15.8.1-gyel).

Mivel eddig csak a térbeli koordinátát diszkretizáltuk, itt is – mint (16.27) esetén – beszélhetünk *egyenesek módszeréről* ill. a (16.89) feladat *szemidiszkréteizációjáról*.

Írjuk fel a (16.92) rendszert mátrixalakban! Legyen

$$A := (a_{ij}) \in \mathbb{R}^{(N-1) \times (N-1)}, \quad \text{ahol } a_{ij} := (w'_j, w'_i),$$

és  $G := (g_{ij}) \in \mathbb{R}^{(N-1) \times (N-1)}$  legyen a  $\{w_i\}$  rendszer Gram-féle mátrixa (a végeelem módszer mechanikai feladatokra való alkalmazásánál szokásos szóhasználatnál élve ez a *tömegmátrix*). Annak elemeit (valamint az  $A$  merevségi mátrixét is) ebben a speciális esetben II-ben számítottuk ki, ld. ott (230), 246. o. Onnan tudjuk, hogy

$$A = hA_h^{(0)}, \quad A_h^{(0)} = \frac{1}{h^2} \text{tridiag}(-1, 2, -1), \quad G = \frac{h}{6} \text{tridiag}(1, 4, 1). \quad (16.93)$$

A tényleges számítás előkészítéséhez még az  $(f, w_i)_0$  skalárszorzatokat alkalmas  $\varphi_i(t)$  közelítésekkel helyettesítjük. Ezután (16.92) a következő alakot veszi fel :

$$G\dot{y} + Ay = \varphi, \quad \varphi := (\varphi_1, \dots, \varphi_{N-1})^T. \quad (16.94)$$

A (16.94) egyenlet stabilitása a kezdetiértékek és a jobboldal szerint kényelmesen bizonyítható be az  $L_2$ -normában (ld. a 15. feladatot). Ebből következik, hogy az  $(f, w_i)_0$  jelzett változtatása  $\varphi$ -re az  $y$  becslhető változásához vezet. Mi itt viszont nem ezt a vonalat követjük, hanem csak majd lejjebb, az egyenlet teljes diszkretizációja után, foglalkozunk az eredő eljárás stabilitásával és pontosságával.

Az említett  $\varphi_i(t)$  közelítések leírásához a rövidebb  $f(x_i, t) =: f_i(t)$  jelölést vezetjük be. Ezután az  $(f, w_i)_0$ -k következő approximációit javasolhatjuk :

$$\varphi_i(t) := \begin{cases} hf_i(t), & \text{vagy} \\ \frac{h}{6}(f_{i+1}(t) + 4f_i(t) + f_{i-1}(t)), & \text{vagy pedig} \\ \frac{h}{12}(f_{i+1}(t) + 10f_i(t) + f_{i-1}(t)). \end{cases} \quad (16.95)$$

Ezek a képletek a következőképpen jönnek létre : az elsőhöz használjuk az

$$(f, w_i)_0 \approx f(x_i, t) \int_0^1 w_i(x) dx = hf_i(t)$$

approximációt, a másodikhoz  $f$ -et először  $V_h$ -beli interpoláltjával helyettesítjük :

$$f(x, t) \approx \sum_{j=1}^{N-1} f_j(t)w_j(x) =: f_h(x, t).$$

Ezután kapjuk, hogy

$$(f_h(\cdot, t), w_i)_0 = \frac{h}{6}(f_{i+1}(t) + 4f_i(t) + f_{i-1}(t)) = (G\vec{f}(t))_i.$$

Itt  $\vec{f}(t)$  az  $f_i(t)$  komponensekből összeállított  $N - 1$ -dimenziós vektor.

Végülis, az

$$(f_h(\cdot, t), w_i)_0 = h \left[ f(x_i, t) + \frac{h^2}{12} \frac{\partial^2 f}{\partial x^2}(x_i, t) + O(h^4) \right]$$

sorfejtés alapján (amelyhez csupán a II-ben a 175. oldalon lévő (60) kifejezést össze kell rakni az  $u_{\bar{x},i}$  szokásos sorfejtésével, ld. ott, 174.o., figyelembe véve, hogy  $u'' = -f$ ) és  $\frac{\partial^2 f}{\partial x^2}$  differenciahányadossal való helyettesítése után kapjuk az utolsó, (16.95)-ben szereplő képletet.

Míg a dimenzió nem túl nagy, (16.94) megoldására célszerű valamelyik, merev rendszerre használatos módszert alkalmazni, v.ö. a (16.27) egyenlet kapcsolatos tárgyalással. Az időderiváltat megszorzó  $G$  mátrixra nézve ld. a II. 10.5.8. pontot (most nincs az a probléma, hogy  $G$  szinguláris lenne).

A merevség itt jellemezhető  $G^{-1}A$  kondíciószámaival, amelyet II. 11.6.4. alapján elég pontosan tudjuk becsülni : onnan ismert az  $Ay = \lambda Gy$  sajátérték feladat megoldása, valamint az, hogy érvényes  $\pi^2 < \lambda < 12/h^2$  annak ott  $\lambda_{k,VE}^h$ -vel jelölt sajátértékeire. Így  $\text{cond}(G^{-1}A) < 12/(\pi^2 h^2)$ .

Alkalmazzuk most (16.94)-re a súlyozott differenciasémát! Ha (16.94) jobboldalát  $t_{j+\sigma} = (j + \sigma)\tau$ -nál vesszük, akkor ezt mint

$$G \frac{y^{j+1} - y^j}{\tau} + A [\sigma y^{j+1} + (1 - \sigma)y^j] = \varphi^{j+\sigma}$$

írhatjuk fel. Normálalakban, az  $y_t := \frac{y^{j+1}-y^j}{\tau}$  jelöléssel :

$$(G + \sigma\tau A)y_t + Ay^j = \varphi^{j+\sigma}. \quad (16.96)$$

Vagyis, most (16.45)-tel összehasonlítva,  $B_h$  helyett  $B = G + \sigma\tau A$  áll és  $A_h$  helyett  $A$ .

A (16.96) eljárást, tehát a végeelem módszer kombinációját a súlyozott differenciasémával, a rövidség kedvéért *végeelem sémának* hívjuk. Képlet-hibájával foglalkozik a 16. feladat.

**16.11. Tétel** (a végeelem séma stabilitása). Ha igaz

$$G + \sigma\tau A \geq \frac{\tau}{2}A,$$

akkor a (16.96) eljárás stabil a kezdetiértékek szerint, és ez a feltétel  $A = hA_h^{(0)}$  esetén  $\sigma \geq \frac{1}{2} - \frac{h^2}{12\tau}$  választásával biztosítható.

**Bizonyítás.** Az első állítás következik a 16.5 tételből. Ha  $A = hA_h^{(0)}$ , akkor vegyük észre, hogy  $G = hI_h - \frac{h^2}{6}A$  (ahol  $I_h \in \mathbb{R}^{N-1 \times N-1}$  az egység-mátrix), tehát

$$B = hI_h + \left(\sigma\tau - \frac{h^2}{6}\right)A \geq \frac{h}{\|A\|}A + \left(\sigma\tau - \frac{h^2}{6}\right)A \geq \frac{\tau}{2}A.$$

Ez teljesül, ha

$$\sigma \geq \frac{1}{2} + \frac{h^2}{6\tau} - \frac{h}{\|A\|\tau}, \text{ ahol } \|A\| = \lambda_{\max}(A) = \frac{4}{h} \cos^2 \frac{\pi h}{2} \approx \frac{4}{h}.$$

Ebből következik a tétel második állítása.  $\square$

**Megjegyzések. 1.** Eszerint  $\sigma \geq 1/2$  esetén az eljárás további feltétel nélkül stabil.  $\sigma = 0$  esetén viszont elegendő, ha  $\tau \leq \frac{h^2}{6}$ . Ez szigorúbb feltétel, mint a véges differencia módszer esetén.

**2.** A magasabbrendű eljárás (ld. 16. feladat) a  $\sigma = \frac{1}{2} + \frac{h^2}{12\tau}$  értéknek felel meg, amelyről tudni érdemes, hogy mikor igaz  $\sigma \leq 1$ . Ez akkor teljesül, amikor  $\tau \geq \frac{h^2}{6}$ . A stabilitás szempontjából  $\sigma = \text{const} > 1$  nem gond, de ekkor az approximáció általában romlik.  $\square$

A (16.96) séma stabilitásából, valamint a (16.89) feladat megoldásának simaságából következik az eljárás konvergenciája (17. feladat). Kevesebb simaság mellett 16.8-ban bizonyítjuk a konvergenciát.

### 16.4.7 Pozitivitástartás, maximumelv, konvergencia a maximum normában

Az alkalmazások szempontjából fontos tudni, vajon eljárásaink a maximum normában konvergálnak-e? Ezzel összefüggő téma az, vajon a diszkretizált feladatra érvényes a maximumelv, oszcillációmentes-e a megoldás, ill. vajon pozitív megoldásokat kapunk, ha ezt várhatjuk a feladat fizikai háttére alapján.

Mint példát az olyan alkalmazott problémákra, ahol ezek a kérdések lényegesek, tekintsük azt az inverz feladatot, hogy a konstans  $k$  hővezetési együtthatót a  $\frac{\partial u}{\partial t} = k \frac{\partial^2 u}{\partial x^2}$  egyenletből meg akarjuk határozni. Ekkor úgy szoktak eljárni, hogy a kísérletben a kezdeti feltételt mint lépcsőfüggvényt választják meg, pl. :

$$u_0(x) = \begin{cases} 1, & \text{ha } 0 \leq x < \frac{1}{2}, \\ 0, & \text{ha } \frac{1}{2} \leq x \leq 1, \end{cases} \quad (16.97)$$

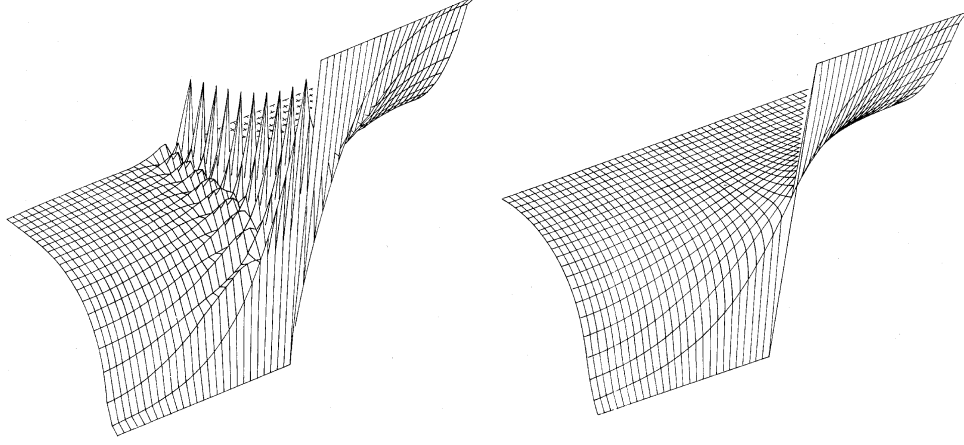
és ezt az  $u(0, t) = 1$ ,  $u(1, t) = 0$  peremfeltétellel egészítik ki. Méréssel állapítják meg az  $u(x, t)$  értékeit rögzített  $t = t_0 > 0$  időben, több  $x$ -értékre. Elegendő a  $\frac{\partial u}{\partial x}$  mérése is  $x = \frac{1}{2}$ -nél.

Ha ezt a feladatot numerikus úton meg akarjuk oldani, akkor nem megfelelő az, ha csak  $L_2(\omega_h)$ -ban vagy akár  $H^1(\omega_h)$ -ban stabil módszert választunk – pl. a Crank–Nicolson módszert  $\tau \approx h$  lépéstávolsággal (ehhez ld. a 18. feladatot). Ekkor ugyanis kicsi  $t$ -re olyan diszkrét megoldást kell biztosítani, amely  $x = \frac{1}{2}$  környékén oszcillációmentes – ami a Crank–Nicolson séma esetén gyakran nem igaz (ld. a 16.4. ábrát).

Ezeket a kérdéseket, tehát a maximum normában való konvergenciát és az oszcillációmentességet, vizsgáljuk az alábbiakban a (16.28) súlyozott differenciaséma esetén. Egyben a (16.96) végeelem sémát is elemezzük.

**Megjegyzés.** A 46. ábrán szemléltetett helyzetben, tehát amikor a kezdeti függvény nem folytonos, a megoldás pontosságának szempontjából érdemes Rannacher cikkének eredményét figyelembe venni, hogy ekkor az elején annyi lépést a tisztán implicit sémával tegyünk meg, amennyi az elliptikus operátor rendje (itt 2), majd ezután váltsunk Crank–Nicolson eljárásra.  $\square$

A (16.6)-(16.9) feladatot approximáló két eljárást a következőképpen írjuk



ábra 16.4: Crank–Nicolson séma (baloldalt) és tisztán implicit séma (jobb-oldalt) eredményei a (16.97) kezdeti függvényen,  $\tau = 1/40$ ,  $h = 1/50$ .

fel összefoglalóan :

$$\begin{aligned}
 B_h y_h^{j+1} &= C_h y_h^j + \tau \varphi_h^j, & j &= 0, 1, \dots, m-1, & (16.98) \\
 B_h &:= I_h + \tilde{\sigma} \tau A_h^{(0)}, & C_h &:= I_h - (1 - \tilde{\sigma}) \tau A_h^{(0)}, \\
 \tilde{\sigma} &:= \sigma - \mu, \quad \mu := \begin{cases} 0, & \text{a (16.28)} \\ \frac{h^2}{6\tau}, & \text{a (16.96)} \end{cases} \text{ séma esetén.}
 \end{aligned}$$

Itt  $A_h^{(0)}$  a (16.27) mátrix, ezenkívül, a  $\gamma := \tau/h^2$  jelöléssel,

$$\varphi_i^j := \begin{cases} f_i^{j+\sigma}, & \text{ha } 2 \leq i \leq N-2, \\ f_1^{j+\sigma} + \gamma[\tilde{\sigma} g_0^{j+1} + (1 - \tilde{\sigma}) g_0^j], & \text{ha } i = 1, \\ f_{N-1}^{j+\sigma} + \gamma[\tilde{\sigma} g_1^{j+1} + (1 - \tilde{\sigma}) g_1^j], & \text{ha } i = N-1. \end{cases} \quad (16.99)$$

A következő lemma elégséges feltételt fogalmaz meg arra, hogy a (16.98) egyenlet megoldása nemnegatív legyen.

**16.12. Lemma** (két súlyozott séma M-mátrix tulajdonsága). Amennyiben

$$1 \geq \tilde{\sigma} \geq \max(1 - 1/(2\gamma), 0), \quad (16.100)$$

akkor (16.98)-ban  $B_h$  M-típusú és  $C_h \geq 0$  elemenként.

**Bizonyítás.** 1) Tisztázzuk vajon  $B_h$  mikor M-mátrix? A főátlón kívüli elemek csak akkor nempozitívak, ha  $\tilde{\sigma} \geq 0$ . Ekkor

$$(B_h e_h)_i = 1 + \begin{cases} 0, & \text{ha } 1 \leq i \leq N-1, \\ \gamma \tilde{\sigma}, & \text{ha } i = 1, N-1, \end{cases}$$

ahol  $e_h := (1, \dots, 1)^T$ . Így  $B_h e_h \geq e_h > 0$ , ha  $\tilde{\sigma} \geq 0$ , azaz  $B_h$  M-mátrix.

2) A  $C_h = (c_{ij})$  mátrix nemnegatív elemeenként akkor, amikor

$$c_{i,i\pm 1} = \gamma(1 - \tilde{\sigma}) \geq 0 \text{ és } c_{ii} = 1 - 2\gamma(1 - \tilde{\sigma}) \geq 0.$$

Innen adódik (16.100).  $\square$

**Megjegyzések. 1.** A (16.28) súlyozott differenciaséma esetén (16.100) azt jelenti, hogy

$$0 \leq \sigma \leq 1, \quad \tau \leq \frac{h^2}{2(1 - \sigma)}. \quad (16.101)$$

Itt most fontos lett az, hogy  $\sigma \in [0, 1]$ . A  $\sigma = 1$  esetet úgy kell érteni, hogy akkor nincs feltétel  $\tau$ -ra.

**2.** Mindkét séma teljesíti a feltételeket, ha  $\sigma = 1$ .

**3.** A végeselem séma esetén (16.100) azt jelenti, hogy vagy

$$1 < \sigma \text{ és } \tau \leq \frac{h^2}{6(\sigma - 1)},$$

vagy pedig

$$\frac{1}{3} \leq \sigma \leq 1 \text{ és } \frac{h^2}{6\sigma} \leq \tau \leq \frac{h^2}{3(1 - \sigma)}.$$

Tehát az explicit séma ki van zárva. Egyébként pl. adódik

$$\frac{h^2}{3} \leq \tau \leq \frac{2h^2}{3}, \text{ ha } \sigma = \frac{1}{2}, \text{ és } \frac{h^2}{6} \leq \tau \leq \frac{5h^2}{6}, \text{ ha } \sigma = \sigma_* = \frac{1}{2} + \frac{h^2}{12\tau}.$$

**4.** Amennyiben  $\sigma > 0$  és  $\tau < h^2/6\sigma$ , a  $(G + \sigma\tau A)^{-1}(G - (1 - \sigma)A)$  mátrix valóban tartalmaz negatív elemeket (sőt : a negatív és pozitív elemek váltakoznak sakktáblaszerű elrendezésben). Tehát  $\tau \geq h^2/6\sigma$  szükséges! Ehhez ld. Faragó I. és tanítványainak cikkeiben a kérdéskör részletes tárgyalását.  $\square$

Differenciálegyenlet-rendszerek megoldásainak *nemnegativitásával* foglalkoztunk II-ben, ld. ott a 10.4. lemmát és a 10.5. tételt. Azok szerint egyébként a (16.27) szemidiszkrétizált eljárás mindig nemnegatív megoldásokat ad (ha nemnegatív a kezdetiérték és a jobboldal), míg a (16.94) nem! Ez utóbbi állítás igazolása a 19. feladat.

Természetes a fogalom átvitele a teljesen diszkrétizált módszerekre.

**Definíció.** A (16.98) eljárást akkor hívjuk *pozitivitástartónak*, ha  $y_h^j \geq 0$  minden  $j$ -re következik abból, hogy érvényes  $y_h^0 \geq 0$  és  $\varphi_h^j \geq 0$  minden  $j$ -re.  $\square$

**16.13. Tétel** (két súlyozott séma pozitívítástartása). A 16.12. lemma feltételei mellett a (16.28) súlyozott differenciaséma és a (16.96) végeelem séma pozitívítástartóak.

**Bizonyítás.** Mivel ekkor (16.99)-ben  $0 \leq \tilde{\sigma} \leq 1$ , így  $\varphi_h \geq 0$  komponensenként. Továbbá,  $B_h^{-1} \geq 0$  (ld. I. 1.3.4., 1.7. lemma), így  $B_h^{-1}C_h \geq 0$  és

$$y_h^{j+1} = B_h^{-1}C_h y_h^j + \tau B_h^{-1}\varphi_h^j \geq 0, \text{ ha } y_h^0 \geq 0. \quad \square \quad (16.102)$$

**Megjegyzések. 1.** A  $\varphi_h$  jobboldal miatt (ld. (16.102))  $B_h^{-1} \geq 0$  szükséges a pozitívítástartáshoz, de  $C_h \geq 0$  nem szükséges, hanem  $B_h^{-1}C_h \geq 0$ .

**2.** A fenti tétel akkor garantálja a pozitívítástartást, ha vagy elsőrendű a módszer (mint a tisztán implicit séma), vagy pedig  $\tau = O(h^2)$ . Magasabbrendű algoritmusokat pl. Horváth Z. II-ben idézett cikkeiben lehet találni.  $\square$

Ehhez a tételhez ld. a 20. feladatot is.

Vannak olyan esetek, amikor a negatív jobboldal mellett pozitív maradt a megoldás. Így az  $\frac{dy}{dt} = qy$  egyenlet megoldása (negatív  $q$  mellett is) pozitív, míg a kezdetiérték az. Általánosabban, tekintsük a

$$\frac{\partial u}{\partial t} = \mathcal{L}u + f \quad (16.103)$$

parabolikus egyenletet akkor, amikor jobboldala a megoldástól is függhet :  $f = f(u, x, t)$ , és

$$f = f_{(+)} - f_{(-)}, \quad f_{(+)} \geq 0, \quad f_{(-)} \geq 0$$

alakú. Ekkor az alábbi, *Patankar*-ra visszamenő fogással biztosíthatjuk a numerikus megoldás pozitívítását : az approximáció során az  $f^j := f(y^j, x, t_j)$  forrástagot a

$$f_{(+)}^j - \frac{f_{(-)}^j}{y^j} y^{j+1} = f_{(+)}^j - \frac{f_{(-)}^j}{y^j} (y^j + \tau y_t) = f^j + O(\tau)$$

kifejezéssel helyettesítjük. Azaz, szokásunk szerint az  $f_1^j, \dots, f_{N-1}^j$  értékekből összeállított vektort  $\vec{f}_h^j$ -vel jelölve, az  $y_t + A_h y_h^{j+1} = \vec{f}_h^j$  tisztán implicit séma helyett (ahol  $A_h$  M-típusú legyen) a következőt vesszük :

$$y_t + A_h y_h^{j+1} = (\vec{f}_{(+)}^j)_h - F_h y_h^{j+1}, \text{ vagyis } y_t + (A_h + F_h) y_h^{j+1} = (\vec{f}_{(+)}^j)_h,$$





alakban írhatjuk le. Mint majoráns vektort a

$$w(x) := 1 + \frac{1}{2}x(1-x) \quad (16.105)$$

függvény  $\omega_h$ -beli értékeiből állítjuk össze először a  $g_h := \vec{w}_h \in \mathbb{R}^{N-1}$  vektort, majd a peremértékeket is figyelembe véve a  $\tilde{g}_h := (1, g_h^T, 1)^T \in \mathbb{R}^{N+1}$  vektort, végül ebből a  $\tilde{g} := (\tilde{g}_h, \tilde{g}_h, \dots)^T$  blokkvektort, tehát minden időrétegen ezek a  $\tilde{g}_h$  vektorok ismétlődnek. Ekkor

$$\mathcal{A}\tilde{g} \geq \tilde{e}, \quad (16.106)$$

ahol  $\tilde{e} := (\tilde{e}_h, \tilde{e}_h, \dots)^T$  az a blokkvektor, amely az  $\bar{\omega}_{h,\tau}$  rács minden pontjához az 1-t tartalmazza. Ugyanis a peremen és a  $t = 0$  rétegen visszkapjuk az  $\mathcal{A}$  blokkmátrixnak a  $\tilde{g}$  vektorral való szorzásánál a  $\tilde{g}$  vektor ottani komponensének 1 értékét, és a belső pontokban a  $B_h g_t + A_h g_h$  differencia-kifejezés működik. Mivel  $\tilde{g}$ -ben minden  $j$ -re ugyanaz a  $\tilde{g}_h$  vektor áll, itt  $g_t = 0$ , ezenkívül  $A_h g_h = e_h$ , mert a perempontokban itt is visszkapjuk az értéket, a belső pontokban viszont  $(A_h g_h)_i = -w''(x_i) = 1$ , v.ö. pl. 15.3.2-ben (15.20)-szal. (16.106) azt jelenti, hogy  $\mathcal{A}$  M-mátrix. Ekkor az I. 1.8. lemma szerint (v.ö. pl. a 15.2. tétellel is) érvényes az

$$\begin{aligned} \|\mathcal{A}^{-1}\|_{C(\bar{\omega}_{h,\tau})} &\leq \frac{\|\tilde{g}\|_{C(\bar{\omega}_{h,\tau})}}{\min_{(x,t) \in \bar{\omega}_{h,\tau}} (\mathcal{A}\tilde{g})} \\ &= \|\tilde{g}_h\|_{C(\bar{\omega}_h)} \leq \|w\|_{C[0,1]} = \frac{9}{8} \end{aligned} \quad (16.107)$$

egyenlőtlenség. Pontosabb becslés érdekében írjuk  $y_h$ -t az  $y_\gamma + y_\omega$  alakban, ahol

$$B(y_\gamma)_t + Ay_\gamma = 0, \quad (y_\gamma)|_{\Gamma_T} = y|_{\Gamma_T},$$

és

$$B(y_\omega)_t + Ay_\omega = \varphi, \quad (y_\omega)|_{\Gamma_T} = 0.$$

Ekkor, 15.3.2-ben a (15.29) becslés bizonyításához hasonlóan eljárva, megkapjuk a (16.104) állítást.  $\square$

**Megjegyzések. 1.** A fenti vizsgálatot a következőképpen lehet egyszerűsíteni. Tekintsük pl. a (16.28)-(16.30) differenciasémát! Először az összes együtthatót írjuk fel, amely az  $(x_i, t_{j+1})$  pont egyenletéhez tartozik :

$$\begin{aligned} y_i^{j+1} &: \frac{1}{\tau} + \frac{2\sigma}{h^2}, & y_i^j &: -\frac{1}{\tau} + \frac{2(1-\sigma)}{h^2}, \\ y_{i\pm 1}^{j+1} &: -\frac{\sigma}{h^2}, & y_{i\pm 1}^j &: -\frac{1-\sigma}{h^2}. \end{aligned}$$

Itt az első,  $y_i^{j+1}$ -vel kapcsolatos együttható kell, hogy pozitív legyen (mert ez kerül a mátrix főátlójára) – ami teljesül, a többi együttható kell, hogy nempozitív legyen. Ebből azonnal adódik a (16.101) feltétel. Ezzel megvan az M-mátrixok előjeleloszlása.

Ezután a (16.105) függvényt magába a (16.103) egyenlet  $\frac{\partial u}{\partial t} - \mathcal{L}u$  differenciálkifejezésébe helyettesítjük be  $u$  helyett, hiszen ez (míg  $w$  legfeljebb harmadfokú polinom  $x$ -ben és lineáris  $t$ -ben) ugyanazt az eredményt adja (ez itt 1), mint a differenciaséma. Mivel majoráns függvényünk a  $\Gamma_T$  peremen sem kisebb 1-nél, így az M-mátrix tulajdonság és a (16.107) becslés következnek, abból tovább az

$$\|y_h\|_{C(\bar{\omega}_{h,\tau})} \leq \frac{9}{8} \max(\|y_h\|_{C(\Gamma_{h,T})}, \|\varphi_h\|_{C(\omega_{h,\tau})})$$

egyenlőtlenség. Ugyanolyan jó szolgálatot a  $w = 1+t$  függvény is tesz. Ehhez ld. a 21. feladatot is.

**2.** Az  $\mathcal{A}$  mátrix M-tulajdonságából az is következik, hogy az eljárás *pozitivitástartó*, v.ö. a 16.13. tétellel (valamint a 15.1. lemmával és a 15.8. tétellel). Így, ha  $f \geq 0$ ,  $g_0 \geq 0$ ,  $g_1 \geq 0$ , akkor  $y \geq 0$ . Továbbá, (16.104) helyett, ha pl.  $f \equiv 0$ , minden  $(x_i, t_j) \in \bar{\omega}_{h,\tau}$ -ra érvényes

$$\begin{aligned} \min \left( \min_{x \in \bar{\omega}_h} u_0(x), \min_{1 \leq k \leq j} (g_0(t_k), g_1(t_k)) \right) \\ \leq y_i^j \leq \max \left( \max_{x \in \bar{\omega}_h} u_0(x), \max_{1 \leq k \leq j} (g_0(t_k), g_1(t_k)) \right). \end{aligned} \quad (16.108)$$

**3.** Eddig az *oszcillációmentességről* nem kaptunk semmi eredményt. Bizonyos oszcillációmentességet viszont már a maximumelv is biztosít! Ennek belátásához képzeljük el, hogy az  $f \equiv 0$  esetben a (16.96) eljárást alkalmazva a  $j+1 \geq 1$  időrétegen, belső pontban kaptunk egy lokális maximumot, azaz

$$\max(y_{i\pm 1}^j, y_i^j, y_{i\pm 1}^{j+1}) < y_i^{j+1}. \quad (16.109)$$

Ekkor az  $y_{i-1}^j, y_i^j, y_{i+1}^j$  adatokat mint kezdetiértéket tekinthetjük, és  $y_{i-1}^{j+1}$ -et mint baloldali,  $y_{i+1}^{j+1}$ -et mint jobboldali peremértéket. A (16.96) eljárást lokálisan alkalmazva, ugyanazt a megoldást kapjuk  $(x_i, t_{j+1})$ -ben mint előbb, tehát  $y_i^{j+1}$ -et, hiszen helyben hárompontos, időben kétréteges sémáról van szó. Ekkor tételünk és (16.108) azt mutatják, hogy (16.109) lehetetlen.

**4.** A Crank–Nicolson eljárás  $\tau \leq h^2$  esetén teljesíti a feltételeket. Máskor csillapított oszcillációkkal számíthatunk (ld. a 16.4. ábrát). Viszont ismert, hogy tetszőleges  $\tau, h$ -ra valamint tetszőleges  $y_h^0$  kezdetiérték esetén, homogén jobboldal és peremértékek mellett igaz az

$$\|y_h^j\|_{C(\omega_h)} \leq M \|y_h^0\|_{C(\omega_h)}$$

becslés. (Szerdjukova eredményei, hogy itt  $M \leq 6$ , és hogy ugyanilyen becslés igaz a (16.37) séma esetén is, ekkor  $M \leq 23$ , ld. a 22. feladatot is.)  $\square$

A diszkrét maximumelvhez nagyon közel áll a következő tulajdonság :

Ha (16.6)-ban  $f \equiv 0$ ,  $\mathcal{L}u \equiv \frac{\partial^2 u}{\partial x^2}$  és  $u_0(x)$  monoton (nem növekszik vagy nem csökken  $x$  növekedésével), akkor megfelelő peremfeltételek esetén az  $u(x, t)$  megoldás is monoton minden rögzített  $t$ -re, tehát a monotonitás megmarad (és oszcillációk nincsenek).

Amennyiben (16.6)-ban viszont

$$\mathcal{L}u := \frac{\partial}{\partial x} \left( k(x) \frac{\partial u}{\partial x} \right),$$

akkor a  $w := k \frac{\partial u}{\partial x}$  behelyettesítéssel látjuk, hogy igaz

$$\frac{\partial}{\partial x} u_t = \frac{1}{k} w_t = \frac{\partial^2 w}{\partial x^2},$$

tehát  $w$ -re teljesül a maximumelv, így  $\frac{\partial u}{\partial x}$  monotonitása biztosítható (megfelelő peremfeltételek esetén).

Ha egy (kétszintes) differenciaséma ugyanezzel a tulajdonsággal rendelkezik, tehát hogy az  $x \in \omega_h$ -ban monoton növekvő  $y^j$  rácsfüggvényhez tartozik ugyancsak monoton növekvő  $y^{j+1}$  rácsfüggvény a következő időszinten, akkor azt mondjuk, hogy a *monotonitás megmarad*, ill. azt, hogy a séma *monoton Godunov értelmében*. Ehhez ld. a \*\*\* 18.10.1. pontot is.

Meg lehet mutatni, hogy a súlyozott differenciaséma esetében, ha  $k \equiv 1$ , a monotonitás megmaradása akkor igaz, amikor érvényes a maximumelv, vagyis ha

$$\text{vagy } \sigma = 1, \text{ vagy } 0 \leq \sigma < 1 \text{ és } \tau \leq h^2/(2(1 - \sigma)).$$

Ha viszont  $0 < k \neq \text{const}$  és  $\mathcal{L}u$  approximációja történik a szokásos módon (ld. II. 11.4.8. vagy 16.5.1), akkor a differenciasémára igaz a maximumelv, ha

$$1 \geq (1 - \sigma) \frac{\tau}{h^2} (k_{i+1/2} + k_{i-1/2}), \quad 1 \leq i \leq N - 1,$$

míg a monotonitás megmarad, ha

$$1 \geq 2(1 - \sigma) \frac{\tau}{h^2} k_{i+1/2}, \quad 1 \leq i \leq N - 2, \text{ és ha } 1 \geq (1 - \sigma) \frac{\tau}{h^2} k_{i+1/2}, \quad i = 1, N.$$

Így lehetséges, hogy a maximumelv teljesül, a monotonitás viszont nem marad meg – és megfordítva. A monotonitás megmaradására \*\*\* 18.10.2-ben térünk vissza.

### 16.4.8 A tömegmátrix kiszámításáról

Ahogy az előbbiekből láthatjuk (vö. a 16.12. lemmával és a pozitivitástartás definíciója előtti megjegyzésekkel), a végeelem módszerrel kapcsolatos szemidiszkretizált ill. teljesen diszkretizált egyenletnek hátrányos tulajdonságai is vannak. Ezeket még tetézi, hogy a (16.94) egyenletben szereplő  $G$  tömegmátrix bár jól kondicionált – ld. II-ben a 11.6.2. pont végét és a megjegyzést a 15.38. lemmához – de (16.94) numerikus megoldása szempontjából kellemetlen: a (16.94) egyenlet implicit  $y_t$ -re nézve, és azt a  $G^{-1}$ -gyel megszorozni nem tanácsos, mert az telt mátrix. Ugyanolyan viszonyokkal találkozunk a (16.6)-(16.9) feladatnál bonyolultabb, pl. többdimenziós feladatok esetén. Hasonló, a tömegmátrixszal kapcsolatos problémák a sajátérték feladatok megoldásánál is fellépnek.

Emiatt a végeelem gyakorlatban lényeges a  $G$  mátrix kiszámítását figyelemmel kísérni, és szükség esetén  $G$ -t olyan  $\tilde{G}$  approximációval helyettesíteni, hogy az diagonális legyen (ekkor  $\tilde{G}$  angolul a „lumped mass matrix”). Ez elérhető, ha (az integrálok egzakt kiszámítása, ill. Gauss-képletek alkalmazása helyett) a kvadratúra képlet súlyai pozitívak és alappontjai egybeesnek a végeelem csomópontjaival (amitől egyben javul a kvadratúra képletek hatékonysága, ld. a 15.7.6. pontot, mert a csomópontok javarészt az elemek peremén fekszenek): ugyanis (15.211) miatt – tehát azért, mert bármely bázisfüggvény az összes csomópontban nulla, kivéve „saját” csomópontjában – érvényes

$$g_{kl} = \int_{\Omega} w_k(x) w_l(x) dx \approx \sum_{i=1}^M \sum_{j=1}^{n_i} \omega_{i,j} p_m(x^{i,j}) p_n(x^{i,j}) = \tilde{g}_{kl} = 0, \text{ ha } k \neq l.$$

Itt  $\omega_{i,j}$  jelöli a kvadratúra képlet az  $x^{i,j}$  alapponthez tartozó súlyát, egyébként használtuk a 15.7.4. és 15.7.7. pont (ld. (15.240)) jelöléseit:  $(m, n)$  és  $(k, l)$  között az az összefüggés, hogy  $k = \text{ind}(i, m)$ ,  $l = \text{ind}(i, n)$ , és ha ilyen összefüggés nincs, akkor már  $w_k(x) w_l(x) \equiv 0$  miatt érvényes  $g_{kl} = \tilde{g}_{kl} = 0$ .

A főátlóbeli  $\tilde{g}_{kk}$  elemek ekkor garantáltan pozitívak. Példák ilyen kvadratúra képletre az egydimenziós esetben a trapézszabály (ha lineáris elemekkel dolgozunk, ld. a 23. feladatot), a kétdimenziós esetben a 15.7.6. pontban a  $T_3$  és az  $R_4$  elemek számára említett (15.236), ill. (15.238) képletek.

Viszont az ugyanott a  $T_6$  elem (és hasonlóan az  $R_8$ ) számára javasolt kvadratúra képlet nem tartozik a fentiek szerint elfogadható képletek közé, ha parabolikus vagy sajátérték feladatokról van szó. Ugyanis a háromszög csúcsai hiányoznak a képletből, a nekik megfelelő  $w_k$  bázisfüggvények ezért nem adhatnak semmi járulékot a tömegmátrixhoz, így  $\tilde{g}_{kk}$  nulla lesz! Ez a kvadratúra képlet még nem is javítható bizonyos értelemben (ld. a 24. feladatot). Ilyen esetekben  $\tilde{G}$  előállításához a  $G$  mátrix  $k$ -adik sorának összegét

írják  $g_{kk}$  helyére és a sor többi elemét lenullázzák – ami gyakran, de nem mindig megfelelő eredményeket ad (ld. ehhez a 25. feladatot).

### 16.4.9 Vizsgálatok Fourier–módszer segítségével

Most vizsgáljuk a

$$B_h y_t + A_h y_h = \varphi_h, \quad B_h = B_h^T > 0, \quad A_h = A_h^T > 0, \quad (16.110)$$

kétréteges differenciaséma megoldását a Fourier–sorfejtés segítségével, amikor a peremfeltételek homogén elsőfajúak. Példaként tekintsük a (16.6)–(16.9) feladat diszkretizációját a (16.28)–(16.30) súlyozott differenciasémával, ill. a (16.96) végeelem sémával.

Emlékeztetünk arra, hogy a

$$\mathcal{L}u := -u'' = \lambda u, \quad 0 < x < 1, \quad u(0) = u(1) = 0 \quad (16.111)$$

sajátérték feladat egzakt megoldása

$$\lambda_k = (k\pi)^2, \quad \mu_k(x) := \sin k\pi x, \quad k = 1, 2, \dots$$

A diszkrét esetben legyen  $v_h^{(k)}$  az a vektor, amelynek komponensei  $\mu_k(x_i)$ . Ezután

$$A_h^{(0)} v_h^{(k)} = \lambda_k^h v_h^{(k)}, \quad \lambda_k^h = \left( \frac{2}{h} \sin \frac{k\pi h}{2} \right)^2, \quad k = 1, \dots, N-1, \quad h = 1/N, \quad (16.112)$$

ahol  $A_h^{(0)}$  a (16.27) mátrix. A (16.111) feladat megoldását az

$$y_h = \sum_{k=1}^{N-1} y_k(t) v_h^{(k)}$$

alakban keressük, ami a diszkrét Fourier–transzformációnak felel meg. Hasonlóan legyen

$$\varphi_h = \sum_{k=1}^{N-1} \varphi_k(t) v_h^{(k)}.$$

a) Véges differencia módszer:  $B_h = I_h + \sigma\tau A_h$ , és legyen  $\sigma \in [0, 1]$ . Ekkor  $(y_k(t_j))$  helyett  $y_k^j$ -t írva stb.)

$$\sum_{k=1}^{N-1} v_h^{(k)} [(1 + \sigma\tau\lambda_k^h)(y_k)_t + \lambda_k^h y_k^j - \varphi_k^j] = 0.$$

Mivel a sajátvektorok ortogonális rendszert alkotnak, ebből következik, hogy

$$y_k^{j+1} = \left(1 - \frac{\tau \lambda_k^h}{1 + \sigma \tau \lambda_k^h}\right) y_k^j + \tau \frac{\varphi_k^j}{1 + \sigma \tau \lambda_k^h} = \rho_k y_k^j + \tau \frac{\varphi_k^j}{1 + \sigma \tau \lambda_k^h},$$

ahol

$$\rho_k := \frac{1 - (1 - \sigma)\tau \lambda_k^h}{1 + \sigma \tau \lambda_k^h}, \quad k = 1, \dots, N-1. \quad (16.113)$$

Mint elégséges stabilitási feltételt, vehetjük a

$$\rho_k \in [-1, 1], \quad k = 1, \dots, N-1,$$

feltételt, ami  $\lambda_k^h > 0$  miatt azt jelenti, hogy  $\sigma \geq \frac{1}{2} - 1/(\tau \lambda_k^h)$ . A legerősebb feltétel a  $k = N-1$  esetben adódik, és  $\lambda_{N-1}^h = \|A_h\|_{(0,h)} = \left(\frac{2}{h} \cos \frac{\pi h}{2}\right)^2$ . (Vegyük figyelembe, hogy  $h \leq \frac{1}{2}$ !) Ekkor kell, hogy

$$\sigma \geq \frac{1}{2} - \frac{1}{\tau \|A_h\|_{(0,h)}}$$

legyen, ld. (16.55). A fenti képletek legtöbbje akkor is érvényes, amikor  $\mathcal{L}$  (16.111) helyett többdimenziós, önadjungált elliptikus operátor és (16.110)-ben  $A_h$  a diszkretizáció szimmetrikus és pozitív definit mátrixa. Speciális esetünkben

$$\sigma \geq \frac{1}{2} - \frac{h^2}{4\tau \cos^2 \frac{\pi h}{2}}, \text{ és picivel erősebben } \sigma \geq \frac{1}{2} - \frac{h^2}{4\tau}$$

a stabilitási feltétel. Továbbá kapjuk (mivel ekkor  $|\rho_k| \leq 1$ ), hogy

$$|y_k^{j+1}| \leq |y_k^j| + \tau |\varphi_k^j| \leq |y_k^0| + \tau \sum_{\ell=0}^j |\varphi_k^\ell|,$$

és így

$$|y_k^{j+1}|^2 \leq 2 \left[ (y_k^0)^2 + \tau^2 \left( \sum_{\ell=0}^j |\varphi_k^\ell| \right)^2 \right] \leq 2 \left[ (y_k^0)^2 + t_{j+1} \sum_{\ell=0}^j \tau (\varphi_k^\ell)^2 \right].$$

Innen következik a Parseval-azonosság, azaz

$$\|y_h^{j+1}\|_{(0,h)}^2 = \sum_{k=1}^{N-1} (y_k^{j+1})^2$$

segítségével, hogy

$$\|y_h^{j+1}\|_{(0,h)}^2 \leq 2 \left[ \|y_h^0\|_{(0,h)}^2 + t_{j+1} \sum_{\ell=0}^j \tau \|\varphi_h^\ell\|_{(0,h)}^2 \right].$$

Vagyis

$$\|y_h^{j+1}\|_{(0,h)} \leq \sqrt{2} \left[ \|y_h^0\|_{(0,h)} + \left( T \sum_{\ell=0}^j \tau \|\varphi_h^\ell\|_{(0,h)}^2 \right)^{1/2} \right].$$

Ez a becslés több vonatkozásban javítható. Pl. a  $\sqrt{2}$  szorzó felesleges. Használva a feladat linearitását, az  $y_h$  megoldást szokásos módon két részre bontjuk fel,  $y_h = v_h + w_h$ .

Itt  $v_h$  az inhomogén kezdeti feltételt elégíti ki,  $w_h$  viszont az inhomogén jobboldalt.

$\varphi_h \equiv 0$  esetén beláthatjuk azt, hogy

$$\|v_h^{j+1}\|_{(0,h)} \leq \|v_h^0\|_{(0,h)} = \|y_h^0\|_{(0,h)},$$

míg  $\varphi_h \not\equiv 0$  esetén következik

$$\|w_h^{j+1}\|_{(0,h)} \leq \|w_h^j\|_{(0,h)} + \tau \|\varphi_h^j\|_{(0,h)}$$

közvetlenül a

$$w_h^{j+1} = B_h^{-1} [(I_h - (1 - \sigma)\tau A_h)w_h^j + \tau\varphi_h^j]$$

egyenletből, figyelembe véve, hogy

$$\|B_h^{-1}(I_h - (1 - \sigma)\tau A_h)\|_{(0,h)} = \max_k |\rho_k| \leq 1$$

és

$$\|B_h^{-1}\|_{(0,h)} = \max_k (1/(1 + \sigma\tau\lambda_k^h)) \leq 1.$$

Ez azt jelenti, hogy

$$\|w_h^{j+1}\|_{(0,h)} \leq \tau \sum_{\ell=0}^j \|\varphi_h^\ell\|_{(0,h)},$$

és  $\|y_h\|_{(0,h)} \leq \|v_h\|_{(0,h)} + \|w_h\|_{(0,h)}$  miatt végülis

$$\|y_h^{j+1}\|_{(0,h)} \leq \|y_h^0\|_{(0,h)} + \tau \sum_{\ell=0}^j \|\varphi_h^\ell\|_{(0,h)}.$$



**Megjegyzés.** Tehát itt derült ki, hogy  $L_2(\omega_h)$ -ban a súlyozott differenciaséma stabil kezdetiértékek és jobboldal szerint, ha igaz a  $\sigma \geq \frac{1}{2} - \frac{1}{\tau\|A_h\|}$  feltétel – ami gyengébb, mint a korábbi  $\sigma \geq \frac{1}{2} - \frac{1-\varepsilon}{\tau\|A_h\|}$ ,  $0 < \varepsilon \leq 1$  feltétel.  $\square$

b) A végeelem módszer esetén  $\frac{1}{h}G = I - \frac{h^2}{6}A_h$  miatt (tehát a  $h$ -val való osztás után) a sajátértékeket  $\lambda_k^h = \lambda_{k,VD}^h = \frac{4}{h^2} \sin^2 \frac{k\pi h}{2}$  helyett a

$$\bar{\lambda}_k^h = \lambda_{k,VE}^h := \frac{\lambda_k^h}{1 - \frac{h^2}{6}\lambda_k^h}$$

képlet adja meg, és II. 11.4.6-ból ismert  $0 < \lambda_k^h < \bar{\lambda}_k^h < \frac{12}{h^2}$ . Ekkor a stabilitási feltétel

$$\bar{\rho}_k := \frac{1 - (1 - \sigma)\tau\bar{\lambda}_k^h}{1 + \sigma\tau\bar{\lambda}_k^h} \in [-1, 1], \quad k = 1, \dots, N-1,$$

ami teljesül, ha  $\sigma \geq \frac{1}{2} - h^2/(12\tau)$  – mint korábban, ld. a 16.11. tételt. Most ez megint a jobboldal szerinti stabilitásra is elegendő.

A fenti eredményekben az tipikus, hogy a *Fourier-módszer szinte azonos, de valamivel pontosabb feltételeket ad az energia-becslések módszeréhez képest.*

**Hosszú időintervallumon való számítás** esetén fontos, hogy az  $y_h^j$  megoldás lényegében úgy viselkedik, mint az első sajátvektornak megfelelő komponense, mert a pontos megoldás is ezzel a tulajdonsággal rendelkezik. Ezt a tulajdonságot a következő definícióval próbáljuk megfogni.

**Definíció.** A (16.28)-(16.30),  $g_0 = g_1 = 0, \varphi_h \equiv 0$  által meghatározott diszkrét parabolikus kezdetiérték feladat  $y_h^j = \sum_{k=1}^{N-1} \rho_k^j y_k^0 v_h^{(k)}$  megoldását akkor hívjuk *aszimptotikusan stabilnak*, ha  $|\rho_k| \leq |\rho_1| \leq 1$  igaz  $k = 2, \dots, N-1$ -re, a (16.113)-mal definiált  $\rho_k$  mennyiségekkel.  $\square$

**Megjegyzés.** A (pozitív)  $\lambda_k^h$  sajátértékek  $k$ -val monoton módon növekednek, amiért  $\rho_{k+1} < \rho_k < \rho_1$  mindig teljesül, és  $\rho_1 < 1$ . Ha  $\rho_1 \leq 0$ , akkor  $\rho_{N-1} < \dots < \rho_2 < \rho_1 \leq 0$  és a megoldás nem lehet aszimptotikusan stabil. Viszont  $\rho_1 > 0$  a  $\tau < 1/((1 - \sigma)\lambda_1^h)$  egyenlőtlenséggel biztosítható. Ennek teljesülését a következőkben mindig feltételezzük. Ezután az aszimptotikus stabilitás lényeges feltétele  $-\rho_1 \leq \rho_{N-1}$ .  $\square$

Számítsuk ki, mit jelent a  $-\rho_1 \leq \rho_{N-1}$  feltétel! Először

$$\begin{aligned} & -(1 + \sigma\tau\lambda_1^h)(1 + \sigma\tau\lambda_{N-1}^h) + \tau\lambda_1^h(1 + \sigma\tau\lambda_{N-1}^h) \\ & \leq (1 + \sigma\tau\lambda_{N-1}^h)(1 + \sigma\tau\lambda_1^h) - \tau\lambda_{N-1}^h(1 + \sigma\tau\lambda_1^h), \end{aligned}$$

ahonnan azt kapjuk, hogy

$$\begin{aligned} \tau(\lambda_1^h + \lambda_{N-1}^h) + 2\sigma\tau^2\lambda_1^h\lambda_{N-1}^h &\leq 2(1 + \sigma\tau\lambda_{N-1}^h)(1 + \sigma\tau\lambda_1^h) \\ &= 2 + 2\sigma\tau(\lambda_1^h + \lambda_{N-1}^h) + 2\sigma\tau^2\lambda_1^h\lambda_{N-1}^h, \end{aligned}$$

tehát

$$0 \leq 2 + (2\sigma - 1)\tau(\lambda_1^h + \lambda_{N-1}^h) + 2\sigma(\sigma - 1)\tau^2\lambda_1^h\lambda_{N-1}^h. \quad (16.114)$$

Amennyiben itt  $\sigma = 1$ , úgy az egyenlőtlenség minden  $\tau > 0$ -ra igaz, tehát a tisztán implicit séma aszimptotikusan stabil.

$0 \leq \sigma < 1$  esetén következik a  $\tau \leq \tau_\sigma$  feltétel, ahol  $\tau_\sigma$  a (16.114) jobb-oldalának (mint  $\tau$ -ban másodfokú polinom) pozitív gyöke. Tehát a súlyozott séma ekkor csak emellett a feltétel mellett aszimptotikusan stabil.

Ha  $\sigma = 0$ , akkor ez azt a feltételt jelenti  $\tau$ -ra, hogy

$$\tau \leq \tau_0 := 2/(\lambda_1^h + \lambda_{N-1}^h).$$

Ez a felső  $\tau$ -határ jól ismert érték : az egyszerű iteráció optimális paramétere, ld. I. 1.6.6.!

Ha  $\sigma = 1/2$ , akkor viszont a feltétel

$$\tau \leq \tau_{1/2} = 2/(\lambda_1^h\lambda_{N-1}^h)^{1/2} = O(h). \quad (16.115)$$

Amennyiben  $0 < \tau \leq \tau_\sigma$ , tehát ha a séma aszimptotikusan stabil, akkor ezt a tulajdonságot  $\tau \in \Omega_\sigma$  alakban is írjuk.

A továbbiakban foglalkozunk a  $\sigma = 1/2$  esettel. Legyen  $\lambda_1$  egy önadjungált  $\mathcal{L}$  differenciáloperátor első sajátértéke és  $\lambda_1^h$  az  $\mathcal{L}$  operátorhoz tartozó diszkretizáció  $A_h$  (szimmetrikus, pozitív definit) mátrixának első sajátértéke. A következő tétel az olyan időlépést ad meg, amely nemcsak az  $\Omega_{1/2}$  aszimptotikus stabilitási tartományban van, hanem amellyel az  $\|y_h^j\|_{(0,h)}$  úgy csökken nullához, mint  $e^{-\lambda_1 t_j}$  – a pontos megoldással egybehangzóan.

**16.15. Tétel** (a Crank–Nicolson eljárás aszimptotikus stabilitása). Ha  $\sigma = 1/2$ ,  $h \leq h_0$ , és

$$\lambda_1^h < \lambda_1 \leq \lambda_1^h + (\lambda_1^h)^2 / (3\lambda_{N-1}^h), \quad \frac{\lambda_{N-1}^h}{\lambda_1^h} = \text{cond}_2(A_h) = O(h^{-2}), \quad (16.116)$$

akkor létezik egy olyan  $\tau = \tau^* = \tau^*(h)$  érték, amelyre

$$\rho_1 = e^{-\lambda_1 \tau}.$$

Erre a  $\tau^*$  értékre teljesül  $\tau^* \in \Omega_{1/2}$  és  $\tau^* = O(h)$ .

[Kommentár: Mivel arra számíthatunk, hogy  $\lambda_1^h = O(1)$ , a (16.116) feltételek azt jelentik, hogy  $|\lambda_1 - \lambda_1^h| = O(h^2)$ , v.ö. a II. 11.6.4-ben (251) előtt álló becsléssel is.]

**Bizonyítás.** Ha  $\sigma = 1/2$ , akkor

$$\rho_1 = \left(1 - \frac{\tau}{2}\lambda_1^h\right) / \left(1 + \frac{\tau}{2}\lambda_1^h\right).$$

Emiatt tekintsük a

$$g(\tau) = \frac{1}{\tau} \ln \frac{1 + \frac{\tau}{2}\lambda_1^h}{1 - \frac{\tau}{2}\lambda_1^h}$$

függvényt. A logaritmus jól ismert sorfejtéséből azt kapjuk, hogy

$$g(\tau) = \frac{\lambda_1^h}{2x} \ln \frac{1+x}{1-x} = \lambda_1^h \sum_{n=0}^{\infty} \frac{x^{2n}}{2n+1}, \quad \text{ha } x := \frac{\tau}{2}\lambda_1^h < 1, \quad (16.117)$$

vagyis ha  $\tau < \frac{2}{\lambda_1^h}$ . (Ezeknek a  $\tau$ -értékeknek csak egy része van  $\Omega_{1/2}$ -ben.) A sorfejtés azt mutatja meg, hogy

$$g(0) = \lambda_1^h, \quad g(\tau) \text{ monoton növekvő, } g\left(\frac{2}{\lambda_1^h}\right) = \infty.$$

Eszerint  $g$  minden,  $\lambda_1^h$ -nál nagyobb értéket pontosan egyszer veszi fel. Így azt kell belátnunk, hogy  $O(h) = \tau^* \in \Omega_{1/2}$  arra a  $\tau^*$  értékre, amelyen  $g(\tau) = \lambda_1$ . Ehhez vizsgáljuk a

$$\bar{\tau} := \left(12 \frac{\lambda_1 - \lambda_1^h}{(\lambda_1^h)^3}\right)^{1/2} \in \left(0, \frac{2}{(\lambda_1^h \lambda_{N-1}^h)^{1/2}}\right] \subseteq \left(0, \frac{2}{\lambda_1^h}\right]$$

értéket – amely (16.115) szerint  $\Omega_{1/2}$  eleme. A (16.117) sorfejtést alkalmazva kapjuk

$$\begin{aligned} g(\bar{\tau}) &= \lambda_1^h \left(1 + \left(\frac{\lambda_1^h}{2}\right)^2 \frac{1}{3} \left(12 \frac{\lambda_1 - \lambda_1^h}{(\lambda_1^h)^3}\right) + \sum_{n=2}^{\infty} \frac{\left(\frac{\bar{\tau}}{2}\lambda_1^h\right)^{2n}}{2n+1}\right) \\ &= \lambda_1 + \lambda_1^h \sum_{n=2}^{\infty} \frac{\left(\frac{\bar{\tau}}{2}\lambda_1^h\right)^{2n}}{2n+1} > \lambda_1 = g(\tau^*). \end{aligned} \quad (16.118)$$

Ezért  $\tau^* < \bar{\tau} \in \Omega_{1/2}$ , tehát  $\tau^* \in \Omega_{1/2}$ . Ezenkívül  $\bar{\tau} = O(h)$ , ugyanis a (16.116) feltételek szerint

$$\frac{\bar{\tau}}{2}\lambda_1^h = \left(3 \frac{\lambda_1 - \lambda_1^h}{\lambda_1^h}\right)^{1/2} \leq \left(\frac{\lambda_1^h}{\lambda_{N-1}^h}\right)^{1/2} = (\text{cond}_2(A_h))^{-1/2} = O(h). \quad (16.119)$$

Tovább láthatjuk be, hogy  $\tau^* = O(h)$  (és nem esetleg  $\tau^* = O(h^2)$ ), amihez abból indulhatunk ki, hogy  $g' > 0$  és

$$g(\bar{\tau}) - g(\tau^*) = (\bar{\tau} - \tau^*)g'(\xi), \quad \xi \in (\tau^*, \bar{\tau}), \quad \text{azaz } \bar{\tau} - \tau^* = \frac{g(\bar{\tau}) - g(\tau^*)}{g'(\xi)}.$$

Ha most  $h \leq h_0$  olyan kicsi, hogy  $O(h^{-2}) = \text{cond}_2(A_h) \geq 2$ , akkor (16.118), (16.119) és hasonló megfontolások, mint fent adják azt, hogy innen következik  $\sqrt{\frac{5}{8}}\bar{\tau} \leq \tau^*$ . Ez  $\tau^* < \bar{\tau}$ -tal együtt már azt mutatja, hogy  $\tau^* = O(h)$ . Végül megemlítjük, hogy  $\bar{\tau} - \tau^* = O(h^3)$ .  $\square$ .

**Megjegyzések. 1.** Ha teljesülnek a tétel feltételei, akkor  $\tau = \tau^*$ -gal valójában igaz  $|\rho_k| < |\rho_1|$ ,  $k \neq 1$ , valamint

$$\|y^j\|_{(0,h)} \leq e^{-\lambda_1 t_j} \|y^0\|_{(0,h)}.$$

**2.** A feltételek érvényesek a (16.111)-(16.112) esetben. Ekkor  $h_0 = 1/3$ ,

$$\tau^* < \tau_{1/2} = h^2 / \sin \pi h \leq \frac{h}{\pi}(1 + (2\pi - 4)h^2) \approx \frac{h}{\pi}(1 + 2.283h^2),$$

és numerikus kísérletek szerint

$$\tau^* \approx \frac{h}{\pi}(1 + 0.329h^2).$$

**3.** A végeelem séma és  $\sigma = 1/2$  esetén nincs olyan  $\tau$ , hogy  $\rho_1 = e^{-\lambda_1 \tau}$ , ugyanis ekkor  $g(\tau) \geq \bar{\lambda}_1^h > \lambda_1$ . Az aszimptotikus stabilitás viszont biztosítható. Ugyancsak nincs olyan  $\tau$  a tisztán implicit differenciaséma esetén, mert

$$\rho_1 = \frac{1}{1 + \tau \lambda_1} > \frac{1}{e^{\tau \lambda_1}} = e^{-\tau \lambda_1}$$

minden  $\tau > 0$ -ra. Hasonlóan  $\sigma = 0$  esetén  $\rho_1 = 1 - \tau \lambda_1 < e^{-\tau \lambda_1}$ .

**4.** Érdekes összevetni az aszimptotikus stabilitást biztosító feltételeket a maximumelv feltételeivel. Míg korábban, (ld. 16.12. lemma – 16.14. tétel) az lehetett a kissé elnagyolt összefoglalás, hogy a hővezetési egyenlet jó diszkrét modellje csak a  $\sigma = 1$ -hez tartozó végeelem vagy differenciaséma (míg  $\sigma = 1/2$  esetén már  $\tau = h^2$  kellett az ott vizsgált, az alkalmazások szempontjából fontos tulajdonságok érvényességéhez), addig most derül ki, hogy a  $\sigma = 1/2$  a legmegfelelőbb, kombinálva azzal, hogy  $\tau = O(h)$ . Viszont a  $\sigma = 1$  séma pontosságának szinten tartásához  $\tau = O(h^2)$  szükséges!

Mivel az aszimptotikus stabilitás fontos nagy időintervallumra vonatkozó parabolikus feladat megoldásakor, a következő stratégiát javasolhatjuk: eleinte a tisztán implicit sémát használjuk, míg a kezdeti függvény esetleges

szakadásai (ehhez ld. a 26. feladatot), a kezdeti- és peremértékek közti ugrások ki nem simultak. Ezután a Crank–Nicolson eljárásra kapcsoljunk át. Ehhez ld. a 16.12 lemma előtti megjegyzést is.

Vannak olyan számítási tapasztalatok is, amelyek szerint célszerű a  $\sigma = 0.55$  érték állandó használata, ill. (ld. Axelsson és Gololobov cikkét) hogy  $\sigma > 1/2$  választása közel  $1/2$ -hez előnyös a stabilitás és pontosság szempontjából még nemlineáris esetben is.  $\square$

A 16.15. tételhez ld. a 27. és 28. feladatot is.

### 16.4.10 A súlyozott differenciaséma stabilitása a nem-szimmetrikus esetben

Ebben a pontban vizsgáljuk a (16.45) séma stabilitását abban az  $L_2$ -féle módszerek alkalmazására – nehezebb esetben, amikor az  $A_h$  mátrix nem feltétlenül szimmetrikus. (Az M-mátrix technikával itt is lehet és kisebb nehézségekkel stabilitási eredményeket bebizonyítani, ld. pl. II. 11.4.10.) Azokra az alkalmazásokban fontos esetekre korlátozzuk a vizsgálatot, hogy  $A_h$  valóban pozitív definit (ld. I. 47. o.) :

$$(A_h y_h, y_h)_{(0,h)} > 0, \text{ ha } y_h \neq 0.$$

Az ilyen mátrix reguláris. Két példa valóban pozitív definit, differenciálegyenletek diszkrétizációjával kapcsolatos mátrixra (homogén peremfeltételek mellett, az egydimenziós  $\omega_h$  rácson és  $a > 0$  és  $b$  konstansokkal, ld. II. 11.4.10.) az az  $A_h$ , amelyre

$$(A_h y_h)_i = -(a y_{\bar{x}x} + b y_x)_i, \quad (16.120)$$

ill. az *upwind-diszkrétizációval* meghatározott

$$(A_h y_h)_i = -(a y_{\bar{x}x} + \frac{b + |b|}{2} y_x + \frac{b - |b|}{2} y_{\bar{x}})_i \quad (16.121)$$

mátrix. Az első esetben a II-beli 11.3. lemmából, valamint az ott a 205/206. oldalon található (135) azonosságból következik ugyanis

$$(A_h y_h, y_h)_{(0,h)} = a |y_h|_{(1,h)}^2.$$

A második esetben viszont az ottani (133), (134) összefüggések megmutatják, hogy

$$(A_h y_h, y_h)_{(0,h)} = (a + \frac{|b|h}{2}) |y_h|_{(1,h)}^2.$$

Itt a kezdetiértékek szerinti stabilitás bizonyítása kézenfekvő. A jobboldal szerinti stabilitás viszont a 16.6. tételből következik.

**16.16. Tétel** (stabilitás valósan pozitív definit mátrix esetén). A (16.45) sémában legyen  $B_h = I_h + \sigma\tau A_h$ , és  $A_h$  legyen valósan pozitív definit.

Ekkor (16.45) stabil a kezdetiértékek és a jobboldal szerint, ha  $\sigma \geq \frac{1}{2}$ , és ebben az esetben érvényes a következő becslés :

$$\|y_h^j\|_{(0,h)} \leq \|y_h^0\|_{(0,h)} + \sum_{k=0}^{j-1} \tau \|\varphi_h^k\|_{(0,h)}.$$

**Bizonyítás.** Szorozzuk meg (16.45)-et  $A_h^{-1}$ -gyel :

$$\tilde{B}_h y_t + y_h^j = \tilde{\varphi}_h^j, \quad j = 0, 1, \dots, m-1, \quad (16.122)$$

$$\tilde{B}_h := A_h^{-1} + \sigma\tau I_h, \quad \tilde{\varphi}_h^j := A_h^{-1} \varphi_h^j. \quad (16.123)$$

Legyen  $y_h = v_h + w_h$ , ahol  $v_h$  a (16.122) azon megoldása, amely  $\tilde{\varphi}_h \equiv 0$  esetén adódik, míg  $w_h$  az  $y_h^0 = 0$  esetnek felel meg. Ekkor  $v_h$  számára (16.122)-ből (úgy mint korábban a 16.5. tétel előtt) a (16.51) azonosság megfelelőjét vezetjük le :

$$2\tau \left( \left( \tilde{B}_h - \frac{\tau}{2} I_h \right) v_t, v_t \right)_{(0,h)} + \|v_h^{j+1}\|_{(0,h)}^2 - \|v_h^j\|_{(0,h)}^2 = 0.$$

Ha  $\sigma \geq \frac{1}{2}$ , akkor itt

$$\left( \left( \tilde{B}_h - \frac{\tau}{2} I_h \right) v_t, v_t \right)_{(0,h)} = \left( (A_h^{-1} + (\sigma - \frac{1}{2})\tau I_h) v_t, v_t \right)_{(0,h)} \geq 0,$$

mert  $A_h$ -val együtt  $A_h^{-1}$  is valósan pozitív definit. Így

$$\|v_h^{j+1}\|_{(0,h)} \leq \|v_h^j\|_{(0,h)} \leq \dots \leq \|v_h^0\|_{(0,h)} = \|y_h^0\|_{(0,h)}. \quad (16.124)$$

$w_h$ -re a 16.6. tétel adja a

$$\|w_h^j\|_{(0,h)} \leq \sum_{k=0}^{j-1} \tau \|(I_h + \sigma\tau A_h)^{-1} \varphi_h^k\|_{(0,h)} \quad (16.125)$$

becslést, ahol figyelembe vettük, hogy (16.123) szerint

$$\tilde{B}_h^{-1} \tilde{\varphi}_h^k = (A_h^{-1} + \sigma\tau I_h)^{-1} \tilde{\varphi}_h^k = (I_h + \sigma\tau A_h)^{-1} \varphi_h^k.$$

Az  $\|(I_h + \sigma\tau A_h)^{-1}\varphi_h^k\|_{(0,h)}$  norma (16.125)-ben viszont felülről becsülhető  $\|\varphi_h^k\|_{(0,h)}$ -val, mert az  $(I_h + \sigma\tau A_h)\psi_h = \varphi_h$  egyenletből következik az  $A_h$  valóban pozitív definitisége alapján (és  $\sigma\tau \geq 0$  mellett)

$$\|\psi_h\|_{(0,h)}^2 \leq ((I_h + \sigma\tau A_h)\psi_h, \psi_h)_{(0,h)} = (\varphi_h, \psi_h)_{(0,h)} \leq \|\varphi_h\|_{(0,h)}\|\psi_h\|_{(0,h)}.$$

Így tehát (16.125)-ből kapjuk a

$$\|w_h^j\|_{(0,h)} \leq \sum_{k=0}^{j-1} \tau \|\varphi_h^k\|_{(0,h)}$$

becslést, amely (16.124)-gyel együtt az állítást adja.  $\square$

A (16.120) ill. (16.121) mátrixokkal felírt (16.96) séma stabilitása (jobb-oldal, kezdeti- és peremértékek szerint) a maximum normában is bebizonyítható (ld. a 29. feladatot), és ekkor a  $\sigma < \frac{1}{2}$  esetben is jutunk eredményre.

### 16.4.11 További súlyozott differenciasémák

Az alábbiakban olyan *háromréteges súlyozott sémát* írunk le, amely a Crank–Nicolson sémával fel tudja venni a versenyt, ami a megvalósítás egyszerűségét és másodrendű pontosságát illeti, de felülmúlja azt a tekintetben, hogy erősen A-stabil változatot tartalmaz (mint közönséges differenciálegyenleteket megoldó módszer) és ekkor kevésbé hajlamos oszcillációkra. A Glowinski és Periaux által más összefüggésben (többdimenziós áramlási problémák megoldására) javasolt séma két paraméterrel ( $\alpha$ -val és  $\theta$ -val) dolgozik. Bevezetve a  $\beta := 1 - \alpha$  és  $\theta' = 1 - 2\theta$  paramétereket is, a séma a következő három részlépésből áll, ha az  $M\dot{u} + A(t)u = f(t)$  egyenletrendszerre alkalmazzuk és  $y^{j+\kappa}$  az  $u((j + \kappa)\tau)$  vektor approximációja :

$$\begin{aligned} (M + \alpha\theta\tau A^{j+\theta})y^{j+\theta} &= (M - \beta\theta\tau A^j)y^j + \theta\tau f^j, \\ (M + \beta\theta'\tau A^{j+1-\theta})y^{j+1-\theta} &= (M - \alpha\theta'\tau A^{j+\theta})y^{j+\theta} + \theta'\tau f^{j+1-\theta}, \\ (M + \alpha\theta\tau A^{j+1})y^{j+1} &= (M - \beta\theta\tau A^{j+1-\theta})y^{j+1-\theta} + \theta\tau f^{j+1}. \end{aligned} \quad (16.126)$$

A séma az  $\alpha = \theta = \frac{1}{2}$  esetben meglehetősen hasonló a Crank–Nicolson sémának három egymást követő,  $\tau/2$  hosszúságú lépésére. De (16.126) nemcsak másodrendű, hanem erősen A-stabil is (amely tulajdonság a Crank–Nicolson sémának nincs), ha

$$\theta = 1 - \sqrt{2}/2, \quad \alpha \in (\frac{1}{2}, 1].$$

A megvalósítás szempontjából viszont különösen érdekes ezen paraméter választás kombinációja az

$$\alpha = (1 - 2\theta)/(1 - \theta) = 2 - \sqrt{2}$$

feltétellel, mert ekkor konstans  $A$  mátrix esetén a (16.126) egyenletek bal oldalán mindig ugyanaz a mátrix áll, a  $j \rightarrow j + 1$  átmenet költsége így megegyezik a Crank–Nicolson séma három lépésének költségével. Ezért, ha most (16.126)-ban  $\tau$ -t  $3\tau$ -val helyettesítünk, akkor a három részlépés után ugyanott vagyunk, mint a Crank–Nicolson séma három  $\tau$  hosszúságú lépése után, de jobb megoldás birtokában.

A háromréteges sémával kapcsolatban ld. a 30. és 31. feladatot is! Az utóbbi feladatban a (16.126)-ra hasonló, kétréteges sémákat konstruálunk, amelyeket többdimenziós parabolikus feladatok megoldására is lehet javasolni. A fent említett háromréteges séma viszont az áramlási feladatokra azért valamivel alkalmasabb, mert az  $e^{i\omega(x-at)}$  alakú függvényeket – a *transport-egyenlet* pontos megoldásait – kevésbé csillapítja, mint a 31. feladatban konstruálandó kétréteges sémák. Ehhez a témához ld. \*\*\* 18.2.4. és 18.10.1. A kétréteges sémák akkor alig csillapítják az említett függvényeket, ha igen közel vagyunk a Crank–Nicolson sémához, v.ö. \*\*\* (179)-cel 18.10.1-ben.

## 16.5 Változó együtthatójú egyenletek

Itt a

$$\frac{\partial u}{\partial \tau} = \frac{\partial}{\partial x} \left( k(x, t) \frac{\partial u}{\partial x} \right) + f(x, t), \quad (16.127)$$

parabolikus egyenlettel foglalkozunk, ill. ennek többdimenziós változatával.

Az ilyen egyenletek a következő alkalmazások miatt érdekesek.

1) kőolaj vagy víz mozgása a földben ill. a talajban (ekkor (16.127)-et *szivárgási egyenletnek* hívjuk). Ekkor tipikusan  $k = k(u)$ , viszont helyben nagyon kiterjedt feladatoknál  $k = k(x, u)$ . Ebben az esetben a talaj helybeli változását vehetjük figyelembe – ha megfelelő adatok állnak rendelkezésre.

2) A hővezetési egyenlet :  $k$  a hővezetési együttható, amely gyakran szakaszonként konstans, ha a vizsgált test különböző anyagokból áll. Ha nagyobb hőmérsékleti intervallumot kell figyelembe venni, akkor  $k$  megint az  $u$ -tól is függ, erről ld. részletesebben a 16.6. pont elejét.

3) Levegőben ill. vízben lezajló turbulens folyamatok modellezéséhez  $x$ - és  $y$ -irányban vesznek konstans együtthatót,  $z$ -irányban (függőlegesen) viszont elég bonyolult függvényt (hőmérséklettől és az áramlási sebességtől függőt). Ide tartozik Richardson törvénye is, amely szerint a turbulencia megfelelő diffúziós együtthatóval durván modellezhető. Ekkor  $t^{2/3}$ -mal növekszik az együttható. (Egy eleinte lokálisan koncentrált szennyeződés a levegőben vagy a tengerben először csak kisskálájú turbulens erők hatása révén terül szét, ami



lassan történik. Szétterülésével viszont a – nagyobb energiájú – nagyskálájú turbulens erők hatása alá kerül és ekkor gyorsabban terül szét.)

A hővezetési feladatok esetén természetes a  $k \geq c_0 > 0$  feltevés. Ekkor – a korábbiakhoz képest – kevés változással alkalmazható az energia-beclsések technikája (mint 16.4.3-ban, felhasználva II. 11.4.8-at is).

### 16.5.1 Egydimenziós hővezetési egyenlet helytől függő együtthatóval

Legyen (16.6)-ban most

$$\mathcal{L}u := \frac{\partial}{\partial x} \left( k(x) \frac{\partial u}{\partial x} \right), \quad 0 < k_0 \leq k(x) \leq K_0, \quad 0 \leq x \leq 1, \quad (16.128)$$

mint (16.127)-ben. Az egyetlen  $x$  helykoordináta annak felel meg, hogy a háromdimenziós tér  $y, z$ -koordinátáiban nincs változás (tehát síkszimmetriát tételezünk fel, v.ö. 16.3-mal is). A (16.128) differenciáloperátorral tekintjük a (16.6)-(16.9) feladatot.

A II. 11.4.8-beli jelölésekkel írjuk a szokásos  $\omega_h$  ekvidisztáns rácson

$$\begin{aligned} (\Lambda(k)y)_i &:= \frac{1}{h} ((ky_x)_{i+1/2} - (ky_x)_{i-1/2}) \\ &= (ky_x)_{\tilde{x},i}, \quad i = 1, \dots, N-1, \end{aligned} \quad (16.129)$$

és ugyanonnan tudjuk (ld. a 11.11. tétel bizonyítását), hogy a (16.6)-(16.9), (16.128) vegyes feladat  $u$  megoldásának megfelelő simasága esetén (ha  $k \in C^3[0, 1]$ ,  $u \in C^{4,0}([0, 1] \times [0, T])$ ) érvényes

$$(\Lambda(k)u)_i = (ku')'(x_i) + O(h^2).$$

A stabilitási elmélet alkalmazásához legyen (16.8)-ban, (16.9)-ben  $g_0 = g_1 \equiv 0$ , a peremen eltűnő rácsfüggvények  $L_2(\omega_h)$  terét használjuk, és ennek skalárszorzata a (16.44) a 16.4.3. pontban. Ekkor a  $\Lambda$  differencia-operátorral definiáljuk az  $A_h = A_h(k)$  mátrixot :

$$(A_h y_h)(x) = -(\Lambda(k)y)(x), \quad x \in \omega_h, \quad (16.130)$$

amihez (16.129)-ben legyen  $y_0 = y_N = 0$ .

**16.17. Lemma** (változó együtthatójú differencia-approximáció tulajdonságai). Az  $L_2(\omega_h)$  térben  $A_h$  önadjungált, és ha igaz (16.128), akkor pozitív definit is.

**Bizonyítás.** A II-beli 11.3. lemma mintájára írhatjuk (mivel  $y_0 = y_N = 0 = v_0 = v_N$ )

$$\begin{aligned} (A_h y_h, v_h)_{(0,h)} &= - \sum_{i=1}^{N-1} [(k y_x)_{i+1/2} - (k y_x)_{i-1/2}] v_i \\ &= \sum_{i=1}^N [(k y_x)_{i-1/2} v_i - (k y_x)_{i-1/2} v_{i-1}] - (k y_x)_{N-1/2} v_N + (k y_x)_{1/2} v_0 \\ &= \sum_{i=1}^N ((k y_x) v_x)_{i-1/2} h =: (k y_x, v_x). \end{aligned}$$

Ezért tehát

$$(A_h y_h, v_h)_{(0,h)} = (y_h, A_h v_h)_{(0,h)},$$

és

$$(A_h y_h, y_h)_{(0,h)} = \sum_{i=1}^N (k(y_x)^2)_{i-1/2} h \geq k_0 \sum_{i=1}^N (y_x)_{i-1/2}^2 h =: k_0 |y_h|_{(1,h)}^2.$$

Alkalmazva erre a II. 11.4.6-beli 11.5. lemmát, az következik, hogy

$$(A_h y_h, y_h)_{(0,h)} \geq 6k_0 \|y_h\|_{(0,h)}^2. \quad \square$$

**Megjegyzések. 1.**  $A_h(k)$  normájára azt kapjuk, hogy

$$(A_h y_h, y_h)_{(0,h)} \leq K_0 \sum_{i=1}^N (y_x)_{i-1/2}^2 h \leq \frac{2K_0}{h^2} \sum_{i=1}^N (y_i^2 + y_{i-1}^2) h \leq \frac{4K_0}{h^2} \|y_h\|_{(0,h)}^2,$$

tehát

$$\|A_h(k)\|_{(0,h)} \leq \frac{4K_0}{h^2}.$$

**2.** A kétdimenziós esetet ld. 15.5.2-ben : a 15.20 tételt és az 1. megjegyzést hozzá.  $\square$

**16.18. Következmény** (változó együtthatójú, súlyozott differenciaséma stabilitása). A homogén peremfeltételek mellett felírt

$$(I_h + \sigma \tau A_h) y_t + A_h y_h = \varphi_h, \quad \varphi_i^j = f_i^{j+\sigma}, \quad (16.131)$$

súlyozott differenciaséma stabil a kezdetiértékek szerint, ha

$$\sigma \geq \frac{1}{2} - \frac{h^2}{4K_0\tau},$$

és stabil a jobboldal szerint, ha

$$\sigma \geq \frac{1}{2} - \frac{1-\varepsilon}{4K_0\tau}h^2, \quad 0 < \varepsilon \leq 1.$$

Ez utóbbi esetben érvényes a következő becslés :

$$2k_0\|y_h^j\|_{C(\omega_h)} \leq k_0\|y_h^j\|_{(1,h)} \leq K_0\|y_h^0\|_{(1,h)} + \left( \frac{1}{2\varepsilon} \sum_{j=0}^{k-1} \tau \|\varphi_h^j\|_{(0,h)}^2 \right)^{1/2}. \quad (16.132)$$

**Bizonyítás.** Az eredmény következik a 16.5., 16.7. tételekből, valamint a (16.61) becslésből. Ezután a II-beli 11.5. lemmára hivatkozunk, hogy a (16.132) becslés előálljon.  $\square$

**Megjegyzések. 1.** Az explicit séma stabilitását érzékenyen érintheti az, ha  $K_0$  nagy, mert ekkor

$$\tau \leq \frac{h^2}{2K_0}$$

biztosítja a stabilitást.

**2.** A maximumelv akkor igaz, ha

$$\frac{1}{\tau} - \frac{1-\sigma}{h^2}(k_{i+1/2} + k_{i-1/2}) \geq 0, \quad i = 1, \dots, N-1, \quad (16.133)$$

vagyis ha

$$\tau \leq \frac{h^2}{2(1-\sigma)K_0}.$$

Ehhez ld. a 32. feladatot.  $\square$

A (16.131) séma egyenleteinek megoldására alkalmazhatjuk a rövidített Gauss-eliminációt (ld. I. 1.3.9).

### 16.5.2 Időtől is függő együttható esete

Tekintsük most azt az esetet, hogy (16.127)-ben  $k = k(x, t)$  és Lipschitz-folytonos  $t$  szerint : van olyan  $L$  konstans, amellyel tetszőleges  $x$ -re teljesül

$$|k(x, t) - k(x, s)| \leq L|t - s|. \quad (16.134)$$

Legyen továbbra is

$$0 < k_0 \leq k(x, t) \leq K_0 \text{ minden } (x, t) \in \overline{Q}_T\text{-re.} \quad (16.135)$$

Erre az általános esetre a stabilitási elmélet bizonyos változtatással alkalmazható.

A differencia-approximáció most függ az időtől :

$$\begin{aligned} (\Lambda y)_i &= (\Lambda^{(j+1/2)}(k)y)_i \\ &= \frac{1}{h} \left[ (k^{(j+1/2)}y_x)_{i+1/2} - (k^{(j+1/2)}y_x)_{i-1/2} \right]. \\ &= (k^{(j+1/2)}y_x)_{\tilde{x},i}. \end{aligned} \quad (16.136)$$

Ehhez a differencia-operátorhoz (16.130) segítségével vezetjük be az  $A_h = A_h^{(j+1/2)}(k)$  mátrixot, feltéve, hogy (16.136) érvényes  $y_0 = y_N = 0$ . Ekkor

$$\begin{aligned} \left| \left( (A_h^{(j+1/2)} - A_h^{(j-1/2)}) y_h, y_h \right)_{(0,h)} \right| &= \left| \sum_{i=1}^N ((k^{(j+1/2)} - k^{(j-1/2)})(y_x)^2)_{i-1/2} h \right| \\ &\leq L\tau \sum_{i=1}^N (y_x)_{i-1/2}^2 h \\ &\leq \frac{L}{k_0} \tau \sum_{i=1}^N (k^{(j-1/2)}(y_x)^2)_{i-1/2} h. \end{aligned}$$

Tehát, ha teljesül (16.134), akkor következik  $A_h$ -ra

$$\left| \left( (A_h^{(j+1/2)} - A_h^{(j-1/2)}) y_h, y_h \right)_{(0,h)} \right| \leq c_2 \tau \left( A_h^{(j-1/2)} y_h, y_h \right)_{(0,h)}, \quad (16.137)$$

ahol  $c_2 \leq L/k_0$ .

**16.19. Tétel** (Szamarszkij; időtől függő kétréteges differenciaséma stabilitása). A súlyozott differenciaséma következő változata :

$$B_h^{(j+1/2)} y_t + A_h^{(j+1/2)} y = \varphi_h^j, \quad j = 0, 1, \dots, m-1, \quad (16.138)$$

stabil a kezdetiértékek és a jobboldal szerint, amennyiben  $A_h^{(j+1/2)}$  szimmetrikus és pozitív definit minden  $j$ -re, érvényes (16.137) és ( $\varepsilon = \text{const} > 0$ -val)

$$B_h^{j+1/2} \geq \frac{\tau}{2} A_h^{j+1/2} + \varepsilon I_h. \quad (16.139)$$

**Bizonyítás.** A 16.5. tétel előtti módon (16.138)-at átalakítva eljutunk a következő azonossághoz :

$$\begin{aligned} 2\tau \left( \left( B_h - \frac{\tau}{2} A_h \right)^{(j+1/2)} y_t, y_t \right)_{(0,h)} &+ \left( A_h^{(j+1/2)} y_h^{j+1}, y_h^{j+1} \right)_{(0,h)} \\ &= \left( A_h^{(j+1/2)} y_h^j, y_h^j \right)_{(0,h)} + 2\tau (\varphi_h, y_t)_{(0,h)}. \end{aligned}$$

A (16.137) becslés segítségével következik

$$\begin{aligned} & 2\tau \left( \left( B_h - \frac{\tau}{2} A_h \right)^{(j+1/2)} y_t, y_t \right)_{(0,h)} + \left( A_h^{(j+1/2)} y_h^{j+1}, y_h^{j+1} \right)_{(0,h)} \\ & \leq (1 + c_2\tau) \left( A_h^{(j-1/2)} y_h^j, y_h^j \right)_{(0,h)} + 2\tau (\varphi_h, y_t)_{(0,h)}. \end{aligned}$$

Itt alkalmazzuk a szokásos  $\varepsilon$ -egyenlőtlenséget :

$$(\varphi_h, y_t)_{(0,h)} \leq \frac{1}{4\varepsilon} \|\varphi_h\|_{(0,h)}^2 + \varepsilon \|y_t\|_{(0,h)}^2,$$

tehát, az  $\left( A_h^{(j-1/2)} y_h^j, y_h^j \right)_{(0,h)} =: \|y_h^j\|_A^2$  egyszerűsítő jelölést bevezetve,

$$\begin{aligned} & 2\tau \left( \left( B_h - \varepsilon I_h - \frac{\tau}{2} A_h \right)^{j+1/2} y_t, y_t \right)_{(0,h)} + \|y_h^{j+1}\|_A^2 \\ & \leq (1 + c_2\tau) \|y_h^j\|_A^2 + \frac{\tau}{2\varepsilon} \|\varphi_h\|_{(0,h)}^2, \end{aligned}$$

és ekkor a (16.139) feltétel azt adja, hogy

$$\|y_h^{j+1}\|_A^2 \leq (1 + c_2\tau) \|y_h^j\|_A^2 + \frac{\tau}{2\varepsilon} \|\varphi_h^j\|_{(0,h)}^2,$$

vagyis

$$\|y_h^{j+1}\|_A^2 \leq (1 + c_2\tau)^{j+1} \left[ \|y_h^0\|_A + \frac{\tau}{2\varepsilon} \sum_{k=0}^j \|\varphi_h^k\|^2 \right].$$

Ez a stabilitást jelenti, hiszen  $t_{j+1} \leq T$  és

$$(1 + c_2\tau)^{j+1} \leq \exp(c_2 t_{j+1}) \leq e^{c_2 T}. \quad \square$$

**Megjegyzés.** Amikor  $A_h$  konkrétan (16.136)-tal definiált, akkor alkalmazható a tétel, ha a (16.134), (16.135) feltételek teljesülnek, és ekkor a séma stabilitása következik a  $c_2 = L/k_0$  konstanssal.  $\square$

### 16.5.3 A hővezetési egyenlet hengerszimmetriában

Tekintsünk most az olyan esetet, amikor (16.128)-ban  $k_0 > 0$  nem érvényes.

Vizsgáljuk azt a feladatot, hogy  $R = 1$  sugarú hengerben olyan hővezetési folyamat zajlik le, amely sugárszimmetrikus,  $f(r, t)$  a hőforrások sűrűsége (amely a szimmetria miatt az  $r$  páros függvénye), és a hőmérsékleti vezetőképesség  $k = \lambda/c_p\rho = 1$ . A folyamat legyen független  $z$ -től is (a henger

hosszúsági koordinátájától). Ekkor a következő matematikai modell alkalmazható :

$$\begin{aligned} t > 0, \quad 0 < r < R : \quad \frac{\partial u}{\partial t} &= \mathcal{L}u + f(r, t); & (16.140) \\ \mathcal{L}u &:= \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) = \frac{1}{r} (ru')'; \\ t = 0, \quad 0 \leq r \leq R : \quad u(r, 0) &= u_0(r); \\ t > 0, \quad r = 0 : \quad \frac{\partial u}{\partial r} &= 0; \quad r = R : \quad u = g(t). \end{aligned}$$

A  $\frac{\partial u}{\partial r} = 0$  feltétel következik a sugárszimmetriából, és az  $r = 0$ -nál szinguláris megoldásokat zárja ki. Az  $\mathcal{L}u = 0$  egyenlet két független megoldása  $u^{(1)} = 1$  és  $u^{(2)} := \ln \frac{1}{r}$ . Emiatt a  $\frac{\partial u}{\partial r} = 0$  feltétel azzal is helyettesíthető, hogy legyen  $u(0)$  véges.

Ezen feltétel még egy másik megjelenési formája az, hogy ha az  $u(r, t)$  megoldást kicsi  $r$ -re Taylor-sor alakjában keressük, akkor abban csak a páros hatványok fordulnak elő :

$$u(r, t) = u(0, t) + \frac{r^2}{2} \frac{\partial^2 u}{\partial r^2}(0, t) + \frac{r^4}{24} \frac{\partial^4 u}{\partial r^4}(0, t) + \dots \quad (16.141)$$

A (16.140) feladat approximációját a következő, az  $r = 0$ -beli szingularitás miatt  $\rho h$ -val eltolt ekvidisztáns rácson adjuk meg :

$$\bar{\omega}_h^{(\rho)} := \{r_i = (\rho + i)h, \quad i = 0, 1, \dots, N; \quad r_N = R = 1\}. \quad (16.142)$$

Itt a  $\rho > 0$ -t később alkalmasan választjuk ki. Ezen a rácson a  $\mathcal{L}u$  kifejezést a következőképpen approximáljuk, amikor  $i \geq 1$  :

$$(\Lambda^{(1)}y)_i := \frac{1}{r_i h} ((ry_r)_{i+1/2} - (ry_r)_{i-1/2}) = \left( \frac{1}{r} (ry_r)_{\bar{r}} \right)_i. \quad (16.143)$$

Ebben a képletben az  $r_{i+1/2}$  és  $(ry_r)_{i+1/2}$  definíciója a szokásos :

$$\begin{aligned} r_{i+1/2} &:= r_i + \frac{h}{2} = \left( \rho + i + \frac{1}{2} \right) h, \\ (ry_r)_{i+1/2} &:= r_{i+1/2} \frac{y_{i+1} - y_i}{h}, \quad i = 0, \dots, N-1. \end{aligned}$$

A következő képleteket ill. sorfejtéseket használjuk  $r = r_i$ -nél, ha  $u \in C^4[0, 1]$  :

$$\begin{aligned} r_{i\pm 1/2} &= r_i \pm \frac{h}{2}, \\ (u_r)_{i\pm 1/2} &= \left( u' \pm \frac{h}{2} u'' + \frac{h^2}{6} u''' \right)_i \pm \frac{h^3}{24} u^{(4)}(r_i + \vartheta_{\pm} h), \end{aligned}$$

tehát

$$\begin{aligned} (ru_r)_{\bar{r},i} &= (ru'' )_i + h^2 r [u^{(4)}(r_i + \vartheta_+ h) - u^{(4)}(r_i + \vartheta_- h)] \\ &\quad + \left( u' + \frac{1}{6} u''' \right)_i + O(h^3), \\ (\Lambda^{(1)}u)_i &= (\mathcal{L}u)_i + O\left(\frac{h^2}{r}\right), \quad i \geq 1, \end{aligned} \quad (16.144)$$

hiszen  $\mathcal{L}u = u'' + \frac{1}{r}u'$ . Az approximáció ezek szerint kis  $r$ -re csak elsőrendű, mivel pl.  $r_1 = (1 + \rho)h$  és így  $h^2/r_1 = h/(1 + \rho)$ .

A homogén peremfeltételekkel rendelkező rácsfüggvények esetére (16.130) mintájára definiáljuk  $\Lambda^{(1)}$  alapján az  $A_h^{(1)}$  mátrixot és vezetjük be a következő skalárszorzatot :

$$(y_h, v_h)_{(r,0)} := \sum_{i=0}^{N-1} r_i y_i v_i h. \quad (16.145)$$

Ezzel egyrészt az  $(\Lambda^{(1)}y)_0$  képletét fogjuk meghatározni, másrészt a stabilitási vizsgálatot készítjük elő.

Amennyiben  $v_N = 0$ , akkor

$$\begin{aligned} (A_h^{(1)}y_h, v_h)_{(r,0)} &= - \sum_{i=1}^{N-1} [(ry_r)_{i+1/2} - (ry_r)_{i-1/2}] v_i + (A_h^{(1)}y_h)_0 r_0 v_0 h \\ &= \sum_{i=0}^{N-1} (ry_r v_r)_{i+1/2} h + v_0 [(ry_r)_{1/2} + (A_h^{(1)}y_h)_0 r_0 h]. \end{aligned}$$

Ha most

$$-(\Lambda^{(1)}y)_0 = (A_h^{(1)}y_h)_0 := -\frac{1}{r_0 h} (ry_r)_{1/2} \quad (16.146)$$

és  $y_N = 0$ , akkor

$$(A_h^{(1)}y_h, v_h)_{(r,0)} = \sum_{i=0}^{N-1} (ry_r v_r)_{i+1/2} h =: (y_h, v_h)_{(r,1)} = (y_h, A_h^{(1)}v_h)_{(r,0)}, \quad (16.147)$$

vagyis  $A_h^{(1)}$  önadjungált. Ezután a  $\rho$ -t abból határozzuk meg, hogy  $(\Lambda^{(1)}u)_0$  legyen alkalmas mint  $\mathcal{L}u$  approximációja  $r = r_0 = \rho h$ -nál.

A szimmetria miatt, ld. (16.141)-et, érvényes

$$\begin{aligned} u(r) &= u_0 + \frac{r^2}{2} u_0'' + \int_0^r \frac{1}{6} (r-s)^3 u^{(4)}(s) ds, \\ \left( u'' + \frac{1}{r} u' \right) (r) &= (\mathcal{L}u)(r) = 2u_0'' + \int_0^r \left[ \frac{(r-s)^2}{2r} + r-s \right] u^{(4)}(s) ds \\ &= 2u_0'' + O(r^2). \end{aligned}$$

Továbbá,

$$u(r_1) - u(r_0) = \frac{(1+\rho)^2 - \rho^2}{2} h^2 u_0'' + O(h^4), \quad (ru_r)_{1/2} = \left(\rho + \frac{1}{2}\right)^2 h^2 u_0'' + O(h^4),$$

és így (16.146) elfogadható mint approximáció,

$$(\Lambda^{(1)}u)_0 = -(A_h^{(1)}\vec{u}_h)_0 = \frac{1}{r_0 h} (ru_r)_{1/2} = \frac{(\rho + \frac{1}{2})^2}{2\rho} \mathcal{L}u + O(h^2), \quad (16.148)$$

ha  $(\rho + \frac{1}{2})^2 / (2\rho) = 1$ . Legyen tehát  $\rho = 1/2$ . Ezzel az  $A_h^{(1)}$  mátrix és a  $\Lambda^{(1)}$  operátor teljesen meg vannak határozva.

Ezután a (16.140) egyenlet súlyozott differencia-approximációja az  $y$  rácsfüggvénnyel

$$y_t = \Lambda^{(1)}y^{(\sigma)} + f^{j+\sigma}, \quad r \in \omega_h^{(1/2)}; \quad y = g, \quad r = R = 1,$$

és standard alakban, a  $g$  inhomogén peremfeltétel eliminációja után, az  $y_h = (y_0, \dots, y_{N-1})^T$  megoldási vektorral,

$$B_h y_t + A_h^{(1)} y_h = \varphi_h, \quad B_h := I_h + \sigma \tau A_h^{(1)}, \quad (16.149)$$

$$\varphi_i^j = f_i^{j+\sigma}, \quad i = 0, 1, \dots, N-2, \quad \varphi_{N-1}^j = f_{N-1}^{j+\sigma} + \frac{g^{(\sigma)}}{h^2} \left( \frac{r_{N-1/2}}{r_{N-1}} \right).$$

Ezen séma képlethibája (16.144) és (16.148) alapján, ha  $u \in C^{4,3}(\overline{Q}_T)$  :

$$\psi_i^j = O\left(\tau|\sigma - \frac{1}{2}| + \tau^2 + \frac{h^2}{r_i}\right). \quad (16.150)$$

Mivel az  $A_h^{(1)}$  és  $B_h$  mátrixokról már tudjuk, hogy önadjungáltak az  $r$ -rel súlyozott (16.145) skalárszorzatban és az  $r = r_N$  peremen eltűnő rácsfüggvényeken, azért a 16.5., 16.7. stabilitási tételek alkalmazhatóságához még annak belátása hiányzik, hogy az  $A_h^{(1)}$  mátrix pozitív definit. Ennek bizonyítására szükségünk van egy, a (16.145) és (16.147) skalárszorzatokhoz tartozó  $\|y_h\|_{(r,0)}$  és  $|y_h|_{(r,1)}$  normákat összekötő beágyazási becslésre.

**16.20. Lemma** (speciális beágyazási tételek). Ha  $y$  tetszőleges olyan rácsfüggvény, amely a (16.142),  $\rho = 1/2$  rácson definiált és annak  $r = r_N$  peremén eltűnik, akkor igazak a következő becslések :

$$\|y_h\|_{(r,0)} \leq \frac{1}{2} |y_h|_{(r,1)}, \quad \|y_h\|_{C(\omega_h^{(\rho)})} \leq \left(\ln \frac{2}{h}\right)^{1/2} |y_h|_{(r,1)}. \quad (16.151)$$



**Bizonyítás.** Abból indulunk ki, hogy  $y_N = 0$  esetén

$$y_i = - \sum_{j=i}^{N-1} (y_r)_{j+1/2} h, \quad i = 0, \dots, N-1,$$

és a Cauchy–egyenlőtlenség segítségével kapjuk, hogy

$$y_i^2 \leq \sum_{j=i}^{N-1} (r(y_r)^2)_{j+1/2} h \sum_{j=i}^{N-1} \frac{h}{r_{j+1/2}} \leq |y_h|_{(r,1)}^2 \sum_{j=i}^{N-1} \frac{h}{r_{j+1/2}}. \quad (16.152)$$

Innen következik, hogy

$$\|y_h\|_{(r,0)}^2 = \sum_{i=0}^{N-1} r_i y_i^2 h \leq |y_h|_{(r,1)}^2 S_N,$$

ahol

$$S_N := \sum_{i=0}^{N-1} \sum_{j=i}^{N-1} \frac{r_i h^2}{r_{j+1/2}} = \sum_{i=0}^{N-1} \sum_{j=i}^{N-1} \frac{(i + \frac{1}{2})h^2}{j+1}. \quad (16.153)$$

Ugyanis  $\rho = 1/2$  miatt  $r_i = (i + 1/2)h$  és  $r_{j+1/2} = (j + 1)h$ . Átrendezéssel azt kapjuk (16.153)-ból, hogy

$$\begin{aligned} S_N &:= \sum_{j=0}^{N-1} \frac{h}{j+1} \sum_{i=0}^j (i + \frac{1}{2})h = \sum_{j=0}^{N-1} \frac{h}{j+1} \frac{(j+1)^2 h}{2} \\ &= \frac{h^2}{4} N(N+1) = \frac{1}{4} (1 - \frac{h^2}{4}) < \frac{1}{4}, \end{aligned}$$

mert  $r_N = (N + 1/2)h = 1$ . Ezzel (16.151) első becslése megvan. Térjünk most vissza (16.152)-höz! Az ottani összeg a következőképpen becsülhető :

$$\sum_{j=i}^{N-1} \frac{h}{r_{j+1/2}} = \sum_{j=i}^{N-1} \frac{1}{j+1} \leq \sum_{j=0}^{N-1} \frac{1}{j+1} < \ln(2N+1) = \ln \frac{2}{h},$$

mivel  $1/h = N + \frac{1}{2}$ , és a lemma be van bizonyítva.  $\square$

**Megjegyzés.** A lemma szerint

$$(A_h^{(1)} y_h, y_h)_{(r,0)} = |y_h|_{(r,1)}^2 \geq 4 \|y_h\|_{(r,0)}^2,$$

tehát  $A_h^{(1)}$  valóban pozitív definit – és  $\sigma \geq 0$  esetén vele együtt  $B_h$  is.  $\square$

**16.21. Tétel** (a súlyozott differenciaséma konvergenciája hengersizmetria esetén). Amennyiben  $u \in C^{4,3}$  a (16.140) feladat megoldása, és ha a

$$\sigma \geq \frac{1}{2} - \frac{h^2(1-\varepsilon)}{4\tau} \quad (16.154)$$

feltétel teljesül, akkor a (16.143), (16.146), (16.149) séma konvergens. Az  $y_h$  megoldására igazak a

$$\|(y_h - \vec{u}_h)^j\|_{(r,1)} = O\left(\tau\left|\sigma - \frac{1}{2}\right| + \tau^2 + h^2\left(2 + \ln\frac{2}{h}\right)^{1/2}\right), \quad (16.155)$$

$$\|(y_h - \vec{u}_h)^j\|_{C(\omega_h^{(\rho)})} = \left(\ln\frac{2}{h}\right)^{1/2} O\left(\tau\left|\sigma - \frac{1}{2}\right| + \tau^2 + h^2\left(2 + \ln\frac{2}{h}\right)^{1/2}\right) \quad (16.156)$$

becslések minden véges  $[0, T]$  intervallumon.

[Kommentár. Emlékeztetünk arra, hogy (16.154)-ben  $\varepsilon = 1$ , ha  $\sigma \geq 1/2$ .]

**Bizonyítás.** A 16.5., 16.7. tételek alkalmazhatók, és azok szerint a kezdetiérték és jobboldal szerinti stabilitás a  $\|\cdot\|_{(r,1)}$  normában következik

$$\sigma \geq \frac{1}{2} - \frac{1-\varepsilon}{4\tau\|A_h^{(1)}\|_{(r,0)}}$$

mellett. Itt  $\|A_h^{(1)}\|_{(r,0)} \leq 4/h^2$ , ld. a 33. feladatot.

A  $z_h := y_h - \vec{u}_h$  hibára igaz az

$$|(y_h - \vec{u}_h)^j|_{(r,1)} \leq \left(\frac{1}{2\varepsilon} \sum_{k=0}^{j-1} \tau \|\psi_h^k\|_{(r,0)}^2\right)^{1/2} \quad (16.157)$$

egyenlőtlenség, ahol  $\psi_h$  a (16.150) szerinti képlethiba. A konvergencia becslés befejezéséhez számítsuk ki a  $\|\psi_h^k\|_{(r,0)}$  normát! A képlethiba  $(\psi_h^{(1)})^k = O\left(\tau\left|\sigma - \frac{1}{2}\right| + \tau^2\right)$  része problémamentesen adja azt, hogy

$$\left(\frac{1}{2\varepsilon} \sum_{k=0}^{j-1} \tau \|(\psi_h^{(1)})^k\|_{(r,0)}^2\right)^{1/2} \leq \left(\frac{T}{2\varepsilon}\right)^{1/2} O\left(\tau\left|\sigma - \frac{1}{2}\right| + \tau^2\right). \quad (16.158)$$

A  $(\psi_h^{(2)})^k$  rész jellemzője  $|(\psi_i^{(2)})^k| \leq Mh^2/r_i$ , így :

$$\begin{aligned} \|(\psi_h^{(2)})^k\|_{(r,0)}^2 &= \sum_{i=0}^{N-1} r_i \left((\psi_i^{(2)})^k\right)^2 h \leq M^2 h^4 \sum_{i=0}^{N-1} \frac{h}{r_i} \\ &\leq M^2 h^4 \left(2 + \ln\frac{2}{h}\right), \end{aligned} \quad (16.159)$$

hiszen

$$\sum_{i=0}^{N-1} \frac{h}{r_i} = \sum_{i=0}^{N-1} \frac{1}{\frac{1}{2} + i} \leq 2 + \int_{1/2}^{N-1/2} \frac{dx}{x} = 2 + \ln(2N-1) < 2 + \ln \frac{2}{h}.$$

(16.157)-(16.159) adja a (16.155) becslést, és a 16.20. lemma segítségével következik (16.156).  $\square$

A (16.156) becslés durva : a  $\ln(2/h)$  tagok feleslegesek benne.

Ezen állítás bizonyításához vizsgáljuk a (16.149) sémát a maximum normában.

**16.22. Tétel** (hengerszimmetrikus eset : konvergencia a maximum normában). Teljesüljenek a 16.21. tétel feltételei. Amennyiben  $\sigma = 1$  vagy  $\tau \leq \frac{h^2}{2(1-\sigma)}$ , akkor a (16.149) séma  $z_h := \vec{u}_h - y_h$  hibájára teljesül a

$$\|z_h\|_{C(\omega_h^{(\rho)})} \leq M \left( \tau \left| \sigma - \frac{1}{2} \right| + \tau^2 + h^2 \right)$$

egyenlőtlenség.

**Bizonyítás.** Az egyszerűség kedvéért a következőképpen becsüljük felülről a képlethibát :

$$|\psi_i^j| \leq M \bar{h}^2 / r_i, \quad \bar{h}^2 := \tau \left| \sigma - \frac{1}{2} \right| + \tau^2 + h^2.$$

Legyen most  $g_i^j := M \bar{h}^2 (1 - r_i)$ . Ekkor

$$\begin{aligned} (A_h^{(1)} g_h^j)_i &= -\frac{1}{r_i h} \left( r_{i+1/2} \frac{g_{i+1} - g_i}{h} - r_{i-1/2} \frac{g_i - g_{i-1}}{h} \right)^k \\ &= -\frac{M \bar{h}^2}{r_i h} (-r_{i+1/2} + r_{i-1/2}) = \frac{M \bar{h}^2}{r_i}, \end{aligned}$$

és ez  $i = 0$  esetén is érvényes ( $r_{-1/2} := 0, r_{1/2} = h$ ). Így

$$(g_t + A_h^{(1)} g_h^{(\sigma)})_i^j = \frac{M \bar{h}^2}{r_i} \geq |\psi_i^j|. \quad (16.160)$$

Az M-mátrixelmélet alkalmazásához ellenőrizzük még a  $(B_h y_t + A_h^{(1)} y_h)_i^j$  kifejezés összes együtthatójának előjelét! Az  $y_i^{j+1}$  együtthatója pozitív, az  $y_{i\pm 1}^{j+1}$  és  $y_{i\pm 1}^j$  együtthatói nempozitívak, de  $y_i^j$  együtthatója

$$-\frac{1}{\tau} + \frac{1-\sigma}{r_i h^2} (r_{i+1/2} + r_{i-1/2}) = -\frac{1}{\tau} + \frac{2}{h^2} (1-\sigma) \leq 0,$$

ha  $\sigma = 1$  vagy  $\tau \leq \frac{h^2}{2(1-\sigma)}$ , és a reláció még  $i = 0$ -ra is igaz ( $r_{-1/2} := 0$ ,  $r_{1/2} = h = 2r_0$ ).

Az előjelek ekkor tehát megfelelőek, a (16.160) feladat  $g_h^j$  megoldási vektora alkalmas mint majoráns vektor, és felülről becsüli a séma hibáját.  $\square$

Ezzel a  $\sigma = 1$  vagy  $\tau \leq \frac{h^2}{2(1-\sigma)}$  esetben bebizonyítottuk, hogy (16.156)-ban nincs szükség az  $\ln(2/h)$  tagokra. Az  $\mathcal{L}$  operátor (16.143), (16.146) approximációjával kapcsolatban ld. a 34. feladatot is. A (16.149) séma egyenleteit megint a rövidített Gauss-eliminációval oldhatjuk meg.

#### 16.5.4 A hővezetési egyenlet gömbszimmetriában

A gömbszimmetrikus esetben érvényes a következő egyenlet :

$$t > 0, \quad 0 < r < R = 1 : \quad \frac{\partial u}{\partial t} = \mathcal{L}u + f(r, t); \quad (16.161)$$

$$\mathcal{L}u := \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial u}{\partial r} \right) = \frac{1}{r^2} (r^2 u')'.$$

Ehhez társulnak a kezdeti és peremfeltételek :

$$t = 0, \quad 0 \leq r \leq R : \quad u(r, 0) = u_0(r); \quad (16.162)$$

$$t > 0, \quad r = 0 : \quad \frac{\partial u}{\partial r} = 0; \quad r = R : \quad u = g(t). \quad (16.163)$$

A szimmetriának megfelelően az  $f(r, t)$  az  $r$  páros függvénye.

A  $[0, 1]$  intervallumot a (16.142) ráccsal osztjuk fel,  $\rho = 1$ -et választva. Ekkor legyen

$$\begin{aligned} (\Lambda^{(2)}y)_i &:= \frac{1}{hr_i^2} ((\bar{r}^2 y_r)_{i+1/2} - (\bar{r}^2 y_r)_{i-1/2}) \\ &= \left( \frac{1}{r^2} (\bar{r}^2 y_r)_{\bar{r}} \right)_i, \quad 1 \leq i \leq N-1, \end{aligned} \quad (16.164)$$

ahol most  $\bar{r}_{i+1/2}^2 = r_i r_{i+1}$ ,  $i = 0, \dots, N-1$ .

Mielőtt az  $i = 0$  esettel foglalkoznánk, mutassuk meg, hogy a (16.164) definícióval

$$(\Lambda^{(2)}u)_i = (\mathcal{L}u)_i + \begin{cases} O(h^2/r_i), & \text{ha } u \in C^4[0, 1], \\ \left( \frac{h^2}{12} \mathcal{L}^2 u \right)_i + O(h^4/r_i), & \text{ha } u \in C^6[0, 1]. \end{cases} \quad (16.165)$$

Ugyanis az  $(u_r)_{i+1/2}$  kifejezést  $r = r_i$  körül Taylor-sorba fejtve kapjuk

$$(u_r)_{i+1/2} = \left( u' \pm \frac{h}{2}u'' + \frac{h^2}{6}u''' \pm \frac{h^3}{24}u^{(4)} + \frac{h^4}{120}u^{(5)} + O(h^5) \right)_i,$$

és  $\bar{r}_{i\pm 1/2}^2 = (r^2 \pm hr)_i$ , tehát

$$\begin{aligned} (\bar{r}^2 u_r)_{i\pm 1/2} &= \left( r^2 u' \pm h \left[ r u' + \frac{r^2}{2} u'' \right] + h^2 \left[ \frac{r}{2} u'' + \frac{r^2}{6} u''' \right] \right. \\ &\quad \left. \pm h^3 \left[ \frac{r}{6} u''' + \frac{r^2}{24} u^{(4)} \right] + h^4 \left[ \frac{r}{24} u^{(4)} + \frac{r^2}{120} u^{(5)} \right] \right)_i + O(r_i h^5). \end{aligned} \quad (16.166)$$

Innen következik

$$(\bar{r}^2 u_r)_{\bar{r},i} = \left( r^2 u'' + 2r u' + h^2 \frac{r^2}{12} \left[ u^{(4)} + \frac{4}{r} u''' \right] \right)_i + O(r_i h^4). \quad (16.167)$$

Ezenkívül

$$\mathcal{L}u = u'' + \frac{2}{r}u', \quad (16.168)$$

$$\mathcal{L}^2 u = \frac{1}{r^2} \left( r^2 \left( u'' + \frac{2}{r}u' \right)' \right)' = u^{(4)} + \frac{4}{r}u'''. \quad (16.169)$$

A (16.167)-(16.169) relációkból következik (16.165) – amely azt jelenti, hogy könnyebb magasabbrendű eljárást megszerkeszteni, mint a hengerszimmetrikus esetben. (Ott olyan sorfejtés, mint a (16.165) érvényes akkor, ha az  $\bar{r}_{i+1/2}$  definíciójához bonyolultabb, logaritmust tartalmazó képletet alkalmazunk. Ehhez ld. Frjazinov 1971-es cikkét. A (16.165) analógiája a síkszimmetria esetével szembetűnő – ld. a szokásos másodrendű differenciahányados II. 11.4.1-beli (16.34) sorfejtését.)

Az  $i = 0$  esetben használhatjuk a  $\Lambda^{(2)}$  változatlan (16.164) definícióját akkor, ha formálisan  $\bar{r}_{-1/2}^2 = 0$ -nak vesszük :

$$(\Lambda^{(2)} y_h)_0 := \frac{1}{hr_0^2} (\bar{r}^2 y_r)_{1/2}. \quad (16.170)$$

Az approximáció oldaláról ez a következők miatt is elfogadható. Abból indulhatunk ki, hogy a szimmetriának köszönhetően

$$u = a_0 + a_1 r^2 + a_2 r^4 + a_3 r^6 + \dots,$$

így

$$\begin{aligned} (\mathcal{L}u)(r) &= 6a_1 + 20a_2 r^2 + 42a_3 r^4 + \dots \\ (\mathcal{L}^2 u)(r) &= 120a_2 + 840a_3 r^2 + \dots, \end{aligned}$$

és az  $r = r_0 = h$  helyen érvényes

$$\left( \mathcal{L}u + \frac{h^2}{12} \mathcal{L}^2 u \right) (r) = 6a_1 + 30a_2 r^2 + 112a_3 r^4 + \dots$$

Ugyanakkor a (16.166) sorfejtés alapján ( $i = 0$ )

$$(\Lambda^{(2)}u)_0 = \frac{1}{hr_0^2} (\bar{r}^2 u_r)_{1/2} = 6a_1 + 30a_2 h^2 + 118a_3 h^4 + \dots,$$

vagyis

$$(\Lambda^{(2)}u)_0 = \left( \mathcal{L}u + \frac{h^2}{12} \mathcal{L}^2 u \right)_0 + 6a_3 h^4 + O(h^6),$$

tehát (16.165) igaz  $i = 0$ -ra is.

Az  $r = R = 1$ -nél adott elsőfajú peremfeltételt eliminálva és (16.130)-cal bevezetve a  $\Lambda^{(2)}$ -nek megfelelő  $A_h^{(2)}$  mátrixot, a (16.161) egyenlet approximációját a súlyozott differenciasémával írhatjuk a következőképpen, amikor  $y_h = (y_0, \dots, y_{N-1})^T$ :

$$\begin{aligned} B_h y_t + A_h^{(2)} y_h &= \varphi_h, & B_h &:= I_h + \sigma \tau A_h^{(2)}, & (16.171) \\ \varphi_i^j &= f_i^{j+\sigma}, \quad i = 0, 1, \dots, N-2, & \varphi_{N-1}^j &= f_{N-1}^{j+\sigma} + \frac{g^{(\sigma)}}{h^2} \left( \frac{\bar{r}_{N-1/2}}{r_{N-1}} \right)^2. \end{aligned}$$

Ha  $u \in C^{4,3}(\bar{Q}_T)$ , akkor ezen séma képlethibájára – figyelembe véve az  $u_t$  és  $u^{(\sigma)}$  sorfejtéseit is (ld. (16.34)) kapjuk (16.165)-ből azt, hogy

$$\psi_i = O \left( \tau \left| \sigma - \frac{1}{2} \right| + \tau^2 + \frac{h^2}{r_i} \right). \quad (16.172)$$

Ezután a további vizsgálat analóg a hengerszimmetrikus esethez. Először a következő skalárszorzatokat vezetjük be:

$$(y_h, v_h)_{(r^2,0)} := \sum_{i=0}^{N-1} r_i^2 y_i v_i h, \quad (y_h, v_h)_{(r^2,1)} := \sum_{i=0}^{N-1} (\bar{r}^2 y_r v_r)_{i+1/2} h. \quad (16.173)$$

Ehhez tartoznak az  $\|y_h\|_{(r^2,0)}^2 := (y_h, y_h)_{(r^2,0)}$  és  $|y_h|_{(r^2,1)}^2 := (y_h, y_h)_{(r^2,1)}$  normák. Következően azt használjuk, hogy teljesül

$$(A_h^{(2)} y_h, v_h)_{(r^2,0)} = (y_h, v_h)_{(r^2,1)} = (y_h, A_h^{(2)} v_h)_{(r^2,0)} \quad (16.174)$$

(ennek bizonyítása a 35. feladat tárgya), és

$$\|y_h\|_{(r^2,0)}^2 \leq \frac{1}{6} |y_h|_{(r^2,1)}^2, \quad |y_h|_{(r^2,1)}^2 \leq \frac{4}{h^2} \|y_h\|_{(r^2,0)}^2. \quad (16.175)$$

Az előbbi eredmények és a 16.5., 16.7. tételek most azt adják, hogy a (16.171) séma stabil az  $|\cdot|_{(r^2,1)}$  normában a kezdetiértékek és a jobboldal szerint, ha

$$\sigma \geq \frac{1}{2} - \frac{1 - \varepsilon h^2}{4\tau}.$$

Ekkor érvényes a következő becslés :

$$|y_h^j|_{(r^2,1)} \leq |y_h^0|_{(r^2,1)} + \left( \frac{1}{2\varepsilon} \sum_{k=0}^{j-1} \tau \|\varphi_h^k\|_{(r^2,0)}^2 \right)^{1/2}. \quad (16.176)$$

A hibabecsléshez megjegyezzük, hogy a  $z_h := \vec{u}_h - y_h$  hiba eleget tesz a

$$B_h z_t + A_h^{(2)} z_h = \psi_h$$

egyenletnek, nulla kezdeti- és peremértékek mellett. A konvergencia tisztázásához az maradt vissza a (16.176) becslés alapján, hogy számítsuk ki a  $\|\psi_h^k\|_{(r^2,0)}^2$  normát. A képlethiba (16.172) alakja miatt elég az  $1/r$  rácsfüggvény normáját meghatározni :

$$\|1/r\|_{(r^2,0)}^2 = \sum_{i=0}^{N-1} h = Nh = 1 - h < 1.$$

Így  $\|\psi_h^k\|_{(r^2,0)}^2 = O\left(\tau \left|\sigma - \frac{1}{2}\right| + \tau^2 + h^2\right)$  és

$$|z^j|_{(r^2,1)} \leq \left(\frac{T}{2\varepsilon}\right)^{1/2} O\left(\tau \left|\sigma - \frac{1}{2}\right| + \tau^2 + h^2\right). \quad (16.177)$$

Emlékeztetünk arra, hogy itt  $\varepsilon = 1$ -nek vehető, ha  $\sigma \geq 1/2$ . Ehhez ld. a 36. és 37. feladatokat is.

Tárgyalásunk végén foglalkozunk azzal a feladattal, hogy (16.161) és (16.162)-(16.163) mellett  $r = R = 1$ -nél adott az elsőfajú peremfeltétel helyett a harmadfajú :

$$\frac{\partial u}{\partial r} + \sigma u = g, \quad r = R, \quad t > 0. \quad (16.178)$$

$\sigma$  a hőátadási tényező, ld. II. 11.2. (16.178) annak felel meg, hogy a gömb felületén Newton-féle (turbulens) hőcsere folyik a külső közeggel (pl. a föld és az atmoszféra között).

A korábbi (16.164)-(16.170) approximációk most is megfelelőek. Ezután elegendő az új peremfeltétel miatt adódó különbségeket felsorolni. Először is,  $r = 1$ -nél a differenciálegyenlet

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial r^2} + 2 \frac{\partial u}{\partial r} + f(1, t).$$

Emiatt most

$$\begin{aligned}
(u_r)_{N-1/2} &= \left( \frac{\partial u}{\partial r} - \frac{h}{2} \frac{\partial^2 u}{\partial r^2} \right)_N + O(h^2) \\
&= \frac{\partial u}{\partial r} - \frac{h}{2} \left( \frac{\partial u}{\partial t} - 2 \frac{\partial u}{\partial r} - f(1, t) \right)_N + O(h^2) \\
&= (1+h) \frac{\partial u}{\partial r} - \frac{h}{2} \left( \frac{\partial u}{\partial t} - f(1, t) \right)_N + O(h^2) \\
&= (1+h)(g - \sigma u) - \frac{h}{2} \left( \frac{\partial u}{\partial t} - f(1, t) \right)_N + O(h^2).
\end{aligned}$$

Innen látjuk, hogy megfelelőek a következő definíciók :

$$(A_h^{(2)} y_h)_N := -\frac{2}{h} [(y_r)_{N-1/2} + (1+h)\sigma y_N], \quad \varphi_N := \frac{2}{h}(1+h)g^{j+\sigma} + f(1, t^{j+\sigma}),$$

amelyekkel a (16.171) sémát  $i = N$ -re egészítjük ki, és  $y_h = (y_0, \dots, y_N)^T$ . A skalárszorzatok a következőkre változnak :

$$(y_h, v_h)_{(r^2, 0)} := \sum_{i=0}^{N-1} r_i^2 y_i v_i h + \frac{h}{2}(1-h)y_N v_N, \quad (16.179)$$

$$(y_h, v_h)_{(r^2, 1)} := \sum_{i=0}^{N-1} (\bar{r}^2 y_r v_r)_{i+1/2} h + \sigma(1-h^2)y_N v_N. \quad (16.180)$$

Az  $(\cdot, \cdot)_{(r^2, 0)}$  skalárszorzatban az  $A_h^{(2)}$  mátrix önadjungált és pozitív definit, mert érvényesek a

$$\frac{18\sigma}{3\sigma+8} \|y_h\|_{(r^2, 0)}^2 \leq |y_h|_{(r^2, 1)}^2 \leq \left( \frac{5}{h^2} + \frac{8\sigma}{3h} \right) \|y_h\|_{(r^2, 0)}^2$$

becslések. Innen következik a kezdetiértékek és jobboldal szerinti stabilitás, ha

$$\sigma \geq \frac{1}{2} - \frac{h^2}{\tau(5 + \frac{8}{3}\sigma h)}.$$

A képlethibát változatlanul (16.172) adja meg, a konvergencia bizonyításához elegendő az  $1/r$  rácsfüggvény  $\|\cdot\|_{(r^2, 0)}$  normáját kiszámítani, és mivel  $\|1/r\|_{(r^2, 0)} < 1$ , így (16.177) most is érvényes. Ennek következménye a maximum normabeli konvergencia, mert

$$\|v_h\|_{C(\bar{\omega}_h^{(1)})} \leq h^{-1/2} \left( 1 + \frac{9}{8\sigma} \right)^{1/2} |v_h|_{(r^2, 1)}.$$

Ezen rész kidolgozása a 38. feladat.



## 16.6 Nemlineáris parabolikus egyenletek

A hővezetési egyenlet bevezetésénél, 16.1.3-ban rámutattunk, hogy annak együtthatói tipikusan a megoldástól is függnék (ld. 16.5. is). Szivárgási feladatok esetén a

$$c \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left( k \frac{\partial u}{\partial x} \right) + f \quad (16.181)$$

egyenlet  $k$  együtthatójának alakja pl.

$$k = k_0 \frac{u^\alpha}{1 + \beta u}$$

lehet (ekkor  $u \geq 0$  a víz- vagy kőolajnyomás), vagy más olyan racionális kifejezés  $u$ -ban, amely  $u$ -val együtt eltűnik. Ilyen feladatoknál tipikus az, hogy a megoldások deriváltjai nem folytonosak olyan pontokban, ahol  $u = 0$  és  $k(0) = 0$ .

Kőolaj esetén  $f = f(u)$ ,  $u$  az egyes frakciók parciális nyomásának vektora, amelynek minden komponensére érvényes egy (16.181) típusú egyenlet,  $f$  is vektor, amely az átmenetet a frakciók között leírja. A  $k$  együttható egyaránt a kőolajtól függ (viszkozitásától), de a kőzettől is (porozitásától). Ilyenkor az időderivált előtti együttható  $c$  helyett  $\rho(u)c_p(u)$ .

Plazmafizikai számításoknál gyakori feltevés az, hogy  $c$  és  $k$  lényegében az  $u$  hatvány-függvényei, és itt is  $k(0) = 0$  :

$$c = c(u) = c_0 u^\alpha, \quad k = k(u) = k_0 u^\beta, \quad k_0, \alpha, \beta > 0. \quad (16.182)$$

Fémek felhevítésének, élelmiszerek befagyasztásának matematikai leírására olyan  $c$  és  $k$  együtthatókat határoznak meg kísérleti úton, amelyek a keresett  $u$  hőmérséklet alacsonyfokú polinomjai, és amelyek minden pozitív  $u$ -ra pozitívak. Kerámiák szárításánál a  $k$  együtthatót  $k(u) = c_0 e^{c_1 u}$  alakban is teszik fel (ekkor  $u$  a nedvességtartalom).

Vegyireakciók esetén olyan tagok lépnek fel  $f$ -ben (v.ö. I. 6.1. ponttal), mint pl. az  $u_2 u_3 k_0 e^{-E/(R u_1)}$ , ahol  $u_1$  a hőmérséklet és  $u_2, u_3$  koncentrációk.

Az ilyen egyenletek numerikus megoldásával a következőkben foglalkozunk röviden. Vagy az egyenleteket teljesen diszkretizáljuk, majd pl. a Newton-módszert használhatjuk. Vagy szemidiszkretizálást alkalmazunk, az időderiváltat meghagyva. Ekkor a kapott közönséges differenciálegyenletek rendszerét többféleképpen lehet hatékonyan megoldani, a Newton-módszerhez szükséges deriváltak kiszámításának megkerülésével is, pl. Runge-Kutta módszert vagy prediktor-korrektor eljárást alkalmazva.

Elsőnek azokat a tennivalókat írjuk le, amelyek a közönséges differenciálegyenleteket megoldó explicit módszerek használatát megelőzik.

A (16.181) egyenlet időderiváltját megszorzó  $c(u)$  együttható miatt új változóhoz megyünk át a

$$\varphi(u) := \int_0^u c(s) ds \quad (16.183)$$

transzformációval (amellyel a (16.181) egyenletet kiegészítő kezdeti- és peremfeltételeket is át kell számítani az új változóba). Ezután

$$\frac{\partial \varphi}{\partial t} = \frac{d\varphi}{du} \frac{\partial u}{\partial t} = c(u) \frac{\partial u}{\partial t}.$$

A transzformáció egyértelműen invertálható, ha  $c(u) > 0$ . Ekkor az új jobboldal

$$g(\varphi) := f(u(\varphi)), \quad \text{továbbá } \frac{\partial \varphi}{\partial x} = c(u) \frac{\partial u}{\partial x}$$

és így

$$\frac{\partial}{\partial x} \left( k(u) \frac{\partial u}{\partial x} \right) = \frac{\partial}{\partial x} \left( \frac{k(u)}{c(u)} \frac{\partial \varphi}{\partial x} \right) = \frac{\partial}{\partial x} \left( a(\varphi) \frac{\partial \varphi}{\partial x} \right),$$

ahol  $a(\varphi) := k(u(\varphi))/c(u(\varphi))$ . Amennyiben érvényes (16.182), akkor pl.

$$\varphi(u) = \frac{c_0}{1+\alpha} u^{1+\alpha}, \quad a(\varphi) = k_1 \varphi^{(\beta-\alpha)/(1+\alpha)}, \quad k_1 := \frac{k_0}{c_0} \left( \frac{1+\alpha}{c_0} \right)^{(\beta-\alpha)/(1+\alpha)}.$$

Az  $\omega_h$  ekvidisztáns rácson a kapott

$$\frac{\partial \varphi}{\partial t} = \frac{\partial}{\partial x} \left( a(\varphi) \frac{\partial \varphi}{\partial x} \right) + g(\varphi)$$

egyenlet szemidiszkretizációját a  $\varphi(u(x_i, t)) \approx y_i(t)$  és

$$\frac{dy_i}{dt} = \frac{1}{h} \left( (ay_x)_{i+1/2} - (ay_x)_{i-1/2} \right) + g(y_i(t)), \quad i = 1, \dots, N-1, \quad (16.184)$$

relációk írják le. Itt  $a_{i\pm 1/2} := \frac{1}{2}[a(y_i) + a(y_{i\pm 1})]$ . A (16.184) rendszerre pl. az explicit Runge–Kutta módszer alkalmazható. Sajnos az elsőrendű módszer

$$\tau \leq \frac{h^2}{2 \max_u a(u)} \leq \frac{h^2 \min_u c(u)}{2 \max_u k(u)}$$

stabilitási korlátjánál alig nagyobb a magasabbrendű módszereké (ld. II. 106. o.). Viszont nagy előny, hogy itt lineáris vagy nemlineáris rendszerek megoldásával nem kell foglalkoznunk.

Következőnek tekintjük a (16.181) egyenlet közvetlen megoldását a súlyozott differenciaséma segítségével, amelyet most iterációval ki kell egészítenünk ( $\ell$  az iterációs számláló) :

$$\begin{aligned} y^{j+1,0} &:= y^j, \\ \ell &:= 0, 1, \dots : \\ c^{(\ell)} &:= c(\sigma y^{j+1,\ell} + (1-\sigma)y^j), \\ k^{(\ell)} &:= k(\sigma y^{j+1,\ell} + (1-\sigma)y^j), \\ f^{(\ell)} &:= f(\sigma y^{j+1,\ell} + (1-\sigma)y^j), \\ c^{(\ell)} \frac{y^{j+1,\ell+1} - y^j}{\tau} &= (k^{(\ell)} \cdot (\sigma y^{j+1,\ell+1} + (1-\sigma)y^j))_{\bar{x}} + f^{(\ell)}. \end{aligned} \quad (16.185)$$

Mivel ezek az egyenletek  $y^{j+1,\ell+1}$ -ben lineárisak, megoldásuk a rövidített Gauss-eliminációval történhet. Az iteráció használatát csak  $\ell = 0$ -ra és  $\ell = 1$ -re javasoljuk, ekkor prediktor-korrektor eljárásnak is tekinthetjük.

Az explicit séma ( $\sigma = 0$ ) az Euler-módszer, amelyet már fent jellemeztük.

Ha simák a feladat adatai : kezdetiértékei, peremfeltételei, jobboldala, és ha  $k(u) > 0$  minden  $u$ -ra, akkor legjobb a Crank-Nicolson eljárást használni ( $\sigma = \frac{1}{2}$ ). Ekkor két iterációt végezzünk (16.185)-tel. Az így kapott megoldás pontosabb, mint a tisztán implicit sémáé. Azokban az esetekben viszont, amikor  $k(u) = 0$  előfordul valamelyik  $u \geq 0$  értékre, a Crank-Nicolson séma alkalmatlan, mert a deriváltak töréspontjainak környezetében oszcillációkat és negatív értékeket szokott produkálni.

Ilyen helyzetben célszerű a tisztán implicit sémát ( $\sigma = 1$ ) használni.

A következőkben vizsgáljuk ezen séma alkalmazását a (16.181) egyenletre, amelyet előzetesen (hasonlóan (16.183)-ra, de) úgy transzformáljuk, hogy mindkét együtthatója eltűnjék. Ehhez először új  $v$  változót vezetünk be  $u$  helyett :

$$v := \int_0^u k(s) ds.$$

Ekkor

$$\frac{\partial v}{\partial x} = \frac{dv}{du} \frac{\partial u}{\partial x} = k(u) \frac{\partial u}{\partial x}, \quad \frac{\partial^2 v}{\partial x^2} = \frac{\partial}{\partial x} \left( k(u) \frac{\partial u}{\partial x} \right),$$

és a szokásos  $k(u) \geq k_0 > 0$  feltétel mellett  $v$  az  $u$  növekvő függvénye, tehát az  $u(v)$  inverz függvény egyértelmű.

Ezután legyen

$$\varphi(v) := \int_0^{u(v)} c(s) ds, \quad \text{azaz} \quad \frac{\partial \varphi}{\partial t} = \frac{d\varphi}{du} \frac{\partial u}{\partial t} = c(u) \frac{\partial u}{\partial t}.$$

Így (16.181) helyett kapjuk

$$\frac{\partial \varphi(v)}{\partial t} = \frac{\partial^2 v}{\partial x^2} + g(v), \quad g(v) := f(u(v)). \quad (16.186)$$

A többdimenziós esetben közvetlenül ellenőrizhető, hogy a

$$c(u) \frac{\partial u}{\partial t} = \operatorname{div}(k(u) \operatorname{grad} u) + f(u)$$

egyenlettel a fenti transzformációk után (16.186) helyett a

$$\frac{\partial \varphi(v)}{\partial t} = \Delta v + g(v) \quad (16.187)$$

egyenletre jutunk.

Azt az esetet vizsgáljuk most, amikor (16.186)-ban a  $g$  forrásfüggvény azonosan nulla. Legyen  $y_i^j$  a  $v(x_i, t_j)$  közelítése. Az eredetileg adott kezdeti- és elsőfajú peremfeltételekből a transzformáció segítségével kapjuk a  $v_0$  kezdetiértéket, ill. a  $g_0(t)$  és  $g_1(t)$  peremfeltételeket. Ekkor a tisztán implicit séma :

$$\begin{aligned} \frac{1}{\tau} [\varphi(y^{j+1}) - \varphi(y^j)] &= y_{\bar{x}x}^{j+1}, \quad x \in \omega_h, \quad j = 0, 1, \dots, m-1, \quad (16.188) \\ y^0 &= v_0, \quad x \in \bar{\omega}_h, \quad y_0^{j+1} = g_0(t_{j+1}), \quad y_N^{j+1} = g_1(t_{j+1}). \end{aligned}$$

Azt tételezzük fel, hogy ezen feladat megoldása minden  $j$ -re létezik.

Az alábbiakban a Newton-módszer alkalmazását (16.188)-ra vizsgáljuk, amikor ott  $j$  rögzített. Ez azt jelenti, hogy az

$$f_h(y_h) := (f_1(y_h), \dots, f_{N-1}(y_h))^T, \quad \text{ahol } f_i(y_h) := \varphi(y_i) - \tau y_{\bar{x}x,i} - \varphi(y_i^j),$$

vektorfüggvény  $y_h = (y_1, \dots, y_{N-1})^T = y_h^{j+1}$  gyökét keressük. Itt az  $y_{\bar{x}x,i}$  kifejezésben  $i = 1$ -re és  $i = N-1$ -re természetesen az  $y_0^{j+1}$ , ill.  $y_N^{j+1}$  peremértéket figyelembe vesszük. Az iterációsátlálót  $\ell$ -l-lel jelölve, a  $0 = f_h(y_h)$  gyökét a

$$0 = f_h(y_h^{(\ell)}) + J_h(y_h^{(\ell)})(y_h^{(\ell+1)} - y_h^{(\ell)})$$

iteráció adja, amelyhez kiindulási értéként  $y_h^{(0)} := y_h^j$ -t vesszük, az előző időréteg  $y_h$  vektorát. Mivel az  $f_h$  Jacobi-mátrixa

$$J_h(y_h) = \varphi'_h(y_h) + \tau A_h, \quad \text{ahol } (A_h y_h)_i = -y_{\bar{x}x,i}$$

homogén peremfeltételeknek felel meg és ahol  $\varphi'_h(y_h)$  az  $\varphi'(y_i)$  értékekből ( $i = 1, \dots, N-1$ ) álló diagonális mátrix, az iteráció végülis felírható mint

$$\varphi_h(y_h^{(\ell)}) = \varphi_h(y_h^{(\ell)}) + \varphi'_h(y_h^{(\ell)}) (y_h^{(\ell+1)} - y_h^{(\ell)}) - \tau y_{\bar{x}x}^{(\ell+1)}, \quad \ell = 0, 1, \dots, \quad (16.189)$$

ahol  $\varphi_h(y_h)$  a  $\varphi(y_i)$  értékekből álló vektor, amikor  $y_h = y_h^j$  vagy  $y_h = y_h^{(\ell)}$ .

Konvergencia esetén érvényes, hogy  $\lim_{\ell \rightarrow \infty} y_h^{(\ell)} = y_h^{j+1}$ .

A (16.189) rendszereket a tridiagonális algoritmus segítségével tudjuk megoldani, ha minden  $u$ -ra

$$\varphi'(v(u)) = \frac{d\varphi}{dv} = \frac{d\varphi}{du} / \frac{dv}{du} = \frac{c(u)}{k(u)} \geq 0, \quad \text{mert akkor } J_h(y_h) \geq \tau A_h$$

elemenként, és így az algoritmus stabilitási feltétele (I. 71. o. (59)) teljesül. Most fordulunk a konvergencia vizsgálatára.

**16.23. Tétel** (Szamarszkij, Newton–módszer konvergenciája nemlineáris parabolikus egyenlet megoldásakor). Teljesüljön

$$\varphi'(v) \geq c_1 > 0 \text{ és } |\varphi''(v)| \leq c_2 \text{ minden } v\text{-re.} \quad (16.190)$$

Ekkor, ha a (16.188) feladat  $y_h^j$  és  $y_h^{j+1}$  megoldásai elég közeliek egymáshoz,

$$\|y_h^{j+1} - y_h^j\|_{C(\omega_h)} < \frac{2c_1}{c_2}, \quad (16.191)$$

akkor a (16.189) Newton–iteráció négyzetesen konvergens.

**Bizonyítás.** Legyen  $z_h^{(\ell+1)} := y_h^{(\ell+1)} - \hat{y}_h$  az  $\ell+1$ -edik iteráció hibavektora, ahol  $\hat{y}_h = y_h^{j+1}$  a keresett vektor. Ekkor, kivonva a

$$\varphi(y_i^j) = \varphi(\hat{y}_i) - \tau \hat{y}_{\bar{x},i}$$

összefüggést az  $x_i$  helyen vett (16.189) egyenletből :

$$0 = \varphi\left(y_i^{(\ell)}\right) + \varphi'\left(y_i^{(\ell)}\right) \left(z^{(\ell+1)} - z^{(\ell)}\right)_i - \varphi(\hat{y}_i) - \tau z_{\bar{x},i}^{(\ell+1)},$$

vagyis, Taylor–sorba fejtve a  $\varphi(\hat{y}_i)$  kifejezést az  $y_i^{(\ell)}$  helyen :

$$0 = \varphi'\left(y_i^{(\ell)}\right) \left(z^{(\ell+1)} - z^{(\ell)}\right)_i + \varphi'\left(y_i^{(\ell)}\right) z_i^{(\ell)} - \frac{1}{2} \varphi''\left(\tilde{y}_i^{(\ell)}\right) \left(z_i^{(\ell)}\right)^2 - \tau z_{\bar{x},i}^{(\ell+1)},$$

ahol  $\tilde{y}_i^{(\ell)}$  egy  $y_i^{(\ell)}$  és  $\hat{y}_i$  közötti érték, így az következik, hogy

$$\varphi'\left(y_i^{(\ell)}\right) z_i^{(\ell+1)} - \tau z_{\bar{x},i}^{(\ell+1)} = \frac{1}{2} \varphi''\left(\tilde{y}_i^{(\ell)}\right) \left(z_i^{(\ell)}\right)^2. \quad (16.192)$$

Ezen egyenlet bal oldalán áll  $(J_h(y_h^{(\ell)}) z_h^{(\ell+1)})_i$ . (16.190) alapján  $J_h(y_h^{(\ell)}) e_h \geq c_1 e_h > 0$ , ahol  $e_h = (1, \dots, 1)^T$ , így az M–mátrix elmélet alkalmazható, amelynek segítségével (16.192)-ből következik

$$\|z_h^{(\ell+1)}\|_{C(\omega_h)} \leq \frac{1}{2c_1} \max_i \left| \varphi''\left(\tilde{y}_i^{(\ell)}\right) \left(z_i^{(\ell)}\right)^2 \right| \leq \frac{c_2}{2c_1} \|z_h^{(\ell)}\|_{C(\omega_h)}^2, \quad (16.193)$$

ahol megint (16.190)-et használtuk.

Ezután ismert módon (16.193)-ból kapjuk a  $z_h^{(\ell)} \rightarrow 0$  konvergenciát (ld. I. 6.4.2.), amennyiben  $\|z_h^{(0)}\|_{C(\omega_h)} = \|y_h^{j+1} - y_h^j\|_{C(\omega_h)}$  elég kicsi, konkrétan, ha (16.191) teljesül.  $\square$

**Megjegyzések. 1.** Mivel  $z_h^{(0)} = y_h^j - y_h^{j+1} = -\tau y_t$ , így arra számíthatunk, hogy nem lesz nehéz a (16.191) feltételnek eleget tenni, csak elég kicsi  $\tau$ -t kell választani. Ha pl.  $\|y_t\| \leq M_1$  ismert (amire akkor számíthatunk, hogyha a feladat pontos megoldása elég sima és nemcsak  $y$  konvergál  $u$ -hoz, hanem  $y_t$  is  $\frac{\partial u}{\partial t}$ -hez), akkor elegendő, hogy

$$\tau < \tau_0 := \frac{2c_1}{M_1 c_2} \quad (16.194)$$

teljesüljön.

**2.** Felhívjuk a figyelmet arra, hogy az egész gondolatmenet változatlanul alkalmazható a többdimenziós esetre is, ha a (16.187) egyenletet úgy approximáljuk, hogy  $M$ -mátrixot nyerünk.  $\square$

A következőkben azt mutatjuk meg, hogy  $\tau = O(h^2)$  esetén elérhető (16.191) teljesülése – simasági feltételek nélkül is, de feltételezve, hogy a  $g_0, g_1$  peremértékek nem függenek  $t$ -től. Abból indulunk ki, hogy  $z_h^{(0)} = y_h^{(0)} - \hat{y}_h$  és

$$\begin{aligned} 0 &= -\tau y_{\bar{x},i}^{j+1} + \varphi(y_i^{j+1}) - \varphi(y_i^j) = -\tau y_{\bar{x},i}^{j+1} - d_i z_i^{(0)}, \quad (16.195) \\ d_i &= d(y_i^j, y_i^{j+1}) := \begin{cases} \varphi'(y_i^j), & y_i^{j+1} = y_i^j \\ \frac{\varphi(y_i^{j+1}) - \varphi(y_i^j)}{y_i^{j+1} - y_i^j}, & y_i^{j+1} \neq y_i^j \end{cases}. \end{aligned}$$

Így (16.190) alapján  $d_i \geq c_1 > 0$ , és mivel  $y_{\bar{x}}^{j+1} = (y^{j+1} - \mu)_{\bar{x}}$ , ahol  $\mu$  az a lineáris függvény, amely a peremfeltételeket interpolálja, azért

$$\|z_h^{(0)}\|_{C(\omega_h)} \leq \frac{4\tau}{h^2 c_1} \|y_h^{j+1} - \bar{\mu}_h\|_{C(\omega_h)}.$$

Vektorformában a (16.195) egyenlet a következőképpen is felírható, a  $\mu$  függvény bevezetése és a pozitív  $d_i$ -re való osztás után :

$$\bar{A}_h (y_h^{j+1} - \bar{\mu}_h) = y_h^j - \bar{\mu}_h, \quad \text{ahol } (\bar{A}_h v_h)_i := v_i - \frac{\tau}{d_i} v_{\bar{x},i}.$$

Mivel  $\bar{A}_h$   $M$ -mátrix és  $\bar{A}_h e_h \geq e_h$ , így

$$\|y_h^{j+1} - \bar{\mu}_h\|_{C(\omega_h)} \leq \|y_h^j - \bar{\mu}_h\|_{C(\omega_h)} \leq \dots \leq \|y_h^0 - \bar{\mu}_h\|_{C(\omega_h)}.$$

Ha most  $y_h^0 \neq \bar{\mu}_h$  és

$$\tau < \frac{h^2 c_1^2}{2c_2 \|y_h^0 - \bar{\mu}_h\|_{C(\omega_h)}}, \quad (16.196)$$

akkor (16.191) teljesül.

A (16.194) és (16.196) feltételeket inkább úgy kell felfogni, hogy van olyan elég kicsi  $\tau_*$ , hogy  $\tau < \tau_*$  esetén konvergens az iteráció.

A fentiekben a (16.189) iterációt teljes konvergenciáig vizsgáltuk. A valóságban érdemes (16.189)-cel 2 iterációt végrehajtani. Ezt ekkor megint prediktor-korrektor eljárásnak foghatjuk fel. Ha csak  $\tau = O(h^2)$  esetén lenne konvergencia, akkor nem lenne érdemes az implicit sémát használni : ekkor ezzel az erővel az explicit sémát is alkalmazhatnánk.

Azoknál a vegyiparban érdekes feladatoknál, amikor (16.181)-ben  $u$  vektor, így  $y_i^j \approx u(x_i, t_j)$  is vektor, de  $c$  és  $k$  konstansok, a (16.185) iterációt a következőképpen érdemes megváltoztatni, a Newton-módszert alkalmazva  $f(\sigma y_i^{j+1, \ell+1} + (1 - \sigma)y_i^j)$ -re ( $J_f$  legyen az  $f(u)$  vektor Jacobi-mátrixa, amelyet az  $y_i^j$  helyen számítjuk ki, aminek köszönhetően  $J_f$  nem függ  $\ell$ -től) :

$$\begin{aligned} y^{j+1,0} &:= y^j, \quad x \in \bar{\omega}_h \\ \ell &:= 0, 1, \dots : \\ c \frac{y^{j+1, \ell+1} - y^j}{\tau} &= k (\sigma y^{j+1, \ell+1} + (1 - \sigma)y^j)_{\bar{x}x} \\ &\quad + \vec{f}_h(y^j) + \sigma J_f \cdot (y^{j+1, \ell+1} - y^j), \quad x \in \omega_h. \end{aligned}$$

Ezek az egyenletek újra lineárisak  $y^{j+1, \ell+1}$ -ben és hatékonyan megoldhatók a Gauss-elimináció blokk-tridiagonális egyenletrendszerekre specializált változatával (\*\*\*) 18.10.3.). A jobboldal differencia-approximációjánál egyébként érdemes lehet a Patankar-fogás alkalmazása (ld. 16.4.7.).

Igazán problematikus lehet a (16.181)-típusú egyenletek megoldása akkor, ha a  $k$  együttható a  $\frac{\partial u}{\partial x}$  függvénye, pl. ha  $k = |\frac{\partial u}{\partial x}|^\alpha$ ,  $\alpha > 1$ . Ilyenkor az egyszerű iteráció vagy a Newton-módszer gyakran nem konvergensek, még a tisztán implicit séma esetén sem. De nem nehéz pl. a  $k = k_0/(1 + |\frac{\partial u}{\partial x}|)$  eset.

A nemlineáris egyenletek megoldása így gyakran nem egyszerű, de adódhatnak kiegészítő lehetőségek is. Mint jellemző példát említjük, hogy az olyan hővezetési feladatokban, amelyeknél a sugárzás lényeges, ott vagy az egyenlet forrástagjaként, vagy a peremfeltételekben a  $\sigma T^4$  tag szerepel, ahol  $\sigma$  a Stefan-Boltzmann konstans és  $T$  a hőmérséklet. Kétréteges sémában való használatra a  $T^4$  diszkrétizációja lehet

$$T^4(x_i, t_{j+1/2}) \approx |T_i^j|^3 T_i^{j+1}, \quad \text{vagy} \quad T^4(x_i, t_{j+1/2}) \approx (T_i^j T_i^{j+1})^2,$$

ahol az első approximáció bár csak elsőrendű, de ezzel egyben linearizálást is hajtottunk végre. A második approximáció nemlineáris, de másodrendű.

Hasonló példát nemlineáris tagok előnyös approximációjára a 17.3. pontban mutatunk. Idetartozik a következő, M. Lees-ra visszamenő, iterációmentes eljárás is, amellyel a (16.181) egyenlet kezdetiérték-peremérték feladatát kényelmesen megoldhatjuk. Ez háromréteges, és így két kezdetiértéket kell megadnunk :

$$y^0 = u_0, \quad x \in \bar{\omega}_h, \quad y^1 = y^0 + \frac{\tau}{c(y^0)} [(k(y^0)y_x)_{\bar{x}} + f(y^0)], \quad x \in \omega_h. \quad (16.197)$$

Ezután  $j = 2, \dots, m = T/\tau$ -re használjuk a következő  $y^{j+1}$ -ben lineáris – sémát :

$$c(y^j)y_t^{\circ} = (k(y^j)(y^{(\sigma)})_x)_{\bar{x}} + f(y^j), \quad x \in \omega_h, \quad (16.198)$$

ahol

$$\begin{aligned} y_t^{\circ} &:= \frac{1}{2}(y_t + y_{\bar{t}}) = \frac{1}{2\tau}(y^{j+1} - y^{j-1}), \\ y_{\bar{t}} &:= (y^j - y^{j-1})/\tau, \quad y_t := (y^{j+1} - y^j)/\tau, \\ y^{(\sigma)} &:= \sigma y^{j+1} + (1 - 2\sigma)y^j + \sigma y^{j-1}. \end{aligned} \quad (16.199)$$

Az  $y^1$  kezdetiérték approximációja a Taylor-sorfejtés alapján másodrendű, a (16.198) egyenleté szimmetriájának köszönhetően ugyancsak másodrendű. Az egyenletben  $y^j$  és  $y^{j-1}$  már ismertek,  $y^{j+1}$  meghatározandó. Ez történhet a tridiagonális algoritmussal, ha  $\sigma > 0$  és

$$c(u, x, t) \geq c_0 > 0, \quad k(u, x, t) \geq k_0 > 0 \quad \text{minden } (u, x, t)\text{-re.}$$

Ezek mellett a feltételek mellett a (lineáris egyenletre alkalmazott) (16.197)-(16.199) séma stabil, ha  $\sigma > \frac{1}{4}$ , ami a \*\*\* 18. fejezetben bizonyításra kerülő 18.6. és 18.7. tételekből következik (ld. ott a 18.8. következményt, valamint a hozzáfűzött megjegyzéseket is). Mivel  $\sigma = \frac{1}{2}$  esetén  $y^j$  csak az együtthatókban fordul elő és annak veszélye áll fenn, hogy a páros és páratlan időrétegek szétváltnak (ld. II 10.4.3-ban a középpont szabály elemzését), azért javasolható a  $\sigma = \frac{1}{3}$  érték (amelyre Lees bizonyította be a másodrendű konvergenciát a nemlineáris esetben). Ehhez a sémához ld. a 40. feladatot is.

Befejezésül ehhez a ponthoz megjegyezzük, hogy a nemlineáris egyenletek esetén már az egydimenziós esetben is érdemes lehet a többrácsos módszer használata.



## 16.7 Többdimenziós parabolikus egyenletek

Tekintsük először a (4) egyenlet  $d$ -dimenziós esetét, a tartomány legyen  $d$ -dimenziós téglá,

$$\Omega := \{0 \leq x_i \leq a_i, \quad 1 \leq i \leq d\}, \quad \text{pereme } \Gamma.$$

Legyen  $k(x, t) \geq k_0 > 0$ ,  $c\rho(x, t) \geq c_0 > 0$ ,  $(x, t) \in Q_T := \{\Omega \times (0, T]\}$ . A  $Q_T$  peremét  $S_T := \Gamma \times (0, T]$ -vel jelöljük.

Keresett olyan  $u : Q_T \rightarrow \mathbb{R}$ , amely eleget tesz a

$$c\rho \frac{\partial u}{\partial t} = \mathcal{L}u + f, \quad (x, t) \in Q_T, \quad \mathcal{L}u := \operatorname{div}(k \operatorname{grad} u), \quad (16.200)$$

egyenletnek és amely rendelkezik az

$$u(x, 0) = u_0(x), \quad x \in \bar{\Omega}; \quad u(x, t) = g(x, t), \quad (x, t) \in S_T, \quad (16.201)$$

kezdetiértékekkel, illetve elsőfajú peremfeltételekkel. A 16.200–16.201 feladatot approximálhatjuk a

$$\begin{aligned} y^0 &= u_0, \quad x \in \bar{\omega}_h, \\ c\rho y_t &= \sum_{\ell=1}^d (k y_{\bar{x}_\ell})_{\bar{x}_\ell}^{(\sigma)} + f^{j+\sigma}, \quad x \in \omega_h; \\ y &= g, \quad x \in \gamma_h = \Gamma \cap \bar{\omega}_h, \quad j = 0, 1, \dots, m-1, \end{aligned} \quad (16.202)$$

kétréteges súlyozott sémával. Az elsőfajú peremfeltételek eliminációja után ennek standard alakját kapjuk :

$$B_h y_t + A_h y_h = \varphi_h, \quad B_h = D_h + \sigma\tau A_h, \quad D_h := \operatorname{diag}_{(x,t) \in \omega_{h,\tau}} (c\rho(x, t)), \quad (16.203)$$

és az erre vonatkozó stabilitási eredményeink (16.5–16.7. tétel) attól függetlenül igazak, hogy a feladat többdimenziós.

Legyen pl. (16.200)-ban  $c\rho \equiv 1$  és  $k \equiv 1$ , tehát  $\mathcal{L}$  a Laplace-operátor, akkor (16.203)-ban az  $A_h$  és  $B_h$  mátrixok képlete

$$B_h = I_h + \sigma\tau A_h, \quad A_h = \sum_{\ell=1}^d A_{\ell h}, \quad (A_{\ell h} y_h)(x) = y_{\bar{x}_\ell x_\ell}, \quad x \in \omega_h. \quad (16.204)$$

Amennyiben most  $\Omega = (0, 1)^d$  és  $\omega_h$  négyzetes rács a  $h$  lépéstávolsággal, akkor

$$A_h \leq \frac{4d}{h^2} I, \quad \text{vagyis } I_h \geq \frac{h^2}{4d} A_h,$$

azaz a 16.5. tétel  $B_h = I_h + \sigma\tau A_h \geq \frac{\tau}{2}A_h$  feltétele teljesül, ha

$$\sigma \geq \frac{1}{2} - \frac{h^2}{4d\tau}.$$

Ennek speciális esete

$$\tau \leq \frac{h^2}{2d} \quad \text{ha} \quad \sigma = 0,$$

ami azt mutatja, hogy az explicit séma stabilitási tartománya szűkül növekvő dimenzióval.

Ha  $k \neq 1$ , akkor a fenti  $\mathcal{L}u \approx \sum_{\ell=1}^d (ku_{\bar{x}_\ell})_{\bar{x}_\ell}^{(\sigma)}$  approximációval (ld. 15.5.2. pontot, valamint a 15.6- és 15.7-beli approximációkat is) hasonló stabilitási eredményekhez jutunk. Ily módon nem a stabilitás a gond. A diszkrétizált egyenletek numerikus megoldása viszont ugyanazokat a problémákat hozza magával, mint az elliptikus egyenleteké : végülis, ha  $\sigma \neq 0$  esetén a (16.203) sémát az

$$\left(\frac{1}{\sigma\tau}D_h + A_h\right)y_h^{j+1} = \Phi_h^j := \left(\frac{1}{\sigma\tau}D_h - \frac{1-\sigma}{\sigma}A_h\right)y_h^j + \frac{1}{\sigma}\varphi_h^j \quad (16.205)$$

alakban írjuk fel, akkor látjuk, hogy éppen diszkrét elliptikus egyenletről van szó. Hasonlóra jutunk a 16.4.11. pontban leírt súlyozott sémával is. Azaz nagyméretű, ritkamátrixú egyenletrendszert kell megoldanunk. Így elsősorban a többrácsos módszert javasolhatjuk (ld. 16.7.4.).

De ezenkívül több olyan algoritmus ismeretes, amely kisebb programozási ráfordításokkal jó pontosságot ad. Ide tartoznak a *váltakozó irányok módszerei* (amelyeknek angol rövidítése ADI : alternating direction iteration).

Egy ilyen módszert a következő pontban mutatunk be.

### 16.7.1 A kétdimenziós eset : Peaceman–Rachford módszer

Tekintsük a (16.200), (16.201) feladat kétdimenziós esetét, amikor  $c\rho = 1$ . Legyen  $\Omega = \{(x_1, x_2) \in (0, a_1) \times (0, a_2)\}$ ,  $\omega_h$  a szokásos ekvidisztáns rács ( $h_\ell = a_\ell/N_\ell$  a rácsállandók,  $\ell = 1, 2$ ) és  $\gamma_h$  annak pereme.

A (16.202),  $\sigma = \frac{1}{2}$ , sémára hasonló pontosságú eljárást fogunk alkalmazni, a továbbiak számára bevezetve a

$$\Lambda_\ell y := (ky_{\bar{x}_\ell})_{\bar{x}_\ell}, \quad x \in \omega_h; \quad \gamma_h^{(\ell)} := (\gamma_h \setminus \gamma_0) \cap \{x_\ell = 0, a_\ell\}, \quad \ell = 1, 2,$$

jelöléseket.  $\gamma_h^{(\ell)}$  az  $\bar{\omega}_h$  rácsnak az  $x_\ell$  irányban szélső pontokat adja, a  $\gamma_0$  sarokpontok nélkül. Így  $\gamma_h = \gamma_0 \cup \gamma_h^{(1)} \cup \gamma_h^{(2)}$ .

A nagyobb számítási hatékonyság érdekében viszont a  $[t_j, t_{j+1}]$  időintervallumot két részre osztjuk fel, és  $t = t_{j+1/2}$ -re az  $y^{j+1/2}$  közbülső megoldást iktatjuk be :

$$\frac{y^{j+1/2} - y^j}{\tau/2} = \Lambda_1 y^{j+1/2} + \Lambda_2 y^j + f^{j+1/2}, \quad x \in \omega_h, \quad y^{j+1/2} = \mu^{j+1/2}, \quad x \in \gamma_h^{(1)}, \quad (16.206)$$

A keresett  $y^{j+1}$  megoldást kapjuk a következő második fél lépésben :

$$\frac{y^{j+1} - y^{j+1/2}}{\tau/2} = \Lambda_1 y^{j+1/2} + \Lambda_2 y^{j+1} + f^{j+1/2}, \quad x \in \omega_h, \quad y^{j+1} = g^{j+1}, \quad x \in \gamma_h^{(2)}. \quad (16.207)$$

Ezeket az egyenleteket  $j = 0, 1, \dots, m-1$ -re írjuk fel, és a (16.201)-nek megfelelő diszkrét kezdeti feltétellel egészítjük ki. A (16.206)-beli  $\mu^{j+1/2}$  peremértéket később pontosítjuk (egyelőre lehet  $\mu^{j+1/2} = g^{j+1/2}$ -re gondolni).  $y^{j+1}$  kiszámításához a  $g^{j+1}$  értékekre csak  $\gamma_h^{(2)}$ -n van szükség, de pl. a megoldás kirajzolásához a többi,  $\gamma_h$ -beli értéket is használjuk.

Ezen eljárás numerikus megvalósításához vegyük észre, hogy a (16.206) – ill. a (16.207) – fél lépésben az  $x_2$  (ill.  $x_1$ ) csak paraméter. Mindkét egyenletrendszer mátrixa tridiagonális (minden rögzített  $x_2$ , ill.  $x_1$  mellett). Ugyanis az elsőben három ( $x_1$ -irányban szomszédos)  $y^{j+1/2}$ -érték össze van kötve (míg  $x_2$  változik, mint külső paraméter). A második egyenletrendszerben mindig három  $x_2$ -irányban szomszédos  $y^{j+1}$ -érték áll kapcsolatban (és ekkor  $x_1$  a külső paraméter).

Vagyis az első, fél lépésben  $x_1$ -irányú diszkrét egyenleteket oldunk meg a tridiagonális algoritmussal :

$$t_j \rightarrow t_{j+1/2} :$$

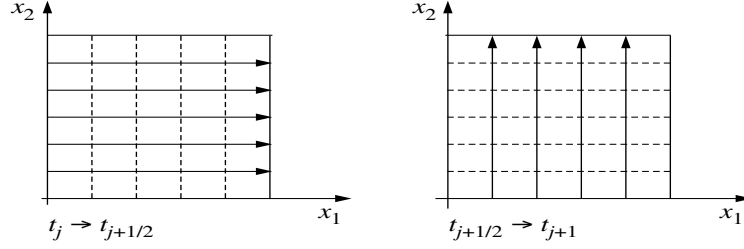
a második fél lépésben pedig  $x_2$ -irányú egyenleteket:

$$t_{j+1/2} \rightarrow t_{j+1} :$$

Innen az eljárás neve : váltakozó irányok módszere.

Fontos észrevenni, hogy nagyságrendileg annyi művelet szükséges egy időlépéshez, mint amennyi rácpontunk van, valamint azt, hogy a többdimenziós feladatot egydimenziós feladatok sorozataira sikerült visszavezetni.

A fenti (16.206)-(16.207) eljárást szerzői után *Peaceman–Rachford módszer*-nek is hívjuk. A két szakember kőolajmezők szimulációival foglalkozott (ld. a 16.6. pont elejét), amikor az eljárást parabolikus és elliptikus egyenletek megoldására javasolták.



ábra 16.5: Váltakozó irányok módszere

Azóta több hasonló módszer látott napvilágot,  $d$ -dimenziós esetre is, (ld. a 16.7.3. pontot). Végeselem approximációval kapcsolatban is foglalkoztak vele (ld. pl. M. Dryja ill. Layton és Rabier cikkét).

A váltakozó irányok módszerének előnyei : könnyen programozhatók, *párhuzamosíthatók* (mert az első féllépés  $x_1$ -irányú, csak  $x_2$ -ben különböző rendszerei mind függetlenül egymástól megoldhatók, és hasonlóan a második féllépés rendszerei is függetlenek), stabilak (ld. lent), másodrendűek, memóriatakarékosak : a megoldást rácsegyenesenként írjuk felül!

A váltakozó irányok módszerei csak eléggé speciális helyzetekben alkalmazhatók : amikor a  $k$  együtthatónak ugrásai (vagy erős gradiensei) vannak, amelyek nem párhuzamosak a koordináta irányokkal.

Az alábbiakban a Peaceman–Rachford módszer elemzését mutatjuk meg. Foglalkozzunk először a képlethibájával! Ehhez legyenek a (16.200), (16.201),  $c\rho = 1$  feladat adatai olyanok, hogy  $u$  megoldása elég sima :  $u \in C^{4,3}(\bar{Q}_T)$ . Ekkor  $\Lambda_\ell u$  az  $\mathcal{L}_\ell u := \frac{\partial}{\partial x_\ell} (k \frac{\partial u}{\partial x_\ell})$  másodrendű approximációja, továbbá

$$u^j = \left( u - \frac{\tau}{2} \frac{\partial u}{\partial t} + \frac{\tau^2}{8} \frac{\partial^2 u}{\partial t^2} \right)^{j+1/2} + O(\tau^3),$$

így, amikor  $x \in \omega_h$ ,

$$\begin{aligned} \psi^{j+1/2} &:= \frac{u^{j+1/2} - u^j}{\tau/2} - \Lambda_1 u^{j+1/2} - \Lambda_2 u^j - f^{j+1/2} \\ &= \left( \frac{\partial u}{\partial t} - \mathcal{L}_1 u - \mathcal{L}_2 u - f \right)^{j+1/2} - \frac{\tau}{4} \frac{\partial^2 u^{j+1/2}}{\partial t^2} + \frac{\tau}{2} \mathcal{L}_2 \frac{\partial u^{j+1/2}}{\partial t} \\ &\quad + O(\tau^2 + h_1^2 + h_2^2) \\ &= -\frac{\tau}{4} \left( \frac{\partial^2 u}{\partial t^2} - 2\mathcal{L}_2 \frac{\partial u}{\partial t} \right)^{j+1/2} + O(\tau^2 + h_1^2 + h_2^2). \end{aligned} \quad (16.208)$$

Hasonlóan

$$\begin{aligned}
\psi^{j+1} &:= \frac{u^{j+1} - u^{j+1/2}}{\tau/2} - \Lambda_1 u^{j+1/2} - \Lambda_2 u^{j+1} - f^{j+1/2} \\
&= \left( \frac{\partial u}{\partial t} - \mathcal{L}_1 u - \mathcal{L}_2 u - f \right)^{j+1/2} + \frac{\tau}{4} \frac{\partial^2 u^{j+1/2}}{\partial t^2} - \frac{\tau}{2} \mathcal{L}_2 \frac{\partial u^{j+1/2}}{\partial t} \\
&\quad + O(\tau^2 + h_1^2 + h_2^2) \\
&= +\frac{\tau}{4} \left( \frac{\partial^2 u}{\partial t^2} - 2\mathcal{L}_2 \frac{\partial u}{\partial t} \right)^{j+1/2} + O(\tau^2 + h_1^2 + h_2^2). \tag{16.209}
\end{aligned}$$

Vegyük észre, hogy a képlethibák elsőrendű tagjai csak az előjelben különböznek, így valójában kioltják egymást, és a pontosság másodrendű lesz.

Ennek bizonyítására tekintsük a  $z := u - y$  hibafüggvényt (amelyet a  $j, j+1$  teljes időrétegeken  $x \in \bar{\omega}_h$ -ra, a közbülső  $j+1/2$  rétegeken viszont csak  $x \in \omega_h \cup \gamma_h^{(1)}$ -re definiálunk). Ez a következő feladat megoldása :

$$\begin{aligned}
z^0 &= 0, \quad x \in \bar{\omega}_h, \\
\frac{z^{j+1/2} - z^j}{\tau/2} - \Lambda_1 z^{j+1/2} - \Lambda_2 z^j &= \psi^{j+1/2}, \quad x \in \omega_h, \tag{16.210}
\end{aligned}$$

$$z^{j+1/2} = (g - \mu)^{j+1/2}, \quad x \in \gamma_h^{(1)}, \tag{16.211}$$

$$\frac{z^{j+1} - z^{j+1/2}}{\tau/2} - \Lambda_1 z^{j+1/2} - \Lambda_2 z^{j+1} = \psi^{j+1}, \quad x \in \omega_h, \tag{16.212}$$

$$z^{j+1} = 0, \quad x \in \gamma_h; \quad j = 0, 1, \dots, m-1.$$

A képlethibát első- és másodrendű részre ( $x \in \omega_h$ ) osztjuk fel :

$$\psi := \psi_{(1)} + \psi_{(2)}; \tag{16.213}$$

$$-\psi_{(1)}^{j+1/2} = \psi_{(1)}^{j+1} = \frac{\tau}{4} \left( \frac{\partial^2 u}{\partial t^2} - 2\mathcal{L}_2 \frac{\partial u}{\partial t} \right)^{j+1/2}, \tag{16.214}$$

$$\psi_{(2)}^{j+1/2} = O(\tau^2 + h_1^2 + h_2^2), \quad \psi_{(2)}^{j+1} = O(\tau^2 + h_1^2 + h_2^2).$$

(16.213)-nak megfelelően a (16.210)-(16.212) feladat megoldását a

$$z = \eta_{(1)} + \eta_{(2)} \tag{16.215}$$

alakban keressük. Itt  $\eta_{(1)}$  a következő feladat megoldása :

$$\begin{aligned} \eta_{(1)}^0 &= 0, \quad x \in \bar{\omega}_h; \\ \frac{\eta_{(1)}^{j+1/2} - \eta_{(1)}^j}{\tau/2} &= \psi_{(1)}^{j+1/2}, \quad x \in \omega_h; \end{aligned} \quad (16.216)$$

$$\begin{aligned} \frac{\eta_{(1)}^{j+1} - \eta_{(1)}^{j+1/2}}{\tau/2} &= \psi_{(1)}^{j+1}, \quad x \in \omega_h, \quad \eta_{(1)}^{j+1} = 0, \quad x \in \gamma_h; \\ j &= 0, 1, \dots, m-1. \end{aligned} \quad (16.217)$$

Az  $\eta_{(1)}^{j+1/2}$  függvény  $\gamma_h^{(1)}$ -beli értékei még nincsenek meghatározva. Ettől függetlenül a (16.216)-(16.217) definíciók szerint, (16.214)-nek köszönhetően, igaz

$$\eta_{(1)}^{j+1/2} = \frac{\tau}{2} \psi_{(1)}^{j+1/2} = -\frac{\tau^2}{8} \left( \frac{\partial^2 u}{\partial t^2} - 2\mathcal{L}_2 \frac{\partial u}{\partial t} \right)^{j+1/2}, \quad x \in \omega_h; \quad \eta_{(1)}^{j+1} = 0, \quad x \in \bar{\omega}_h,$$

minden  $j$ -re. Ez azt is jelenti, hogy

$$\eta_{(1)}^{j+1/2} = O(\tau^2), \quad j = 1, 2, \dots, m-1.$$

A (16.216)-(16.217) és (16.210)-(16.212) egyenletek különbségéből adódik az a feladat, amelynek  $\eta_{(2)}$  a megoldása. Ezen feladat jobb oldalán szerepel a  $\Lambda_1 \eta_{(1)}^{j+1/2}$  tag, mivel  $\Lambda_1 z^{j+1/2} = (\Lambda_1 \eta_{(1)} + \Lambda_1 \eta_{(2)})^{j+1/2}$ .  $\Lambda_1 \eta_{(1)}^{j+1/2}$  akkor lesz másodrendű  $\tau$ -ban és  $h_1$ -ben, ha  $u$  elég sima (mégpedig  $\mathcal{L}_2(\frac{\partial^2 u}{\partial t^2} - 2\mathcal{L}_2 \frac{\partial u}{\partial t}) \in C(\bar{Q}_T)$ ) és ha  $\eta_{(1)}^{j+1/2}$  képlete a  $\gamma_h^{(1)}$  peremen ugyanaz, mint  $\omega_h$ -n, vagyis ha

$$\eta_{(1)}^{j+1/2} = -\frac{\tau^2}{8} \left( \frac{\partial^2 u}{\partial t^2} - 2\mathcal{L}_2 \frac{\partial u}{\partial t} \right)^{j+1/2} = O(\tau^2), \quad x \in \gamma_h^{(1)}. \quad (16.218)$$

Most (16.211)-ből következik

$$(g - \mu)^{j+1/2} = \eta_{(1)}^{j+1/2} + \eta_{(2)}^{j+1/2}, \quad x \in \gamma_h^{(1)},$$

avagy (figyelembe véve, hogy  $u = g$ ,  $x \in \gamma_h$ )

$$\mu^{j+1/2} = g^{j+1/2} + \frac{\tau^2}{8} \left( \frac{\partial^2 g}{\partial t^2} - 2\mathcal{L}_2 \frac{\partial g}{\partial t} \right)^{j+1/2} - \eta_{(2)}^{j+1/2}, \quad x \in \gamma_h^{(1)}.$$

Választhatunk pl. (ehhez ld. a 42.a feladatot is)

$$\mu^{j+1/2} = g^{j+1/2} + \frac{\tau^2}{8} \left( \frac{\partial^2 g}{\partial t^2} - 2\mathcal{L}_2 \frac{\partial g}{\partial t} \right)^{j+1/2}, \quad \eta_{(2)}^{j+1/2} = 0, \quad x \in \gamma_h^{(1)}, \quad (16.219)$$

vagy

$$\mu^{j+1/2} = g^{j+1/2}, \quad \eta_{(2)}^{j+1/2} = -\eta_{(1)}^{j+1/2} = O(\tau^2), \quad x \in \gamma_h^{(1)}, \quad (16.220)$$

vagy

$$\mu^{j+1/2} = \frac{1}{2}(g^{j+1} + g^j) - \frac{\tau}{4}\Lambda_2(g^{j+1} - g^j), \quad \eta_{(2)}^{j+1/2} = O(\tau^2 + h_1^2), \quad x \in \gamma_h^{(1)}. \quad (16.221)$$

Ezután  $\eta_{(2)}$  feladatában, a (16.219) választás mellett :

$$\eta_{(2)}^0 = 0, \quad x \in \bar{\omega}_h;$$

$$\frac{\eta_{(2)}^{j+1/2} - \eta_{(2)}^j}{\tau/2} - \Lambda_1 \eta_{(2)}^{j+1/2} - \Lambda_2 \eta_{(2)}^j = \bar{\psi}_{(2)}^{j+1/2}, \quad x \in \omega_h, \quad (16.222)$$

$$\bar{\psi}_{(2)}^{j+1/2} := \psi_{(2)}^{j+1/2} + \Lambda_1 \eta_{(1)}^{j+1/2}; \quad \eta_{(2)}^{j+1/2} = 0, \quad x \in \gamma_h^{(1)};$$

$$\frac{\eta_{(2)}^{j+1} - \eta_{(2)}^{j+1/2}}{\tau/2} - \Lambda_1 \eta_{(2)}^{j+1/2} - \Lambda_2 \eta_{(2)}^{j+1} = \bar{\psi}_{(2)}^{j+1}, \quad x \in \omega_h, \quad (16.223)$$

$$\bar{\psi}_{(2)}^{j+1} := \psi_{(2)}^{j+1} + \Lambda_1 \eta_{(1)}^{j+1/2}; \quad \eta_{(2)}^{j+1} = 0, \quad x \in \gamma_h;$$

$$j = 0, 1, \dots, m-1,$$

a  $\bar{\psi}_{(2)}$  jobboldali rácsfüggvények mind  $O(\tau^2 + h_1^2 + h_2^2)$  nagyságrendűek, ha  $u$  elég sima.

Ha most a (16.222)-(16.223) feladat stabilitása pl. a  $\|\cdot\|_{(0,h)}$  normában ismert :

$$\|\eta_{(2)}^j\|_{(0,h)} \leq C \sum_{k=0}^{j-1} \tau \left( \|\bar{\psi}^{k+1/2}\|_{(0,h)} + \|\bar{\psi}^{k+1}\|_{(0,h)} \right), \quad j = 1/2, 1, 3/2, 2, \dots, m,$$

akkor (16.215) és  $\eta_{(1)} = O(\tau^2)$  miatt következik a háromszög egyenlőtlenség segítségével

$$\|z^j\|_{(0,h)} \leq \|\eta_{(1)}^j\|_{(0,h)} + \|\eta_{(2)}^j\|_{(0,h)} = O(\tau^2 + h_1^2 + h_2^2), \quad j = 1/2, 1, 3/2, 2, \dots, m.$$

A  $\mu^{j+1/2}$  peremérték kényelmesebb (16.220) és (16.221) választása esetén viszont kell egy peremérték szerinti stabilitási eredmény, ezt ld. a szerző 1971-es cikkében.

A Peaceman–Rachford eljárás stabilitását akkor nem nehéz megmutatni, amikor homogének a peremfeltételek és érvényes

$$A_{1,h} = A_{1,h}^T > 0, \quad A_{2,h} = A_{2,h}^T > 0, \quad A_{1,h}A_{2,h} = A_{2,h}A_{1,h}, \quad (16.224)$$

ahol a  $(A_{\ell,h}y_h)(x) := (\Lambda_\ell y)(x)$ ,  $x \in \omega_h$ , relációkkal a differencia-operátorokhoz a homogén peremfeltételeknek megfelelő mátrixokat vezettük be.

Az első két feltétel teljesül, ha  $\mathcal{L}_1$  és  $\mathcal{L}_2$  is önadjungált, pozitív definit operátorok, és ha végeelem approximációt vagy megfelelő véges differencia approximációt használunk úgy, hogy az  $A_{1,h}, A_{2,h}$  mátrixok is ezekkel a tulajdonságokkal rendelkeznek. A harmadik feltétellel még nem találkoztunk. Ez a felcserélhetőség pl. akkor teljesül, ha

$$\mathcal{L}_\ell u = \frac{\partial}{\partial x_\ell} \left( k_\ell(x_\ell) \frac{\partial u}{\partial x_\ell} \right), \quad \ell = 1, 2, \quad (16.225)$$

tehát ha  $k_1$  nem függ  $x_2$ -től és  $k_2$  nem függ  $x_1$ -től;

$$\text{a tartomány téglalap;} \quad (16.226)$$

és ha

$$\Lambda_\ell y := (k_\ell y_{x_\ell})_{\tilde{x}_\ell}, \quad x \in \omega_h, \quad (16.227)$$

a  $\mathcal{L}_\ell u$  standard-approximációja az  $\omega_h$  kétdimenziós,  $h_1, h_2$  lépéstávolságú ekvidisztáns rácson.

Ehhez ld. a 41. feladatot. A (16.224) feltételek mellett  $A_{1,h}$  és  $A_{2,h}$  közös, ortonormált

$$\{w_i = w_{i_1, i_2} = w_{i_1} \cdot w_{i_2}\}_{i_1, i_2=1}^{N_1-1, N_2-1} \quad (16.228)$$

sajátvektor-rendszerrel rendelkeznek (ld. pl. Gantmaher I-ben idézett könyvét), a („diszkrét”) sajátértékek pozitívak :

$$\begin{aligned} A_\ell w_i &= \lambda_{i_\ell}^{h_\ell} w_i, & \lambda_{i_\ell}^{h_\ell} &> 0, \\ i &= (i_1, i_2), \quad i_\ell = 1, \dots, N_\ell - 1, & \ell &= 1, 2. \end{aligned}$$

**16.24. Tétel** (a Peaceman–Rachford eljárás stabilitása). Ha teljesülnek a (16.224) feltételek és a peremértékek homogének :  $(g - \mu)^{j+1/2} = 0$ , akkor a Peaceman–Rachford eljárás stabil az  $L_2(\omega_h)$ -normában. Ekkor a (16.210)-(16.212) feladat megoldása eleget tesz a következő becslésnek :

$$\|z^j\|_{(0,h)} \leq \left\{ \frac{T}{2} \sum_{k=0}^{j-1} \tau (\|\psi^{k+1/2}\|_{(0,h)}^2 + \|\psi^{k+1}\|_{(0,h)}^2) \right\}^{1/2}. \quad (16.229)$$

**Kommentár.** Feltételeink mellett a (16.210)-(16.212) és a (16.222)-(16.223) feladatok csak jelölésekben különböznek.  $\square$



**Bizonyítás.** A (16.210)-(16.212) feladat megoldását a

$$z^j(x) = \sum_{i=(i_1, i_2)} c_i^j w_i(x), \quad x = (x_1, x_2) \in \omega_h,$$

alakban keressük, és ehhez a  $\psi$  jobboldalt is a  $\{w_i\}$  rendszer szerint sorba fejtjük :

$$\psi^k(x) = \sum_{i=(i_1, i_2)} d_i^k w_i(x), \quad x \in \omega_h, \quad k = j + 1/2, j + 1.$$

Skalárisan megszorozva a (16.210)-(16.212) egyenleteit  $w_\ell$ -l (ahol  $\ell = (\ell_1, \ell_2)$ ), kapjuk, hogy

$$\begin{aligned} \frac{c_\ell^{j+1/2} - c_\ell^j}{\tau/2} + \lambda_{\ell_1}^{h_1} c_\ell^{j+1/2} + \lambda_{\ell_2}^{h_2} c_\ell^j &= d_\ell^{j+1/2}, \\ \frac{c_\ell^{j+1} - c_\ell^{j+1/2}}{\tau/2} + \lambda_{\ell_1}^{h_1} c_\ell^{j+1/2} + \lambda_{\ell_2}^{h_2} c_\ell^{j+1} &= d_\ell^{j+1}. \end{aligned}$$

Innen

$$\begin{aligned} \left(1 + \frac{\tau}{2} \lambda_{\ell_1}^{h_1}\right) c_\ell^{j+1/2} &= \left(1 - \frac{\tau}{2} \lambda_{\ell_2}^{h_2}\right) c_\ell^j + \frac{\tau}{2} d_\ell^{j+1/2}, \\ \left(1 + \frac{\tau}{2} \lambda_{\ell_2}^{h_2}\right) c_\ell^{j+1} &= \left(1 - \frac{\tau}{2} \lambda_{\ell_1}^{h_1}\right) c_\ell^{j+1/2} + \frac{\tau}{2} d_\ell^{j+1}, \end{aligned}$$

vagyis, a

$$\rho_k = \rho_{\ell_k}^{h_k} := \frac{1 - \frac{\tau}{2} \lambda_{\ell_k}^{h_k}}{1 + \frac{\tau}{2} \lambda_{\ell_k}^{h_k}}, \quad k = 1, 2,$$

jelöléssel,

$$c_\ell^{j+1} = \rho_1 \rho_2 c_\ell^j + \frac{\tau}{2 \left(1 + \frac{\tau}{2} \lambda_{\ell_2}^{h_2}\right)} \left[ \rho_1 d_\ell^{j+1/2} + d_\ell^{j+1} \right].$$

A sajátértékek pozitivitása miatt mindig érvényes  $|\rho_1|, |\rho_2| < 1$ , és így

$$|c_\ell^{j+1}| \leq |c_\ell^j| + \frac{\tau}{2} \left[ |d_\ell^{j+1/2}| + |d_\ell^{j+1}| \right].$$

Ezt az egyenlőtlenséget  $j = 0$ -ig követjük vissza, figyelembe véve, hogy  $c_\ell^0 = 0$  :

$$|c_\ell^{j+1}| \leq \sum_{k=0}^j \frac{\tau}{2} \left[ |d_\ell^{k+1/2}| + |d_\ell^{k+1}| \right]. \quad (16.230)$$

Most azt használjuk, hogy a (16.228) rendszer ortonormáltsága miatt érvényes

$$\|z^j\|_{(0,h)}^2 = \sum_{\ell=(\ell_1,\ell_2)} |c_\ell^j|^2, \quad \|\psi^k\|_{(0,h)}^2 = \sum_{\ell=(\ell_1,\ell_2)} |d_\ell^k|^2, \quad k = j + 1/2, j + 1,$$

stb. Ezért először négyzetre emeljük (16.230)-ot, majd a Cauchy–egyenlőtlenséget alkalmazzuk :

$$\begin{aligned} |c_\ell^{j+1}|^2 &\leq \left( \sum_{k=0}^j \frac{\tau}{2} [ |d_\ell^{k+1/2}| + |d_\ell^{k+1}| ] \right)^2 \leq \frac{\tau^2(j+1)}{4} \sum_{k=0}^j [ |d_\ell^{k+1/2}| + |d_\ell^{k+1}| ]^2 \\ &\leq t_{j+1} \frac{\tau}{4} \sum_{k=0}^j 2 [ |d_\ell^{k+1/2}|^2 + |d_\ell^{k+1}|^2 ]. \end{aligned}$$

Innen a (16.229) becslés következik  $\ell = (\ell_1, \ell_2)$  feletti összegzéssel :

$$\|z^{j+1}\|_{(0,h)}^2 = \sum_{\ell} |c_\ell^{j+1}|^2 \leq \frac{t_{j+1}}{2} \sum_{k=0}^j \tau (\|\psi^{k+1/2}\|_{(0,h)}^2 + \|\psi^{k+1}\|_{(0,h)}^2). \quad \square$$

A fenti feltételek mellett a 16.5–16.7. tételek segítségével is megmutatható a stabilitás (ld. a 42.b feladatot), de van olyan bizonyítás is (ld. a 43. feladatot), amely a felcserélhetőség nélkül is működik, tehát ekkor a tartomány téglalapokból összeállított is lehet.

Befejezésül megjegyezzük, hogy a Peaceman–Rachford módszert Frjazinov és Gulin tetszőleges kétdimenziós tartományra általánosították : azt bizonyították be, hogy ekkor a pontosság legalább  $O(\tau + h^{3/2})$ , és ehhez a (16.224)-beli felcserélhetőségre nem volt szükségük.

A végeelem diszkretizáció egyébként akkor alkalmasabb arra, hogy a váltakozó irányok módszeréhez áttérjünk, ha tenzorszorzatok a bázisfüggvényei (mint egyes Lagrange–elemek esetén, pl. az  $R_4$  elem ilyen, de az  $R_8$  nem). De háromszöges végeelem diszkretizáció esetére is általánosították a váltakozó irányok módszerét. Ebben az esetben a vizsgálatot tetszőleges kétdimenziós tartományra is elvégezték, és sikerült bebizonyítani, hogy legalább  $O(\tau + h)$  sebességgel konvergál az eljárás, ld. M. Dryja dolgozatát 1979-ből.

### 16.7.2 A Peaceman–Rachford módszer mint simító eljárás

Az előbb bemutatott váltakozó irányok módszere igen hasznos mint simító eljárás a többrácsos módszer keretében. Mint ilyen – a Gauss–Seidel módszerrel összehasonlítva – nem éppen olcsó, de igen robusztus, és mint simító

főként azért kedvelt, mert anizotróp feladatok kezelésére is alkalmas. Anizotróp a (16.200)-típusú feladat akkor, ha a  $k$  egységes vezetőképesség helyett irányonként (erősen) különböző  $k_\ell$  együtthatók állnak,  $\ell = 1, 2$ .

Példa ilyen anizotróp feladatra az is, hogy megoldandó a

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + f = 0, \quad 0 \leq x \leq 1, \quad 0 \leq y \leq L \gg 1, \quad (16.231)$$

egyenlet, mondjuk homogén elsőfajú peremfeltételekkel. Új független változók bevezetésével láthatóvá tehetjük, hogy itt valóban van probléma :

$$x_1 = x, \quad x_2 = y/L : \quad 0 \leq x_1, x_2 \leq 1;$$

ugyanis ezután az egyenlet

$$\frac{\partial^2 u}{\partial x_1^2} + \varepsilon \frac{\partial^2 u}{\partial x_2^2} + f = 0, \quad \varepsilon := 1/L^2. \quad (16.232)$$

Ekkor a többrácsos módszer általában rosszul konvergál, mert  $x_2$  irányban ilyenkor a szokásos simító iterációk hatása nem megfelelő, a kis  $\varepsilon$  miatt (• amit az „Ellipsis” programmal azonnal ellenőrizhetjük).

Speciális esetünkben azt lehet tenni, hogy az  $x_1$ -irányban pontosan oldjuk meg az egyenleteket (16.206) alakú számítási lépésekkel (és  $x_2$  szolgál mint külső paraméter), hiszen így az  $\varepsilon = 0$  esetben is gyorsan a pontos megoldáshoz közeledünk. Ez az elv nem változik, ha a  $\frac{\partial^2 u}{\partial x_2^2}$  szorzója 1, de  $\frac{\partial^2 u}{\partial x_1^2}$  szorzója  $\varepsilon$ . Ekkor tehát (16.207)-féle lépésekkel iterálhatunk.

Amikor viszont mindkét második derivált előtt áll egy-egy olyan függvény, amely hol nagy hol kicsi értéket vesz fel (ami nemlineáris esetben még előre ismeretlen helyeken is történhet), akkor a (16.206)-(16.207) váltakozó irányok módszere segít.

Elemezzük a váltakozó irányok módszerének alkalmazását (16.231)-re! Mivel fenti feladatunk nem időtől függő, így (16.206)-(16.207) mindkét fél-lépésében  $\tau/2$  helyett más-más fiktív időlépést vehetünk. Tehát a váltakozó irányok módszere mint simító iteráció :

$$\begin{aligned} \frac{v^{j+1/2} - v^j}{\tau_1} + A_{1,h}v^{j+1/2} + A_{2,h}v^j &= f, \\ \frac{v^{j+1} - v^{j+1/2}}{\tau_2} + A_{1,h}v^{j+1/2} + A_{2,h}v^{j+1} &= f, \end{aligned}$$

ahol  $(A_{1,h}v)_{k,\ell} = (-v_{\bar{x}x})_{k,\ell}$ ,  $(A_{2,h}v)_{k,\ell} = (-v_{\bar{y}y})_{k,\ell}$  a homogén peremfeltételeknek megfelelően, és  $\tau_\ell$  az iterációs paraméterek,  $\ell = 1, 2$ .

Már a Peaceman–Rachford séma stabilitási elemzésénél feltűnhetett, hogy az tulajdonképpen a Crank–Nicolson séma vizsgálatára vezetett vissza, tehát az egydimenziós esetre. A sajátvektorok szerinti sorfejtés során láttuk, hogy döntő a  $\rho_1 \rho_2 = \rho_{\ell_1}^{h_1} \rho_{\ell_2}^{h_2}$  szorzat. Legyen most  $N_1$  és  $N_2$  kettőre osztható, és – figyelembe véve, hogy

$$\lambda_{N_k/2}^{h_k} = 2/h_k^2, \quad \lambda_{N_k-1}^{h_k} \lesssim 4/h_k^2, \quad k = 1, 2,$$

– követeljük meg azt (v.ö. (15.88)-cal 15.4.2-ben), hogy a hatékony simítás érdekében legyen

$$\rho_{N_k/2}^{h_k} = \left(1 - \frac{2\tau_k}{h_k^2}\right) / \left(1 + \frac{2\tau_k}{h_k^2}\right) = - \left(1 - \frac{4\tau_k}{h_k^2}\right) / \left(1 + \frac{4\tau_k}{h_k^2}\right), \quad k = 1, 2.$$

Ekkor ugyanis

$$|\rho_\ell^{h_k}| \leq \rho_{N_k/2}^{h_k}, \quad \ell = N_k/2, N_k/2 + 1, \dots, N_k - 1, \quad k = 1, 2.$$

Ez a követelmény azt jelenti  $\kappa = \kappa_k := \tau_k/h_k^2$  számára ( $k = 1, 2$ ), hogy  $1 + 2\kappa - 8\kappa^2 = -(1 - 2\kappa - 8\kappa^2)$  vagyis

$$\kappa = \kappa_1 = \kappa_2 = \frac{\sqrt{2}}{4}.$$

Ekkor

$$\rho_{N_k/2}^{h_k} = \frac{2 - \sqrt{2}}{2 + \sqrt{2}} = 3 - 2\sqrt{2} \approx 0.172$$

mindkét irányban ( $k = 1, 2$ ), függetlenül  $L \gg 1$ -től. Így  $\rho_{N_1/2}^{h_1} \rho_{N_2/2}^{h_2} = 17 - 12\sqrt{2} \approx 0.0296$  és

$$\rho_{\ell_1}^{h_1} \rho_{\ell_2}^{h_2} \lesssim \begin{cases} 1, & 1 \leq \ell_k < N_k/2, \quad k = 1, 2, \\ 0.172, & 1 \leq \ell_1 < N_1/2, \quad N_2 \leq \ell_2 \leq N_2 - 1, \\ 0.172, & N_1 \leq \ell_1 \leq N_1 - 1, \quad 1 \leq \ell_2 < N_2/2, \\ 0.0296, & N_k \leq \ell_k \leq N_k - 1, \quad k = 1, 2. \end{cases}$$

Változó együtthatók esetén, az optimális  $\tau_1, \tau_2$  paraméterek meghatározásának nehézsége miatt, inkább másképpen értelmezzük a váltakozó irányok módszerét, a parabolikus egyenletre való hivatkozás nélkül. Legyen az elliptikus feladat pl. (16.231) és

$$A_{1,h}y + A_{2,h}y_h = f_h$$

a véges differencia eljárással előállított lineáris rendszer.

$A_{\ell,h}$  tehát az  $\ell$ -edik koordinátairányhoz tartozik. Legyen

$$A_{\ell,h} = D_{\ell,h} - B_{\ell,h}, \quad \ell = 1, 2,$$

az additív felbontása főátlós részre és maradékra. Ekkor

$$(A_{1,h} + D_{2,h})v_h^{j+1/2} = B_{2,h}v_h^j + f_h, \quad (16.233)$$

$$(A_{2,h} + D_{1,h})v_h^{j+1} = B_{1,h}v_h^j + f_h, \quad j = 0, 1, \dots, \quad (16.234)$$

a javasolt iteráció, amely a tapasztalatok szerint jól alkalmazható. Ha pl. a (16.232) feladat többrácsos megoldás alatti simítására használjuk, akkor  $\varepsilon = 0$  esetén egy (16.233)-(16.234) lépéssel a pontos megoldást érjük el, de ugyanaz igaz, ha (16.232)-ben az  $\varepsilon = 0$  együttható az  $x_1$  szerinti második derivált előtt állna.

Vegyük észre, hogy (16.233)-(16.234) majdnem a Peaceman–Rachford sémával esik egybe, hiszen (16.233)-(16.234) ekvivalens alakja

$$\begin{aligned} D_{2,h}(v_h^{j+1/2} - v_h^j) + A_{1,h}v_h^{j+1/2} + A_{2,h}v_h^j &= f_h, \\ D_{1,h}(v_h^{j+1} - v_h^{j+1/2}) + A_{1,h}v_h^{j+1/2} + A_{2,h}v_h^{j+1} &= f_h. \end{aligned}$$

Amennyiben két olyan  $\kappa_\ell$  konstans van, hogy  $D_{\ell,h} = \frac{2\kappa_\ell}{h_\ell^2}I_h$ ,  $\ell = 1, 2$ , akkor viszont a (16.233)-(16.234) iteráció pontosan a Peaceman–Rachford eljárásnak felel meg, ahol ekkor  $\tau_\ell/2 = h_\ell^2/(2\kappa_\ell)$  a fél időlépés.

Befejezésül hasonlítsuk össze a csillapított Jacobi-iterációt és a váltakozó irányok módszerét a kétdimenziós izotróp esetben (amikor (16.232)-ben  $\varepsilon = 1$ ), mint a többrácsos módszer két simító iterációját!

Ekkor, négyzet alakú rácson és az  $\ell$ -edik szinten ( $h_\ell = 1/N_\ell$  lépéstávolsággal)  $\omega_\ell = \frac{h_\ell^2}{6}$  választással  $\rho = 0.6$  lesz a csillapított Jacobi-iteráció hibacsökkenési rátája (ld. a 16. feladatot 15.11-ben). Egy iterációhoz  $5N_\ell$  összeadás és  $3N_\ell$  szorzás kell.

A váltakozó irányok módszere az előzőek szerint  $\rho = 3 - 2\sqrt{2} \approx 0.172$  csökkenést biztosít a magasfrekvenciájú hibakomponenseknél. Ekkor  $2N_\ell^{1/2}$  összeadásba és  $3N_\ell^{1/2}$  szorzásba kerül a tridiagonális algoritmus alkalmazása  $x_1$ -irányban, rögzített  $x_2$ -re (ahol az algoritmus  $\alpha_i$  és  $d_i = 1/(b_i - a_i\alpha_i)$  együtthatóinak kiszámítását elhagytuk, mivel azt csak egyszer kell elvégezni). Így az első fellépéshez  $2N_\ell$  összeadás és  $3N_\ell$  szorzás szükséges, és ennek duplája a két fellépés.

A csillapított Jacobi-iterációnak 3-4. lépésében elérjük a váltakozó irányok módszere 1 lépésének simítási teljesítményét, és ehhez körülbelül a dupla műveletszám szükséges. Ezért a kétdimenziós esetben a váltakozó irányok

módszere hatékonyabb mint a csillapított Jacobi–iteráció. De mindkettőnél hatékonyabb a Gauss–Seidel módszer, így a váltakozó irányok módszerét valóban csak az erősen anizotróp esetben használjuk (amikor a másik két eljárásnál érezhetően jobb).

### 16.7.3 Kettőnél több dimenziós parabolikus egyenletek megoldása

Röviden vázoljuk a több mint kétdimenziós (16.200)–(16.201),  $c\rho = 1$  feladatok numerikus megoldását a váltakozó irányok módszereivel.

Legyen a (16.200)-beli  $\mathcal{L}u$  approximációja a  $h_1, \dots, h_d$ -lépéstávolságú rácsra

$$\Lambda y = \sum_{\ell=1}^d \Lambda_{\ell} y, \quad (\Lambda_{\ell} y)_i := -(k_{\ell} y_{\bar{x}_{\ell}})_{\bar{x}_{\ell}, i}, \quad \text{és } \varphi^j := f^{j+\sigma}. \quad (16.235)$$

Ezekkel a jelölésekkel a (16.202) sémát a következő alakba írhatjuk át :

$$(I + \sigma\tau\Lambda)y^{j+1} = (I - (1 - \sigma)\tau\Lambda)y^j + \tau\varphi^j,$$

és ( $O(\tau^2)$ -rendű tagokat hozzáadva) innen kapjuk, (16.235)-öt figyelembe véve,

$$\prod_{\ell=1}^d (I + \sigma\tau\Lambda_{\ell})y^{j+1} = \prod_{\ell=1}^d (I - (1 - \sigma)\tau\Lambda_{\ell})y^j + \tau\varphi^j. \quad (16.236)$$

Ezen séma pontossága általában csupán  $O(\tau + h^2)$ , ahol  $h^2 := h_1^2 + \dots + h_d^2$ , de  $\sigma = \frac{1}{2}$  esetén  $O(\tau^2 + h^2)$ . Speciális esetekben (16.236) rendje még magasabb. Ha  $\mathcal{L}$  a Laplace–operátor és az egyes koordinátairányokban más-más súlyt veszünk :

$$\sigma \implies \sigma_{\ell} = \frac{1}{2} - \frac{h_{\ell}^2}{12\tau}, \quad \text{valamint } \varphi^j := \left( I - \sum_{\ell=1}^n \frac{h_{\ell}^2}{12} \Lambda_{\ell} \right) f^{j+1/2}$$

(v.ö. (16.37)-tel), akkor  $O(\tau^2 + h^4)$ -rendű a séma.

Végül is, ha minden koordinátairányban ugyanazt a  $h$  lépéstávolságot használjuk és  $\mathcal{L}$  a Laplace–operátor, akkor itt is (mint az egydimenziós esetben) van hatodrendű séma, ld. a 44. feladatot.

A (16.236)-ban fellépő szorzatok miatt *faktorizált sémáról* is beszélünk. A séma numerikus megoldása szempontjából az a döntő, hogy az egyes  $I + \sigma\tau\Lambda_{\ell}$  szorzók olyan differenciáloperátornak felelnek meg, amely csupán az  $x_{\ell}$

koordinátára hat. Továbbá, minden  $I + \sigma\tau\Lambda_\ell$  differencia operátorhoz tartozó mátrix tridiagonális. Ez azért lényeges, mert (16.236) szétbontható

$$(I + \sigma\tau\Lambda_\ell)v_\ell = w_\ell, \quad \ell = 1, \dots, d,$$

alakú rendszerekre, amelyeknek mindegyike a rövidített Gauss-eliminációval oldható meg. Ilyen szétbontásra különféle módok vannak.

a) Janenko „közbülső lépések” módszere :

$$(I + \sigma\tau\Lambda_\ell)y^{j+\ell/d} = (I - (1 - \sigma)\tau\Lambda_\ell)y^{j+(\ell-1)/d} + \tau\delta_{1i}\varphi^j, \quad (16.237)$$

$$\ell = 1, \dots, d; \quad (\delta_{1i} = 0, \quad i \neq 1, \quad \delta_{11} = 1).$$

Ezen eljárás stabilitása egyszerűen az egydimenziós súlyozott differenciaséma stabilitásából következik : a  $\Lambda_\ell$  operátorok felcserélhetősége nem kell hozzá.

b) Djakonov változata :

$$(I + \sigma\tau\Lambda_1)y^{j+1/d} = \prod_{\ell=1}^d (I - (1 - \sigma)\tau\Lambda_\ell)y^j + \tau f^{j+1/2}, \quad (16.238)$$

$$(I + \sigma\tau\Lambda_\ell)y^{j+\ell/d} = y^{j+(\ell-1)/d}, \quad \ell = 2, \dots, d. \quad (16.239)$$

Ezen eljárás stabilitása a 16.24. tétel mintájára megmutatható.

c) A szerző változata :

$$(I + \sigma_\ell\tau\Lambda_\ell)y^{j+\ell/d} = (I - (1 - \sigma_\ell)\tau\Lambda_{\ell-1})y^{j+(\ell-1)/d} + \tau\delta_{1i}\varphi^j, \quad (16.240)$$

$$\ell = 1, \dots, d, \quad \varphi^j := \left( I + \frac{\tau}{2} \sum_{k=1}^{d-1} \Lambda_k - \tau\left(\frac{1}{2} - \sigma_d\right)\Lambda_d \right) f^{j+1/2}.$$

Ez a változat a peremfeltételek szempontjából kényelmesebb :  $y^{j+\ell/d}$  meghatározására csak az  $x_\ell$  irányban kell peremfeltételeket előírni. Stabilitása a 43. feladatban bemutatott módon következik, ha minden  $\Lambda_\ell$  operátor valóban pozitív definit és  $\sigma_\ell \geq 1/2$ . Az operátorok felcserélhetőségére nincs szükség.

Akármelyik változatról lenne szó, az  $y^{j+\ell/d}$  függvények számára kétféleképpen lehet peremfeltételeket megadni. Az egyik lehetőség az, hogy a (16.236) alakban az  $y^{j+1}$  és  $y^j$  peremfeltételeit elimináljuk. Ekkor viszont a differencia operátorok szorzatai hatnak  $g$ -re is. Az elimináció után a  $\Lambda_\ell$  operátoroknak megfelelő  $A_{\ell,h}$  mátrixok és az  $y_h^{j+1}, y_h^j$  vektorok bevezethetők, amelyeknek nincsenek peremértékei (formálisan homogének). Ezután a  $\Lambda_\ell \rightarrow A_{\ell,h}, y \rightarrow y_h$  stb. változtatásokkal a (16.237), (16.238), (16.240) algoritmusok alkalmazhatók.

A másik lehetőség abból indul ki, hogy az  $y^{j+\ell/d}$  közbülső megoldások kapcsolatba hozhatók  $u(x, t_{j+\ell/d})$ -nel. Így a (16.206)-(16.207) Peaceman-Rachford séma esetén az  $u|_{\Gamma} = g(t)$  inhomogén peremfeltételek diszkrét megfelelője

$$y^{j+1/2} = g^{j+1/2}, \quad x \in \gamma_h^{(1)}; \quad y^{j+1} = g^{j+1}, \quad x \in \gamma_h,$$

(részletesebben ld. a 16.7.1. pontot), míg (16.240) esetén a diszkrét peremfeltételek

$$y^{j+\ell/d} = g^{j+\sigma_\ell} - \tau \left[ (1 - \sigma_\ell) \sum_{k=1}^{\ell-1} \Lambda_k - \sigma_\ell \sum_{k=\ell+1}^{d-1} \Lambda_k + (1 - \sigma_\ell - \sigma_d) \Lambda_d \right] g^{j+\sigma_\ell} + \tau(1 - \sigma_\ell) f^{j+1/2},$$

$$x \in \gamma_h^{(\ell)}, \quad \ell = 1, \dots, d-1; \quad y^{j+1} = g^{j+1}, \quad x \in \gamma_h.$$

Mindhárom esetben az  $y^{j+1}$  megoldást úgy kapjuk  $y^j$ -ből, hogy először az  $y^{j+\ell/d}$  segédfüggvényeket határozzuk meg,  $\ell = 1, \dots, d-1$ , majd  $y^{j+1}$ -et. Mindegyik  $y^{j+\ell/d}$  előállításához az egész megoldási tartományon végig kell mennünk : a számítás  $x_\ell$ -irányban folyik, az  $x_1, \dots, x_{\ell-1}, x_{\ell+1}, \dots, x_d$  koordináták külső ciklusban változó paraméterek.

Amiután ezzel elkészültünk, következik egy hasonló,  $x_{\ell+1}$ -iránnyal kapcsolatos lépés, stb. Az egész tartományon végül is  $d$ -szer megyünk végig, és mivel a rövidített Gauss-elimináció műveletigénye lineárisan nő az ismeretlenek számával, összesen  $O(d \cdot N)$  művelet szükséges az  $y^j \rightarrow y^{j+1}$  átmenet-höz, ahol  $N$  a térbeli megoldási tartomány összes diszkrétizációs pontnak a száma.

Az első váltakozó irányok módszere volt a 16.7.1-ben tárgyalt Peaceman-Rachford séma. Ez a c) algoritmus kétdimenziós esete, ha  $\sigma_\ell = 1/2$ . Azzal a tulajdonságával, hogy az  $y^{j+1/2}$ ,  $y^{j+1}$  megoldásvektorai mind közelítik a (16.200)-(16.201) feladat megoldását (ld. (16.208), (16.209)), a fenti (16.237)-(16.240) algoritmusok  $y^{j+\ell/d}$  vektorai általában csak  $\ell = 0$  esetén rendelkeznek.

A fenti váltakozó irányok módszerei akkor különösen javasolhatóak, amikor a feladat igen nagyméretű, forrástagja nem tartalmaz erős nemlinearitást. Ha az utolsó körülmény lép fel, akkor vagy térjünk át a többrácsos módszerhez, vagy pedig, kevesebb dimenziójú feladatok esetén, a szemidiszkrétizált egyenleteket használjuk (ld. a 16.4.1., 16.4.6. és 16.8. pontokat). Az onnan eredő,

$$G\dot{y} + Ay = \varphi \tag{16.241}$$

alakú, közönséges differenciálegyenletekből álló rendszerekre az a jellemző, hogy merevek : az  $A$ , ill.  $G^{-1}A$  mátrixok sajátértékei pozitívak és  $\lambda_{\max}/\lambda_{\min}$



nagy – ami a II. 10.5-ben tárgyalt módszereknek a bevetését szükségessé teszi – amelyek jól működnek az  $\varphi$ -beli nemlinearitások jelenlétében is.

Amikor a feladatok mérete viszont túl nagy (ha két- vagy háromdimenziós a tartomány, és különösen akkor, amikor már maga az  $u$  is vektor) és ha ugyanakkor erős nemlinearitások jelen vannak (mint légszennyeződési számításoknál, ahol az  $u$  vektor különböző kémiai komponensek koncentrációit tartalmaz), akkor ezek a módszerek túl költségesek. Ekkor ajánlható, hogy használjuk a *fizikai jelenségek szerinti felbontás* (avagy *splitting*) elvét! Ezt először a (16.241) rendszerrel szemléltetjük, amikor ott  $G = I$ , a differencia módszernek megfelelően.

Első lépésben, pl. kétréteges sémával, oldjuk meg az

$$\dot{y} + Ay = 0$$

egyenletet a  $t_j \leq t \leq t_{j+1}$  intervallumon, ahol  $y(t_j)$  adott mint kezdetiértékek vektora. Így kapjuk az  $\bar{y}(t)$  közbülső megoldást, ami pl. a diffúzió (vagy, az  $A$  interpretációjától függően, a hővezetés vagy a transzport) figyelembevételét jelentheti. Ezután második lépésben az

$$\dot{y} = \varphi(y, t)$$

rendszert oldjuk meg  $\bar{y}(t_{j+1})$ -ből mint kezdetiértékből indulva ki, ugyancsak a  $t_j \leq t \leq t_{j+1}$  intervallumon. Tehát csak most vesszük figyelembe a kémiai kinetikát vagy a nemlineáris hőforrásokat – de „lekapcsoltuk” a diffúzióból adódó merevséget.

Erre a második lépésre gyakran Adams- ill. explicit Runge–Kutta-módszer is alkalmas lehet. A Gear-féle vagy más merev egyenletekre alkalmas módszerekhez csak akkor kell folyamodnunk, amikor a kémiai kinetika nagyon különböző időskálákban lefolyó reakciókat tartalmaz.

Szembetűnő, hogy a vázolt eljárás pontossága csak elsőrendű  $\tau$ -ban, előnye viszont az, hogy itt modulokra bontható szét az összetett feladat, és minden egyes részlépésben jobban alkalmazkodhatunk a speciális körülményekhez. Így lehet, hogy az elméletileg magas pontossági rendet nem, de a fizikailag jobban interpretálható eredményekhez jutunk. (Másodrendű eljárást úgy kapunk, ha minden második lépésben a részlépések sorrendjét megfordítjuk.)

Vegyük észre, hogy a váltakozó irányok módszerei is értelmezhetők mint a fizikai jelenségek szerinti felbontás : pl. a (16.237) Janenko-módszer szerint először az első, majd a második stb. irányban lefuttatjuk a hővezetés folyamatát. (Szamarszkij 1984-es könyvében (284–286. o.) több példa található arra, hogy a folytonos kezdetiérték-peremérték feladat ilyen felbontása a pontos megoldásra vezethet.)

Összetettebb példát a fizikai jelenségek szerinti felbontásra a következő feladat nyújtja. Oldjuk meg, megfelelő kezdeti- és peremértékek mellett, az alábbi  $d$ -dimenziós egyenletet!

$$\frac{\partial T}{\partial t} + \operatorname{div}(T\vec{u}) = \operatorname{div}(k \operatorname{grad} T) + f(T),$$

ahol  $\vec{u} = (u_1, \dots, u_d)^T$ . Ilyen egyenletek approximációjával pl. \*\*\* 17.4-ben foglalkozunk, ld. ott (52). A  $c\rho\frac{\partial T}{\partial t} + \operatorname{div}(T\vec{u})$  rész a hő szállítását írja le, a  $\operatorname{div}(k \operatorname{grad} T)$  a hő diffúzióját, és végül  $f(T)$  a hőtermelést adja. Ennek megfelelően az egyenlet megoldásánál pl. a következő, egészen különböző modulokra lehet gondolni.

1. Transzport-lépés :

$T(t_j)$  adott mint kezdetivektor az  $\omega_h$  rácson, és a  $t_j \leq t \leq t_{j+1}$  intervallumon oldjuk meg a hiperbolikus

$$\frac{\partial T}{\partial t} + \operatorname{div}(T\vec{u}) = 0$$

transzport-egyenletet. Ennek során arra figyelünk, hogy az esetlegesen jelenlevő meredek hőmérsékleti gradiensek nem kenődnek szét, hiszen a diffúzió most ki van kapcsolva. Erre pl. a karakterisztikák módszere alkalmas (ld. \*\*\* 18.10.1. pontban), ill. annak differencia-approximációja, amely a

$$\tau \leq h / \max_{(x,t) \in Q_T} |\vec{u}|$$

Courant-feltétel mellett stabil (ld. \*\*\* 18.4.1-2.). Az eredmény legyen  $\bar{T}(t_{j+1})$ .

2. Hővezetési lépés :

$T(t_j) := \bar{T}(t_{j+1})$  adott mint kezdetivektor az  $\omega_h$  rácson, és a  $t_j \leq t \leq t_{j+1}$  intervallumon oldjuk meg a parabolikus

$$\frac{\partial T}{\partial t} = \operatorname{div}(k \operatorname{grad} T)$$

egyenletet. Az explicit séma  $\tau \leq h^2 / \max_{(x,t) \in Q_T} k(x,t)$  stabilitási korlátját elkerülve itt például a (16.236) faktorizált sémának valamelyik megoldási algoritmusát alkalmazzuk. Az eredmény legyen  $\bar{\bar{T}}(t_{j+1})$ .

3. A nemlineáris hőforrások figyelembevétele :

$T(t_j) := \bar{\bar{T}}(t_{j+1})$  adott mint kezdetivektor az  $\omega_h$  rácson, és a  $t_j \leq t \leq t_{j+1}$  intervallumon és az  $\omega_h$  rács minden pontjában oldjuk meg a közönséges

$$\frac{dT}{dt} = f(T)$$

differenciálegyenletet. A merevségtől függően válasszunk (nem túl magasrendű) Runge–Kutta vagy Gear- vagy Radau–módszert (ld. II. 10.5.). Az eredmény  $T(t_{j+1})$  a keresett hőeloszlás az új időszinten.

A fizikai jelenségek szerinti felbontáshoz ld. a fif-1-feladatot is.

#### 16.7.4 Többdimenziós parabolikus egyenletek többrácsos megoldása

Ahogy rámutattuk, a legcélszerűbb gyakran az, ha többrácsos módszerrel oldjuk meg a (16.205) alakú lineáris egyenletrendszert, amely a (16.200),  $c\rho = 1$  egyenlet súlyozott differencia-approximációjánál keletkezik (hasonlót kapunk a 16.4.11-ben tárgyalt sémából kiindulva is). Ennek a rendszernek a csak  $t$  szerint diszkrétizált megfelelője  $\sigma = 1$  esetén

$$\operatorname{div}(k \operatorname{grad} u)^{j+1} - \frac{c\rho}{\tau} u^{j+1} + \left( f^{j+1/2} + \frac{c\rho}{\tau} u^j \right) = 0, \quad x \in \Omega, \quad u|_{\Gamma} = g.$$

Ez az eredeti parabolikus differenciálegyenletre nézve a *Rothe-módszer* alkalmazását jelenti, amelynek segítségével elliptikus feladatok sorozatára vezetjük vissza. Hasonló elliptikus feladatokra jutunk  $0 < \sigma < 1$  esetén.

A 15.4. pontban részletesen megtárgyaltuk a többrácsos módszert, így itt csak néhány különbségre elég lesz felhívni a figyelmet.

A parabolikus feladatoknál két irányban kell haladnunk a többrácsos módszer keretében : finomabb rácsok és növekvő idő felé. Ekkor a (teljes) többrácsos módszer felgyorsítását eredményezi, ha (ld. a 15.4.3-ban közölt algoritmus 3. lépését)

$$y_\ell := \tilde{P}_{\ell-1}^\ell y_{\ell-1}$$

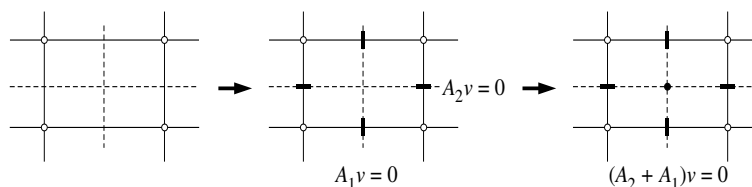
helyett, amikor az *új időszintre* emelkedünk, ott kezdeti közelítésnek minden új,  $\ell$ -edik rácsszinten az

$$y_\ell(t_{j+1}) := y_\ell(t_j) + \tilde{P}(y_{\ell-1}(t_{j+1}) - y_{\ell-1}(t_j))$$

képlet (azaz Euler–prediktor) segítségével számítjuk ki a nulladik közelítést. (Ezt  $j = 0$  esetén nem tehetjük, de nincs is szükség rá, mert ott adott a kezdetiérték).

Amennyiben a feladat olyan, hogy  $u$  beáll stacionárius értékre :  $u(x, t) \rightarrow u_\infty(x)$ , amikor  $t \rightarrow \infty$ , akkor a magas sorszámú szintekre csak ritkán kell felemelkedni, ami lényeges munkamegtakarítást jelent.

Azokban a speciális helyzetekben, amikor a váltakozó irányok módszerei alkalmatlanok (ha a  $k$  hővezetési együtthatónak szakadásai vagy erős gradiensei vannak, amelyek nem párhuzamosak a koordinátairányokkal) – kivéve



ábra 16.6: Mátrixfüggő interpoláció

az egydimenziós esetet, amikor az egzakt differenciasémából lehet kiindulni – javasolható a végelem approximáció, kombinálva újra a többrácsos módszerrel a diszkrét egyenletek megoldására. Ekkor célszerű olyan interpoláció használata, amely a mátrix elemeitől függ. Ezt, ha téglalapos rácson alapuló differencia-approximációról van szó, a következőképpen hozzuk létre (de az ötlet háromszöges végelem rácra általánosítható).

Legyen  $A_h := A_{1,h} + \dots + A_{d,h}$  az elliptikus operátornak megfelelő mátrix szétbontása az egyes koordinátairányok szerint (ld. (16.204)). Feladatunk a  $v_h$  finomrácsú javítás kiszámítása a durvarácsú  $y_H$  megoldás értékeiből. Ekkor először a javítás összes olyan értékét szokásos módon a finom rácra visszük át, amelynek pontja egyben a finom rácselem is. Ezután minden ( $i$ -vel jelölt) koordinátairányban a durvarácsú pontok közti  $(v_h)_\ell$  értékeket töltjük be az  $(A_{i,h}v_h)_\ell = 0$  lokális egyenletek megoldásával, felhasználva mint bal- és jobboldali peremértéket a már átvitt durvarácsú  $v$ -értékeket.

Ezután következnek azok a pontok, amelyek nem a durva rácsból valók, de a finom rácsnak egy  $i$ -irányú egyenesén fekszenek. Ekkor az  $((A_{i,h} + A_{j,h})v_h)_\ell = 0$  lokális egyenleteket oldjuk meg minden  $j \neq i$ -re, megint a már átvitt értékeket mint peremértékeket felhasználva.

A kétdimenziós esetben ezzel kész vagyunk (és a leírt interpoláció akkor, amikor a mátrix éppen (16.204), pontosan a bilineáris interpolációnak felel meg), háromdimenziós feladat esetén lezárjuk a folyamatot a  $0 = (A_h v)_\ell = ((A_{1,h} + A_{2,h} + A_{3,h})v)_\ell$  lokális egyenletek megoldásával azokra a finomrácsú pontokra, amelyek az elemi téglák közepén vannak.

A restriktio pedig legyen a leírt interpoláció adjungáltja.

## 16.8 Variációs eljárás

### 16.8.1 Szemidiszkretizáció

A parabolikus feladatok 16.3. megoldhatósági tétele és a végelem módszer elmélete egyaránt a (16.90) variációs megfogalmazásból indul ki (ld. (16.15)-öt is). Ezt a következőképpen kapjuk meg az  $n$ -dimenziós hővezetési egyenlet

esetén.

Legyen  $\Omega$  az  $\mathbb{R}^d$  véges, konvex tartománya és  $\Gamma$  annak poligonális pereme. Vizsgáljuk a

$$\frac{\partial u}{\partial t} = \sum_{i=1}^d \frac{\partial}{\partial x_i} \left( k_i(x) \frac{\partial u}{\partial x_i} \right) + f(x, t), \quad (x, t) \in Q_T := \Omega \times (0, T), \quad (16.242)$$

egyenletet homogén peremfeltételek mellett :

$$u(x, t) = 0, \quad t > 0, \quad x \in S_T := \Gamma \times (0, T); \quad (16.243)$$

a kezdeti feltétel legyen

$$u(x, 0) = u_0(x), \quad t = 0, \quad x \in \bar{\Omega}. \quad (16.244)$$

Feltesszük, hogy igaz

$$0 < k_0 \leq k_i(x) \leq K_0 \text{ minden } i\text{-re és minden } x \in \Omega\text{-ra.} \quad (16.245)$$

A (16.242) egyenletet megszorozzuk az  $\Omega$ -n definiált és folytonosan differenciálható  $v$  függvénnyel, a szorzatot integráljuk  $\Omega$  felett. Így kapjuk parciális integrálás útján (v.ö. 15.7.1-gyel) a

$$\begin{aligned} \left( \frac{\partial u}{\partial t}, v \right)_0 + a(u, v) &= (f, v)_0, \quad t > 0, \\ u(x, 0) &= u_0(x), \quad t = 0, \quad x \in \bar{\Omega} \end{aligned} \quad (16.246)$$

feladatot, ahol  $(u, v)_0$  az  $L_2$ -féle skalárszorzat és

$$a(u, v) := \int_{\Omega} \sum_{i=1}^d k_i(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} dx$$

bilineáris forma. Erre, (16.245)-nek köszönhetően, teljesül

$$m_a |u|_1^2 \leq a(u, u), \quad a(u, v) \leq M_a |u|_1 |v|_1, \quad (16.247)$$

ahol  $m_a := k_0$ ,  $M_a := K_0$ . Tehát az  $a$  forma  $H_0^1(\Omega) =: V$ -ben elliptikus és folytonos.

A (16.246) egyenlet akkor is értelmes, ha  $f(\cdot, t) \in H := L_2(\Omega)$  rögzített  $t$ -re, ha a  $v$  függvény és a rögzített  $t$ -re vizsgált  $u(\cdot, t)$  függvény  $V$ -ből valók.  $u(\cdot, t)$  normája

$$|u(\cdot, t)|_1^2 := \int_{\Omega} \sum_{i=1}^d \left( \frac{\partial u(x, t)}{\partial x_i} \right)^2 dx.$$

Az  $u(t) = u(\cdot, t)$ -t tekintjük mint  $(0, T) \rightarrow V$  leképzést és  $f(\cdot, t)$  mint  $(0, T) \rightarrow H$  leképzést, minden rögzített  $t$ -hez kapunk a  $V$ , ill.  $H$  térnek egy elemét. Mint  $t$  függvénye viszont pl.  $L_2$ -ből valók. Az ilyen függvények Hilbert-tereit  $L_2(0, T; V)$ -vel, ill.  $L_2(0, T; H)$ -val jelöljük (ehhez ld. részletesebben Zeidler könyveinek 23. fejezetét). Ezeknek normája

$$\|f\|_{L_2(0, T; H)} := \left( \int_0^T \|f(t)\|_0^2 dt \right)^{1/2}, \quad \|u\|_{L_2(0, T; V)} := \left( \int_0^T \|u(t)\|_1^2 dt \right)^{1/2}.$$

$L_2(0, T; H)$  helyett azt is írhatjuk, hogy  $L_2(\Omega \times (0, T))$ .

Ha az  $u$ -t mint a  $t$  ilyen függvényét tekintjük, akkor helyesebb (16.246)-ban  $\left(\frac{\partial u}{\partial t}, v\right)_0$  helyett  $\frac{d}{dt}(u, v)_0$ -t vagy  $(u', v)_0$ -t írni. A legáltalánosabb az, hogy itt  $u'$  alatt a  $V$ -n definiált funkcionált értjük. Ekkor  $(u', v)_0 = \frac{d}{dt}(u, v)_0$  ezen funkcionál értéke a  $v$  függvényen.

Mint ilyen általánosított derivált az  $u'$  nem kell, hogy minden  $t$ -re létezzen. Ennek megfelelően megengedjük, hogy a (16.246) variációs egyenlet nem teljesül egyes, nullmértékű halmazból vett  $t$ -értékekre.

Hasonlóan  $f$  is lehet a  $V$  téren definiált funkcionál. Mivel  $L_2(0, T; V)$  duális tere  $L_2(0, T; V')$ , így  $f, u' \in L_2(0, T; V')$ .

Ezután a  $H^1(0, T; V, H)$  teret vezethetjük be :

$$H^1(0, T; V, H) := \{u \in L_2(0, T; V), \quad u' \in L_2(0, T; V')\},$$

amelynek normája

$$\|u\|_{H^1} := \|u\|_{L_2(0, T; V)} + \|u'\|_{L_2(0, T; V')}.$$

$H$  azért szerepel a  $H^1$  tér jelölésében, mert  $V'$ -t a  $H = L_2$  tér skalárszorozata alapján definiáltuk. Így pl.

$$\|f\|_{V'} = \sup_{v \in V} \frac{(f, v)_0}{\|v\|_1}.$$

Lényeges, hogy ezen  $H^1(0, T; V, H)$  tér elemei  $(u(\cdot, t))$ -nek nullmértékű halmazon való esetleges megváltoztatása után) folytonosak  $t$ -ben mint  $H$ -ba való leképezés :

$$u \in H^1(0, T; V, H) \Rightarrow u \in C(0, T; H).$$

Ez a beágyazási tételek következménye és egyrészt azt jelenti, hogy a  $\|u(t)\|_0$  norma  $t$ -ben majdnem mindenütt folytonos és  $\|u(t)\|_0$  lényeges szupremuma véges, másrészt azt is vonja maga után, hogy így világos lesz, milyen értelemben írhatunk elő kezdeti feltételt :  $u_0$  lehet  $H$  eleme.

Ekkor a következő variációs feladatot (16.242)-(16.245) helyett fogalmazhatjuk meg.

Keressünk olyan  $u \in H^1(0, T; V, H)$  függvényt, hogy

$$(u', v)_0 + a(u, v) = (f, v)_0 \quad \text{minden } v \in V\text{-ra,} \quad (16.248)$$

$$u(0) = u_0 \in H. \quad (16.249)$$

Ezen feladat megoldását – ezt nevezzük a (16.242) *gyenge megoldásának* – jellemzi a következő tétel (ld. pl. Zeidler könyvét) :

**16.25. Tétel** (evolúciós egyenlet gyenge megoldása). Amennyiben az  $a$  forma bilineáris, folytonos,  $V$ -elliptikus :

$$a(u, v) \geq m_a |u|_1^2, \quad \text{minden } u \in V\text{-re,} \quad m_a = \text{const} > 0, \quad (16.250)$$

és ha  $u_0 \in H$ ,  $f \in V'$ , akkor a (16.248)-(16.249) feladatnak pontosan egy (gyenge) megoldása van, és ez stabil, azaz folytonosan függ a feladat  $u_0$ ,  $f$  adataitól : érvényes

$$\|u(t)\|_0^2 \leq \|u_0\|_0^2 + \frac{1}{2m_a} \int_0^t \|f(\eta)\|_{V'}^2 d\eta. \quad \square \quad (16.251)$$

**Megjegyzés.** Ez a becslés a következőképpen látható be : mivel minden rögzített  $t$ -re az  $u(t)$  gyenge megoldás  $V$  eleme, így behelyettesíthetjük (16.248)-ba. Ekkor figyelembe vehetjük, hogy

$$(u', u) = \frac{1}{2} \frac{d}{dt} \|u\|_0^2.$$

Ezután használjuk a (16.250)-et, valamint a Schwarz- és az  $\varepsilon$ -egyenlőtlenséget ( $\varepsilon = m_a$ -val) :

$$\frac{1}{2} \frac{d}{dt} \|u\|_0^2 + m_a |u|_1^2 \leq \|f\|_{V'} |u|_1 \leq m_a |u|_1^2 + \frac{1}{4m_a} \|f\|_{V'}^2. \quad (16.252)$$

Innen (16.251) adódik.  $\square$

Amennyiben  $f \in L_1(0, T; H)$ , akkor egy másik becslés vezethető le

$$(f, u)_0 \leq \|f\|_0 \|u\|_0 \quad \text{és} \quad \frac{1}{2} \frac{d}{dt} \|u\|_0^2 = \|u\|_0 \frac{d}{dt} \|u\|_0$$

alapján. Ekkor használhatjuk azt is, hogy  $V$  sűrűn és folytonosan be van ágyazva  $H$ -ba, tehát érvényes  $\|u\|_0 \leq c|u|_1$  és így következik (16.250)-ből

$$\beta \|u\|_0^2 \leq a(u, u), \quad \beta := m_a/c^2, \quad (16.253)$$

(de közvetlen becsléssel esetleg jobb konstans kapható). Ekkor (16.252) helyett

$$\|u\|_0 \frac{d}{dt} \|u\|_0 + \beta \|u\|_0^2 \leq \|f\|_0 \|u\|_0,$$

érvényes, és ezt osztva  $\|u\|_0$ -val, megszorozva  $e^{\beta t}$ -tel kapjuk, hogy

$$\begin{aligned} \frac{d}{dt} (e^{\beta t} \|u\|_0) &\leq \|f(t)\|_0 e^{\beta t}, \\ e^{\beta t} \|u(t)\|_0 - \|u(0)\|_0 &\leq \int_0^t e^{\beta \eta} \|f(\eta)\|_0 d\eta, \end{aligned}$$

vagyis

$$\|u(t)\|_0 \leq e^{-\beta t} \|u(0)\|_0 + \int_0^t e^{-\beta(t-\eta)} \|f(\eta)\|_0 d\eta. \quad (16.254)$$

A kezdetiérték befolyása tehát növekvő  $t$ -vel egyre csökken.

Egy harmadik stabilitási becslés vezethető le abban az esetben, ha  $u'(t) \in H^1(0, T; V)$ ,  $u(0) \in V$ .

Ekkor, (16.248)-ba  $v = u'$ -t behelyettesítve,

$$\|u'\|_0^2 + a(u, u') = (f, u')_0,$$

és innen (felhasználva  $a(u, v)$  alakját és azt a feltételünket, hogy a  $k_i$  együttműködőkkel együtt az  $a$  forma sem függ  $t$ -től) kapjuk, hogy

$$\|u'\|_0^2 + \frac{1}{2} \frac{d}{dt} a(u, u) \leq \|f\|_0 \|u'\|_0 \leq \|u'\|_0^2 + \frac{1}{4} \|f\|_0^2.$$

Így következik, hogy

$$a(u, u)(t) \leq a(u, u)(0) + \frac{1}{2} \int_0^t \|f(\eta)\|_0^2 d\eta,$$

vagyis

$$|u(t)|_1^2 \leq \frac{M_a}{m_a} |u(0)|_1^2 + \frac{1}{2m_a} \int_0^t \|f(\eta)\|_0^2 d\eta, \quad (16.255)$$

ahol az  $a$  forma ellipticitását és folytonosságát is felhasználtuk, ld. (16.247).

A  $V = H_0^1(\Omega)$  tér diszkrét altere legyen most  $V_h := \text{span}\{w_j\}_{j=1}^N$ , ahol  $\{w_j\}_{j=1}^N$  alkalmas,  $\Omega$ -n definiált végeselem bázis, ld. 15.7.4.

A diszkrétizált variációs feladat a következő, (16.248)-(16.249) helyett :



Keressünk olyan  $u_h \in H^1(0, T; V_h, H)$  függvényt, hogy

$$(u'_h, v_h)_0 + a(u_h, v_h) = (f, v_h)_0 \quad \text{minden } v_h \in V_h\text{-re,} \quad (16.256)$$

és hogy

$$u_h(0) = u_{h0} \in V_h.$$

Ezt a kezdeti feltételt pl. a végeselem interpolációval definiálhatjuk. Ekkor  $u_{h0}(x^\ell) = u_0(x^\ell)$  a trianguláció  $x^\ell$  belső csomópontjaiban,  $u_{h0}|_\Gamma = 0$ , és általában

$$u_{h0}(x) = \sum_{\ell=1}^N u_0(x^\ell) w_\ell(x). \quad (16.257)$$

(Más lehetőségekre \*\*\* 18.3.2-ben térünk ki.)

Szokásos módon elegendő a (16.256) egyenlőséget csak a  $V_h$  bázisára megkövetelni. Ekkor az  $u_h = u_h(t) := \sum_{\ell=1}^N y_\ell(t) w_\ell$  végeselem megoldás  $y = (y_1(t), \dots, y_N(t))^T$  együttható vektorára a

$$Gy' + Ay = \varphi, \quad y(0) = y^0, \quad (16.258)$$

feladatot kapjuk (ahol  $\varphi = \varphi(t) = ((f, w_1)_0, \dots, (f, w_N)_0)^T$  az  $f$  jobboldal Galjorkin-projekciója). Ez  $y$  számára egy implicit differenciálegyenlet!

Elsőnek vizsgáljuk ezután azt a kérdést, vajon mekkora az  $u \rightarrow u_h$  átmenettel kapcsolatos hiba?

Ennek becslésére az  $(u - u_h)(t) = u - u_h$  eltérést felbontjuk az  $a$  bilineáris forma segítségével definiált ortogonális  $V \rightarrow V_h$  projekció (a *Ritz-projekció*) felhasználásával :

$$\|u - u_h\|_0 \leq \|u - R_h u\|_0 + \|R_h u - u_h\|_0,$$

ahol az

$$R_h : V \rightarrow V_h$$

projekciót az

$$a(R_h u, v_h) = a(u, v_h) \quad \text{minden } v_h \in V_h\text{-re} \quad (16.259)$$

variációs feladat határozza meg. Itt tehát  $u$  a rögzített megoldásunk és  $R_h u$  keresett.

Elsőnek az  $u - R_h u$  eltérést fogjuk becsülni – ami az elliptikus feladatoknál látott módon megtehető. Az  $R_h u - u_h$  eltérés viszont  $V_h$  eleme, és ennek becslésekor a (16.256) variációs egyenlet lesz majd segítségünkre.

a) A Ritz-projekció definíciója alapján minden  $t \in [0, T]$ -re

$$a(u - R_h u, v_h) = 0 \quad \text{minden } v_h \in V_h\text{-ra,} \quad (16.260)$$

így

$$\begin{aligned} a(u - R_h u, u - R_h u) &= a(u - R_h u, u - v_h + (v_h - R_h u)) = \\ &= a(u - R_h u, u - v_h). \end{aligned}$$

Innen kapjuk (16.250) és (16.247) segítségével, a Céa-lemma bizonyítását ismételve,

$$|u - R_h u|_1 \leq \frac{M_a}{m_a} |u - v_h|_1.$$

Amennyiben  $u(t) \in H^r(\Omega)$ ,  $r > 1$ , úgy az  $u - v_h$  eltérés ( $v_h = \Pi_h u$  választása esetén) becsülhető  $c_0 h^{r-1} \|u\|_r$  segítségével, ld. a 15.31. tételt, azaz

$$|(u - R_h u)(t)|_1 \leq c_1 h^{r-1} \|u(t)\|_r, \quad c_1 := c_0 \frac{M_a}{m_a}, \quad (16.261)$$

minden  $t \in [0, T]$ -re.

Ha a Nitsche-fogás alkalmazható (v.ö. II. 11.6.3-mal és 15.7.3-mal), pl. a (16.242) egyenlet  $k_i$  együtthatóinak simasága és  $\Omega$  konvexitása esetén, akkor  $L_2$ -ben is kapunk becslést :

$$\|(u - R_h u)(t)\|_0 \leq c_2 h^r \|u(t)\|_r. \quad (16.262)$$

A lineáris végeelem bázis pl.  $r = 2$ -t enged meg, és a továbbiakban erre az esetre korlátozzuk a tárgyalást.

b) Most legyen  $w_h = w_h(t) := (R_h u - u_h)(t)$  és  $w = w(t) := (R_h u - u)(t)$ . Vezessük le (16.256) felhasználásával azt a variációs egyenletet, amelynek  $w_h$  tesz eleget!  $w_h$  definíciójából következik

$$\begin{aligned} \frac{d}{dt}(w_h, v_h)_0 + a(w_h, v_h) &= \frac{d}{dt}(R_h u, v_h)_0 + a(R_h u, v_h) \\ &\quad - \frac{d}{dt}(u_h, v_h)_0 - a(u_h, v_h) \\ &= \frac{d}{dt}(R_h u, v_h)_0 + a(R_h u, v_h) - (f, v_h)_0 \\ &= \frac{d}{dt}(R_h u, v_h)_0 + a(u, v_h) - (f, v_h)_0, \end{aligned}$$

ahol a Ritz-projekció (16.259) tulajdonságát is használtuk. Ezután viszont (16.248)-(16.249)-ből következik (mivel  $v_h \in V_h \subset V$ )

$$(w'_h, v_h)_0 + a(w_h, v_h) = \frac{d}{dt}(R_h u, v_h)_0 - \frac{d}{dt}(u, v_h)_0 = (w', v_h)_0. \quad (16.263)$$

Most hivatkozva arra, hogy  $w_h = w_h(t) \in V_h$ , választjuk  $v_h = w_h(t)$ -nek  $\frac{d}{dt}(w_h, v_h)_0 = (w'_h, v_h)_0$ -ban, és ekkor érvényes

$$(w'_h, w_h)_0 = \frac{1}{2} \frac{d}{dt} \|w_h\|_0^2 = \|w_h\|_0 \frac{d}{dt} \|w_h\|_0.$$

(16.263) és (16.250) alapján

$$\|w_h\|_0 \frac{d}{dt} \|w_h\|_0 + \beta \|w_h\|_0^2 \leq \|w'\|_0 \|w_h\|_0,$$

ld. (16.253). Innen kapjuk (ld. (16.253) után (16.254)-ig), hogy

$$\|w_h(t)\|_0 \leq e^{-\beta t} \|w_h(0)\|_0 + \int_0^t e^{-\beta(t-\eta)} \|w'(\eta)\|_0 d\eta. \quad (16.264)$$

Itt  $w_h(0) = u(0) - u_{h0} + (R_h u - u)(0)$ , tehát

$$\|w_h(0)\|_0 \leq \|u_0 - u_{h0}\|_0 + c_2 h^2 \|u_0\|_2, \quad (16.265)$$

ld. (16.262). Ha  $u' \in V$ , akkor differenciálhatjuk a (16.260) egyenletet  $t$  szerint, és mivel az  $a$  bilineáris forma nem függ  $t$ -től, azt kapjuk, hogy

$$(R_h u)' = R_h u' \in V_h,$$

valamint

$$\|w'\|_0 = \|u' - R_h u'\|_0 \leq c_2 h^2 \|u'\|_2. \quad (16.266)$$

Most foglalhatjuk össze eredményeinket, a (16.264)-(16.266) és (16.262) ( $r = 2$ -re vonatkozó) becsléseket, az utóbbit még a következőképpen folytatva :

$$\|u(t)\|_2 = \left\| u(0) + \int_0^t u'_t(\eta) d\eta \right\|_2 \leq \|u_0\|_2 + \int_0^t \|u'(\eta)\|_2 d\eta.$$

**16.26. Tétel** (szemidiszkretizáció hibabecslése). Legyen  $\Omega$  konvex, pereme poligonális, az  $a$  bilineáris forma  $V$ -elliptikus, folytonos és független  $t$ -től, és a (16.248)-(16.249) variációs feladat  $u$  megoldása legyen elég sima :  $u \in H^1(0, T; V, H)$ ,  $u'(t) \in V$  és  $r \geq 1$ -gyel  $u(t) \in H^r(\Omega)$  minden  $t$ -re. Ekkor igaz a következő hibabecslés :

$$\|u(t) - u_h(t)\|_0 \leq e^{-\beta t} \|u_0 - u_{h0}\|_0 + 2c_2 h^r \left\{ \|u_0\|_r + \int_0^t \|u'(\eta)\|_r d\eta \right\}. \quad (16.267)$$

Ehhez a becsléshez ld. a 45. feladatot is.

**Megjegyzések. 1.** Ha az  $u_{h0}$ -t Ritz-projekcióval számítjuk ki ( $u_0$ -ból), akkor (16.267)-ben az  $\|u_0 - u_{h0}\|_0$  tag nem is keletkezik. Amennyiben viszont  $u_{h0}$ -t (16.257)-nek megfelelően határozzuk meg, akkor a 15.31. tétel szerint  $\|u_0 - u_{h0}\|_0 \leq ch^2|u_0|_2$ .

**2.** Amilyen mértékben simább a (16.248)-(16.249) variációs feladat megoldása ( $u(t) \in H^r(\Omega)$ ,  $r \geq 2$ ) – v.ö. (16.267)-tel – olyan mértékben emelkedik a térbeli approximáció rendje, ha legalább  $(r - 1)$ -edfokú polinomokat használunk. Ezzel együtt viszont egyre több nemnulla elem jelenik meg a Gram-féle mátrixban, ami a (16.258) egyenlet megoldását megnehezíti és tulajdonságait megváltoztatja. A lineáris végelemek esetén ismert, hogy a pontosság csorbítása nélkül  $G$  helyett vehető

$$\tilde{G} := \text{diag}((Ge)_i), \quad e := (1, \dots, 1)^T$$

azaz (16.258)-ban helyettesítjük a Gram-féle mátrixot azzal a diagonális mátrixszal, amelynek főátlóján állnak  $G$  sorösszegei. Ehhez ld. a 16.4.8. pontot.

**3.** Mivel kvázi egyenletes trianguláción és a lineáris végelemekkel definiált  $V_h$  térben (ld. a 15.5.3. pontot is) igaz

$$\|v_h\|_{L_\infty} \leq c |\ln h|^{1/2} |v_h|_1,$$

így a fenti eredmények, valamint a 46. feladat megoldásának birtokában  $L_\infty$ -beli becsléseket is vezethetünk le, az előző becslést  $v_h := R_h u - u_h \in V_h$ -ra alkalmazva.  $\square$

## 16.8.2 Teljes diszkretizáció

Most a teljes diszkretizáció vizsgálatára térünk rá, amit (16.256)-ból úgy nyerünk, hogy a  $t_j := j\tau$  ekvidisztáns időpontokhoz ( $\tau > 0$  legyen az időlépés) az

$$u_h^j \approx u(t_j) =: u^j, \quad \frac{1}{\tau}(u_h^j - u_h^{j-1}) \approx u'(t_j), \quad f^j := f(t_j)$$

értékeket rendeljük hozzá.

A tisztán implicit módszerek (avagy *implicit Euler-Galjorkin módszer*-nek) megfelelően ezután az

$$\left( \frac{u_h^{j+1} - u_h^j}{\tau}, v_h \right)_0 + a(u_h^{j+1}, v_h) = (f^{j+1}, v_h)_0, \quad j = 0, 1, \dots, \quad (16.268)$$

$$u_h^0 = u_{h0},$$

egyenleteket írjuk fel. Ekkor legyen  $w_h^j := R_h u(t_j) - u_h^j$  úgy, hogy a teljes diszkretizáció hibája

$$u(t_j) - u_h^j = (u - R_h u)(t_j) + w_h^j.$$

A további gondolatmenet szorosan követi azt, amellyel a (16.267) becslést vezettük le. Először is megjegyezzük, hogy (16.262) továbbra is érvényes. Ezután vegyük észre, hogy  $w_h^j$ -re (definíciója szerint) teljesül

$$\begin{aligned} & \left( \frac{w_h^{j+1} - w_h^j}{\tau}, v_h \right)_0 + a(w_h^{j+1}, v_h) = \\ & \left( \frac{(R_h u)^{j+1} - (R_h u)^j}{\tau}, v_h \right)_0 - \left( \frac{u_h^{j+1} - u_h^j}{\tau}, v_h \right)_0 + a((R_h u)^{j+1}, v_h) - a(u_h^{j+1}, v_h), \end{aligned}$$

és itt

$$\left( \frac{u_h^{j+1} - u_h^j}{\tau}, v_h \right)_0 + a(u_h^{j+1}, v_h) = (f^{j+1}, v_h)_0 = (u'(t_{j+1}), v_h)_0 + a(u^{j+1}, v_h).$$

Így, mivel  $a((R_h u)^{j+1} - u^{j+1}, v_h) = 0$ ,

$$\left( \frac{w_h^{j+1} - w_h^j}{\tau}, v_h \right)_0 + a(w_h^{j+1}, v_h) = \left( \frac{(R_h u)^{j+1} - (R_h u)^j}{\tau} - u'(t_{j+1}), v_h \right)_0.$$

Ezt az egyenletet az

$$\left( \frac{w_h^{j+1} - w_h^j}{\tau}, v_h \right)_0 + a(w_h^{j+1}, v_h) = (r^j, v_h)_0 \quad (16.269)$$

alakban írjuk fel, ahol

$$\begin{aligned} r^j & := \frac{R_h u^{j+1} - R_h u^j}{\tau} - u'(t_{j+1}) \\ & = \left( \frac{R_h u^{j+1} - R_h u^j}{\tau} - \frac{u^{j+1} - u^j}{\tau} \right) + \left( \frac{u^{j+1} - u^j}{\tau} - u'(t_{j+1}) \right). \end{aligned}$$

Most (16.269)-be  $v_h = w_h^{j+1}$ -et helyettesítjük be. A bilineáris forma ekkor nemnegatív értéket vesz fel, tehát

$$\begin{aligned} \|r^j\|_0 \|w_h^{j+1}\|_0 & \geq \left( \frac{w_h^{j+1} - w_h^j}{\tau}, w_h^{j+1} \right)_0 \\ & = \frac{1}{\tau} \left\{ \|w_h^{j+1}\|_0^2 - (w_h^j, w_h^{j+1})_0 \right\} \geq \frac{1}{\tau} \left\{ \|w_h^{j+1}\|_0^2 - \|w_h^j\|_0 \|w_h^{j+1}\|_0 \right\}, \end{aligned}$$

vagyis

$$\|w_h^{j+1}\|_0 \leq \|w_h^j\|_0 + \tau \|r^j\|_0 \leq \|w_h^0\|_0 + \tau \sum_{k=0}^j \|r^k\|_0. \quad (16.270)$$

Itt a  $w_h^0 = w_h(0)$  becsléssel már (16.265)-ben foglalkoztunk.  $r^j$  becsléséhez felhasználjuk, hogy egyrészt

$$\begin{aligned} u^{k+1} - u^k - \tau u'(t_{k+1}) &= \int_{t_k}^{t_{k+1}} [u'(s) - u'(t_{k+1})] ds \\ &= - \int_{t_k}^{t_{k+1}} \int_s^{t_{j+1}} u''(\eta) d\eta ds \\ &= - \int_{t_k}^{t_{k+1}} \int_{t_k}^{\eta} u''(\eta) ds d\eta = \int_{t_k}^{t_{k+1}} (t_k - \eta) u''(\eta) d\eta, \end{aligned}$$

és ennek alapján

$$\left\| \frac{u^{k+1} - u^k}{\tau} - u'(t_{k+1}) \right\|_0 \leq \int_{t_k}^{t_{k+1}} \|u''(\eta)\|_0 d\eta. \quad (16.271)$$

Másrészt

$$\frac{1}{\tau} \left\{ (R_h u - u)^{k+1} - (R_h u - u)^k \right\} = \frac{1}{\tau} \int_{t_k}^{t_{k+1}} (R_h u - u)'(\eta) d\eta,$$

tehát

$$\sum_{k=0}^j \|r^k\|_0 \leq \int_0^{t_{j+1}} \|u''(\eta)\|_0 d\eta + \frac{1}{\tau} \int_0^{t_{j+1}} \|(R_h u - u)'(\eta)\|_0 d\eta. \quad (16.272)$$

Az előző (16.271) egyenlőtlenség és a (16.262), (16.265), (16.266), (16.270), (16.272) becslésekből (figyelembe véve, hogy (16.262) a  $t$  szerinti deriváltra is igaz, mert a bilineáris forma nem függ  $t$ -től, ld.  $R_h$  (16.259) definícióját) következik a kívánt eredmény.

**16.27. Tétel** (az implicit Euler–Galjorkin módszer hibabecslése). Ha a (16.248)–(16.249) variációs feladat megoldása elég sima abban az értelemben, hogy teljesül  $u \in H^1(0, T; V, H^2(\Omega))$ ,  $u'' \in L_1(0, T; H)$ , akkor az implicit Euler–Galjorkin módszer segítségével és lineáris végelemekkel előállított közelítő megoldás hibája másodrendű  $h$ -ban és elsőrendű  $\tau$ -ban. Igaz a következő egyenlőtlenség :

$$\begin{aligned} \|u(t_j) - u_h^j\|_0 &\leq \|u_0 - u_h^0\|_0 + 2c_2 h^2 \left\{ \|u_0\|_2 + \int_0^{t_j} \|u'(\eta)\|_2 d\eta \right\} + \\ &\quad + \tau \int_0^{t_j} \|u''(\eta)\|_0 d\eta. \quad \square \end{aligned}$$

Ehhez a becsléshez ld. az előző tételhez fűzött első megjegyzést is.

## 16.9 Összefoglalás

A parabolikus egyenletek megoldására többféle módszert mutatunk be. Csak a térbeli változókat diszkrétizálva, azaz *szemidiszkrétizációt* végrehatjva, az egyenletet közönséges differenciálegyenletek rendszerére vezethetjük vissza. A rendszer merevsége miatt A-stabil módszerek alkalmazása szükséges. A II-ben ismertetett magasabbrendű A-stabil módszerek itt különösen akkor javasolhatók, ha az egyenlet jobboldala nemlineáris, és ha a rendszer dimenziója nem túl nagy.

Teljes diszkrétizációból kiindulva részletesen vizsgáljuk az

$$\frac{y^{j+1} - y^j}{\tau} + A_h(\sigma y^{j+1} + (1 - \sigma)y^j) = \varphi^j$$

alakú *súlyozott kétréteges differenciasémát*, ahol  $\tau$  az időlépés,  $y^j$  a keresett megoldásértékek vektora a  $j$ -edik időrétegen és  $\sigma \in [0, 1]$  a súly, míg  $A_h$  az egyenlet elliptikus operátorának (differencia vagy végelem módszerrel levezetett) approximációja. Ezen módszer család legismertebb két speciális esete a Crank–Nicolson séma ( $\sigma = 1/2$ , a közönséges differenciálegyenletek szóhasználatával a trapézsabályról van szó), és a tisztán implicit séma ( $\sigma = 1$ , azaz implicit Euler–módszer).

A részletes stabilitási vizsgálat eredményeként belátjuk, ha  $A_h$  szimmetrikus és pozitív definit, hogy a súlyozott differenciaséma stabil, amikor

$$\sigma \geq \frac{1}{2} - \frac{1}{\tau \|A_h\|}, \quad \text{így pl. a } \sigma \geq \frac{1}{2}$$

esetben. Ekkor a séma stabil függetlenül a merevségtől (azaz  $\text{cond } A_h$ -tól). Ennek következménye, hogy a Crank–Nicolson séma másodrendben konvergens  $\tau$ -ban és  $h$ -ban, a tisztán implicit séma pedig elsőrendben  $\tau$ -ban, másodrendben  $h$ -ban.

Ezek az eredmények akkor is érvényesek, amikor az eredeti parabolikus egyenlet térben többdimenziós. Ilyenkor a diszkrétizált egyenletek általános megoldó algoritmusaként a *többrácsos módszer* javasolható, de sok esetben alkalmazható a *váltakozó irányok módszere* is, amely térben egydimenziós feladatokra vezeti vissza a többdimenziós feladatot, könnyen programozható és párhuzamosítható.

## 16.10 Feladatok

Az alábbi feladatok megoldásánál a „Hipy” program parabolikus része lehet segítségére.

1. feladat. Transzformáljuk a (16.2) egyenletet új változó bevezetésével a (16.3) alakra!

[Útmutatás. Az  $u(x, \tau) = e^{\alpha x + \beta \tau} v(x, \tau)$  összefüggésben válasszuk alkalmasan az  $\alpha$  és  $\beta$  paramétereiket.]

2. feladat. Bizonyítsuk be a 16.2. következményt a 16.1. tételhez!

[Útmutatás. Tekintsük  $u := u_1 - u_2$ -t és alkalmazzuk a 16.1. tételt!]

3. feladat. Tekintsük a II. 11.4.12. pont szerint megoldhatatlan elliptikus peremérték feladat kiterjesztését az időtől függő, parabolikus esetre :

$$\begin{aligned} t > 0, \quad 0 < x < 1 : \quad \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + \pi^2 u + \sin \pi x; \\ t = 0, \quad 0 \leq x \leq 1 : \quad u(x, 0) &= \sin \pi x; \\ t > 0 : \quad u(0, t) &= u(1, t) = 0. \end{aligned}$$

Keressük ezen vegyes feladat megoldását  $u(x, t) = c(t) \sin \pi x$  alakban! Mi történik  $t \rightarrow \infty$  esetén?

4\*. feladat. Hányadfokú hő-polinomokra egzakt a következő, (16.18)-nak megfelelő képlet a (16.6)-(16.9) feladat közelítő megoldására az  $x = \frac{1}{2}, t = 1$  pontban :

$$u\left(\frac{1}{2}, 1\right) = \frac{4}{9}(u(0, 1) + u(1, 1)) + \frac{1}{9}u\left(\frac{1}{2}, 0\right).$$

4. feladat. Hozzánk jön egy mérnök a következő, egy nedves lemez szárítását leíró egyenletrendszerrel :

$$\frac{\partial T}{\partial t} - \varepsilon \frac{\partial X}{\partial t} = a \frac{\partial^2 T}{\partial z^2}, \quad (16.273)$$

$$\frac{\partial X}{\partial t} = b \frac{\partial^2 T}{\partial z^2} + c \frac{\partial^2 X}{\partial z^2}, \quad 0 < z < L, \quad t > 0. \quad (16.274)$$

Itt  $T$  a lemez hőmérséklete,  $X$  a nedvességtartalma,  $t$  az idő és  $L$  a lemez vastagságának fele. A (16.273)-(16.274) rendszerhez kezdetiértékek is tartoznak  $T$  és  $X$  számára :

$$t = 0, \quad 0 \leq z \leq L : \quad T(z, 0) = T_0(z), \quad X(z, 0) = X_0(z), \quad (16.275)$$

valamint a következő peremfeltételek :

$$z = 0, \quad t > 0 : \quad \frac{\partial X}{\partial z} = 0, \quad \frac{\partial T}{\partial z} = 0, \quad (16.276)$$

$$z = L, \quad t > 0 : \quad \lambda \frac{\partial T}{\partial z} + aT = g_0(t), \quad \frac{\partial}{\partial z} (bT + cX) = g_1(t). \quad (16.277)$$



Az  $\varepsilon, a, b, c, \lambda$  mind pozitív konstansok. A feladat abból áll (mivel a fenti rendszer megoldhatóságáról vagy numerikus megoldásáról nem találunk irodalmat), hogy

a) transzfomáljuk a rendszert a következő alakra :

$$\frac{\partial y}{\partial t} = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} \frac{\partial^2 y}{\partial x^2}, \quad y := \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \quad (16.278)$$

Ehhez alkalmas a  $\begin{pmatrix} T \\ X \end{pmatrix} = Dy$  transzformáció, ahol  $D \in \mathbb{R}^{2 \times 2}$  konstans elemű mátrix.

b) Bizonyítsuk be, hogy (16.278)-ban  $k_1, k_2 > 0$  !

c) Mutassuk meg, hogy a transzformáció alatt a (16.276)-(16.277) peremfeltételek a következőkbe mennek át :

$$t > 0, \quad z = 0 : \quad \frac{\partial y}{\partial z} = 0, \quad (16.279)$$

$$t > 0, \quad z = L : \quad \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} \left( \frac{\partial y}{\partial z} + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} y \right) = f + \begin{pmatrix} 0 & \sigma_2 k_2 \\ \sigma_1 k_1 & 0 \end{pmatrix} y.$$

Itt  $\sigma_1, \sigma_2 > 0$ , és  $f = (f_1, f_2)^T$  az adott  $g_0$ -ból és  $g_1$ -ből számítható ki.

Az eredeti rendszernek hány lényeges paramétere van, ha  $T_0, X_0, g_0, g_1$  is konstansok és ha dimenziótlan alakra hozzuk a (16.273)-(16.277) relációkat? [Ehhez legyen pl.  $x := z/L$ ,  $\theta := T/T_0$ ,  $\xi := \varepsilon X/T_0$  és  $\tau := at/L^2$ .]

5. feladat. Bizonyítsuk be, hogy a (16.28) séma képlethibája  $O(h^6)$ , ha  $u$  elég sima, ha

$$\sigma = \frac{1}{2} - \frac{h^2}{12\tau} = \frac{3 - \sqrt{5}}{6}, \quad \tau = \frac{h^2}{2\sqrt{5}} \quad (16.280)$$

és a séma  $\varphi$  jobboldala megfelelő!

6. feladat. • A „Hipy” programban válasszuk az  $u(x, t) = 1 + x^3 \sin t$  pontos megoldáshoz a többi adatot, a peremfeltételek legyenek elsőfajúak.  $N = m = 10, 20, 40$ -nel futatva a programot, állítsunk össze táblázatot a kapott numerikus megoldás maximum normabeli pontosságáról a  $\sigma = 1/2, \sigma_*, 1$  esetekre, valamint az előző feladat speciális feltételeinek megfelelően. Hasonlítsuk össze eredményeinket a (16.38) ill. az 5. feladat szerint várható pontossággal!

7. feladat. • Futassuk a „Hipy” programot a beépített feladatával és  $N = 5, 10, 20, \dots, 320$ -t választva! Ismételjük a kísérletet  $\sigma = 0$ -val,  $\sigma = 1/2$ -del és  $\sigma = 1$ -gyel!

8. feladat. a) Írjuk fel az  $\dot{y}_h = -A_h^{(0)} y_h \in \mathbb{R}^{N-1}$  egyenletrendszerre az explicit Euler módszer képletét, ortogonális transzformációval kapjunk ebből  $N - 1$  skaláris egyenletet. Innen vezessük le az A-stabilitás feltételét!

b) Vizsgáljuk a  $\frac{3}{2}y_t - \frac{1}{2}y_{\bar{t}} + A_h^{(0)}y_h = 0$  séma stabilitását, aszimptotikus stabilitását!

c) Vizsgáljuk az  $y_t + \sigma\tau y_{\bar{t}} + A_h^{(0)}\hat{y}_h = \varphi_h$  séma stabilitását!

d) Vizsgáljuk az  $y_t^\circ + A(\sigma_1\hat{y}_h + (1 - \sigma_1 - \sigma_2)y_h + \sigma_2\check{y}_h) = \varphi_h$  séma stabilitását! ( $0 \leq \sigma_1, \sigma_2 \leq 1$ ,  $A_h = A_h^* > 0$ .)

9. feladat. • A „Hipy” program beépített feladatával, de  $T = 8$  és  $\sigma = 0$  esetén, növeljük az  $m$  időlépések számát 7-ről addig, míg az eredmények elfogadhatóknak nem válnak: a numerikus megoldás ne oszcilláljon a pontos megoldás körül a grafikonon. Hasonlítsuk össze a kapott időlépést a  $\tau \leq h^2/2$  stabilitási határral!

10. feladat. Kiindulva a (16.28)-(16.32) képletekből, valamint az I. 1.3.9-ben közölt rövidített Gauss-elimináció képleteiből, számítsuk ki pontosan a súlyozott differenciaséma műveletigényét, megkülönböztetve a  $\sigma = 0$ ,  $\sigma > 0$  esetén és az első és a többi időszinten adódó műveletigényt is!

11. feladat.(tra-2 \*\*\*) Írjuk fel a (16.278)-(16.279) egyenletrendszerre a tisztán implicit sémát és elemezzük a stabilitását, konvergenciáját (feltéve, hogy (16.278)-(16.279) megoldása létezik és elég sima)! A diszkrétizált egyenletek numerikus megoldására dolgozzunk ki hatékony eljárást, a rövidített Gauss-eliminációt megfelelően változtatva! Ekkor  $z = L$ -nél kapunk egy  $4 \times 4$ -es, konstans mátrixú lineáris rendszert, amelynek megoldása után beindíthatjuk a visszahelyettesítést. Mutassuk meg ezen mátrix regularitását (felhasználva ehhez az I.-beli 1.16. tételt)!

12. feladat (11.\*\*\*). Legyen  $u \in C^{6,3}$  a (16.63)-(16.66) feladat megoldása. Mutassuk meg, hogy  $\sigma = \sigma_*$  esetén, és ha  $\varphi$ -t a belső rácspontokban (16.37) szerint határozzuk meg (tehát amikor a képlethiba  $\psi_i = O(\tau^2 + h^4)$ ,  $1 \leq i \leq N - 1$ ), a (16.68)-(16.71) séma egyenletei a perempontokban úgy megváltoztathatók (a differenciacsillag kiterjesztése nélkül), hogy a képlethiba ott is (a  $\frac{h}{2}$ -vel való osztás előtt)  $O(\tau^2 + h^4)$ -rendű legyen!

13. feladat. Mutassuk meg, hogy igazak a (16.83) és (16.78) becslések, ahol  $A_h$  a (16.67)-tel definiált mátrix.

[Útmutatás. Ld. II-ben a 11.7. lemma bizonyítását, amely szerint

$$(A_h(\sigma_0, \sigma_1)v_h, v_h)_{(0,h)} \geq \frac{1}{\left(\frac{1}{2} + \frac{1}{\min(\sqrt{\sigma_0}, \sqrt{\sigma_1})}\right)^2} \|v_h\|_{C(\bar{\omega}_h)}^2.$$

A bizonyítást finomítva  $\sigma_0 + \sigma_1 > 0$  esetén, kapjuk (16.83)-ot, majd (16.78)-at is.]

14. feladat (12.\*\*\*). Bizonyítsuk be, hogy a (16.68)-(16.71) séma stabil a maximum normában, ha  $1 \geq \sigma \geq 1 - h^2/(2\tau(1 + h \max(\sigma_0, \sigma_1)))$ . Ezután vonjuk le a következtetést, hogy emellett a  $\sigma$ -feltétel mellett a séma pontossága

$O(h^2)$  a maximum normában. Tehát a 16.9'. következményhez tartozó megjegyzéssel és a 16.10. tétellel összehasonlítva eltűnt a  $h^{-1/2}$  szorzó!

[Útmutatás. A perempontokban szorozzuk meg  $\frac{h}{2}$ -vel a differencia-egyenleteket, majd használjuk a 16.14. tétel bizonyítását.]

15. feladat. a) A (16.91) összefüggés alapján bizonyítsuk be, hogy ha  $\|\cdot\|_0$  az  $L_2$ -norma és  $(\cdot, \cdot)$  az euklideszi skalárszorzat, akkor  $\|u_h(\cdot, t)\|_0^2 = (Gy(t), y(t))$  minden  $t \in [0, T]$ -re.

b) A (16.94) egyenlet  $y(t)$  megoldásvektora eleget tesz az

$$\|y(t)\|_G \leq \|y(0)\|_G + \int_0^t \|\varphi(\eta)\|_{G^{-1}} d\eta$$

becslésnek minden  $t \in [0, T]$ -re, ahol  $\|y(t)\|_G^2 := (Gy(t), y(t))$ .

c) Legyen minden rögzített  $t$ -re  $\psi_h(t) = \sum_{j=1}^{N-1} \psi_j(t) w_j$  az  $f(\cdot, t) \in L_2(0, 1)$  legjobb közelítése  $V_h$ -ban és  $\psi = \psi(t)$  a  $\psi_j(t)$  együtthatókból összeállított vektor. Ekkor

$$\|\psi(t)\|_G = \|\psi_h(t)\|_0 \leq \|f(\cdot, t)\|_0.$$

Mivel a (16.94)-beli  $\varphi$  és  $\psi$  között az az összefüggés, hogy  $G\psi = \varphi$ , így

$$\|\varphi\|_{G^{-1}} = \|\psi\|_G \leq \|f(\cdot, t)\|_0,$$

ami a b) szerinti stabilitási eredménnyel kombinálható.

16. feladat. Legyen  $u$  a (16.89) feladat megoldása és  $\vec{u}_h^j$  az a vektor, amely az  $u$  megoldás  $\omega_h$ -n felvett értékeit tartalmazza. Taylor-sorfejtés segítségével ellenőrizzük, hogy a (16.96), (16.95) eljárás

$$\psi_i := (G + \sigma\tau A)\vec{u}_i + A\vec{u}^j - \varphi^{j+\sigma}$$

képlethibája eleget tesz a következő relációknak, ha  $\varphi_i^{j+\sigma} = f_i^{j+\sigma}$  :

$$\psi_i^j = \begin{cases} O(\tau + h^2), & \text{ha } u \in C^{4,2}, \\ O(\tau|\sigma - \frac{1}{2}| + \tau^2 + h^2), & \text{ha } u \in C^{4,3}. \end{cases}$$

Mutassuk meg azt is, hogy itt is van magasabbrendű eljárás, de a  $\sigma$ -érték nem (16.37), mint a (16.28) séma esetén : Érvényes  $\psi_i^j = O(\tau^2 + h^4)$ , ha  $u \in C^{6,3}$  és  $\sigma = \frac{1}{2} + \frac{h^2}{12\tau}$ , viszont  $\varphi_i^{j+\sigma}$  helyett vesszük a (16.95) utolsó képletét a  $t = t_{j+1/2}$  időpontban.

17. feladat. Az előző feladat és a 16.11 tétel feltételei mellett lássuk be, hogy a (16.96) végeelem séma konvergens abban az értelemben, hogy  $\|\vec{u}_h - y\| \rightarrow 0$ , ha  $\tau, h \rightarrow 0$  és ha a  $\|\cdot\|$  norma a diszkrét  $L_2$ - vagy maximum norma! Az  $u_h$  (16.91) képletének segítségével mutassuk meg, hogy az előző

eredményekből az is következik, hogy  $\|u_h(\cdot, t) - u(\cdot, t)\| \rightarrow 0$ , amikor a norma az  $L_2(0, 1)$ - vagy a  $C[0, 1]$ -norma és  $t \in (0, T]$  tetszőleges.

18. feladat. • Kísérletezzünk a Crank–Nicolson módszerrel (és a  $\tau = h$  lépéstávolsággal) abban az esetben, amikor a kezdeti függvény a (16.97) képlettel megadott. Ezt a következőképpen valósíthatjuk meg a „Hipy” program keretében :  $u_0(x) = 0.5 * (1 - (x - 0.499)/|x - 0.499|)$ .

Figyeljük meg az első néhány időlépés grafikonjait  $h = 1/10, 1/20, 1/40$  esetén!

19. feladat. Bizonyítsuk be, hogy a (16.27) szemidiszkretizált eljárás pozitívástartó, de (16.94) általában nem!

[Útmutatás. Használjuk az II-beli 10.4. lemmát és a 10.5. tételt. A (16.94) (végeelem módszerrel szemidiszkretizált) eljárással kapcsolatban lássuk be, hogy  $G^{-1}$  olyan telt mátrix, amelynek elemei váltakozva pozitívak és negatívak sakktáblaszerű elrendezésben. Ebből következtessük, hogy  $G^{-1}A$  elemeinek előjelei is sakktáblaszerűen rendeződnek el.]

20. feladat. • Használjuk a „Hipy” programot a következő feladat megoldására! Legyen (a program jelöléseivel)  $u_0(x) = (2x - 1)/|2x - 1| - (17x - 9)/|17x - 9|$ ,  $m = 1$ ,  $N = 41$ . A peremfeltételek legyenek homogének, elsőfajúak, a jobboldal  $f = 0$ . Kezdve  $T = 1/80$ -tól variáljuk az  $1/T$  egész számot addig, míg a megoldás már nem tartalmaz negatív értéket! Mi köze van a kapott értéknek  $N^2$ -hez és a 16.13. tételhez?

21. feladat. Követve a 16.14. tétel első megjegyzésének a gondolatmenetét vezessük le, milyen becslés adódik a maximum normában akkor, ha a (16.105) majoráns függvény helyett  $w(t) = 1 + t-t$  vesszük !

22. feladat. • Végezzünk numerikus kísérleteket a 16.14. tétel 4. becslésében előforduló konstans meghatározására!

23. feladat. Számítsuk ki, hogy mi adódik (16.94) helyett, ha a (16.92) szemidiszkretizált egyenlet integráljait a trapézszabály segítségével közelítjük!

24. feladat. a) Bizonyítsuk be, hogy nincs olyan, a  $T_6$  elem összes polinomját pontosan integráló kvadratura képlet, amelynek alappontjai az elem hat csomópontja és súlyai pozitívak!

b) Számítsuk ki az  $R_8$  elemhez a közelítő, lokális tömegmátrixot, ha a numerikus integrálás a 15.7.6. pont (15.239) képlet segítségével történik! Blokkdiagonális lesz-e ez a tömegmátrix?

[Útmutatás. Használjuk az  $R_8$  elem bázisfüggvényeinek 15.7.4-ben közölt (15.213) képleteit!]

25. feladat. Számítsuk ki analitikusan az  $R_8$  elem lokális tömegmátrixát, majd ebből készítsünk úgy diagonális mátrixot, hogy a főátlóra a sorössze-

geket írjuk. Pozitív definit-e a kapott mátrix? Milyen lesz a globális tömegmátrix és az abból hasonlóan elkészített diagonális mátrix?

26. feladat. • Határozzuk meg kísérletileg az optimális  $\sigma$ -t a következő feladat súlyozott differenciasémával való numerikus megoldásához:

$$\begin{aligned}\frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < 1, \quad 0 < t \leq T = 0.02, \\ u_0(x) &= \begin{cases} 0, & 0 \leq x < \frac{1}{2}, \\ 1, & \frac{1}{2} \leq x \leq 1. \end{cases} \\ u(0, t) &= 0, \quad u(1, t) = 1, \quad 0 < t \leq T.\end{aligned}$$

A „Hipy” programban állítsuk be a következő értékeket :  $m = 5$  (időlépések száma),  $N = 250$ , azaz ekkor  $\tau = h$ , legyen  $\sigma = 0.4$  (0.1) 1. A kezdeti függvény miatt v.ö. a 18. feladattal. A pontos megoldás néhány értéke :  $u(0.5, t) = 0.5$ ,  $u(0.75, t) = 1 - u(0.25, t)$ , és

t	0.005	0.01	0.02	0.03	0.04	0.05
u(0.25,t)	0.00621	0.03855	0.10556	0.15262	0.18438	0.20578

27. feladat. Elemezzük a (16.96) végeselem séma  $\sigma = 1$  esetét az aszimptotikus stabilitás szempontjából!

28. feladat. • Futassuk a „Hipy” programot a homogén ( $f \equiv 0$ ) (16.6) hővezetési egyenlet  $u(x, t) = e^{-\pi^2 t} \sin \pi x$  pontos megoldásával, homogén elsőfajú peremfeltételekkel és  $h = 1/20$ ,  $\sigma = 1/2$  esetén! Határozzuk meg azt a  $\tau \in [0.005, 0.05]$  értéket, amely mellett  $T = 1$ -re a legkisebb maximum normabeli hibát kapjuk! Hasonlítsuk össze ezt az értéket a 16.15. tételben szereplő  $\bar{\tau}$ -tal, valamint a tételhez fűzött 2. megjegyzés szerinti  $\tau^*$ -gal is!

29. feladat. Bizonyítsuk be a maximum normában a (16.120) ill. (16.121) mátrixokkal felírt (16.96) séma jobboldal, kezdeti- és peremértékek szerinti stabilitását, ha  $\sigma \in [0, 1]$  alkalmas feltételnek tesz eleget!

[Útmutatás. Kövessük a 16.14. tételhez fűzött első megjegyzésének a gondolatmenetét!]

30. feladat. Bizonyítsuk be, hogy a háromréteges (16.126) séma, alkalmazva az  $\dot{u} + \lambda u = 0$  közönséges differenciálegyenletre, akkor másodrendű, amikor  $\theta = 1 - \sqrt{2}/2$ , és erősen A-stabil (ld. II. 106. o.), ha ezenkívül még  $\alpha \in (\frac{1}{2}, 1]$  !

31. feladat. Tekintsük a következő kétréteges sémát az  $\dot{u} + Au = f(t)$  egyenletrendszer megoldására (ahol  $y^{j+\kappa}$  az  $u((j + \kappa)\tau)$  vektor approximációja) :

$$(I + \alpha\tau A)y^{j+1/2} = (I - \beta\tau A)y^j + \delta_1\tau f^{j+1/2}, \quad (16.281)$$

$$(I + \alpha\tau A)y^{j+1} = (I - \gamma\tau A)y^{j+1/2} + \delta_2\tau f^{j+1/2}. \quad (16.282)$$

Válasszuk ki  $\alpha$  függvényében a többi paramétert úgy, hogy valóságosak legyenek és az eljárás másodrendű!

Bizonyítsuk be ezután azt, hogy ezek mellett a paraméterek mellett a séma

- a) A-stabil pontosan akkor, amikor  $\alpha \geq 1/4$ ;
- b) erősen A-stabil, amikor  $\alpha > 1/4$  és  $\alpha \neq 1/2$ ;
- c) L-stabil, ha  $\alpha = 1 \pm \frac{\sqrt{2}}{2}$  !

Mi a séma hátránya akkor, amikor  $1 - \frac{\sqrt{2}}{2} < \alpha < 1 + \frac{\sqrt{2}}{2}$  ?

Vizsgáljuk azt is, hogy milyen feltételek mellett pozitivitástartó a másodrendű (16.281)-(16.282) séma, ha

- a)  $A = \lambda I$ ,  $\lambda > 0$  és  $f \equiv 0$ , de  $\tau > 0$  tetszőleges;
- b)  $A = \frac{1}{h^2}$  tridiag $(-1, 2, -1)$ ,  $\tau > 0$  elégedően kicsi, de  $f \geq 0$  tetszőleges!

Végülis, harmadrendű milyen paraméter értékekre a (16.281)-(16.282) séma?

32. feladat. Bizonyítsuk be, hogy a (16.129), (16.131), (16.29), (16.30) sémára igaz a maximumelv, ha teljesülnek a (16.133) feltételek!

33. feladat. a) Legyen  $A_h = A_h^{(1)}$  a (16.143), (16.146) mátrix,  $(\cdot, \cdot)_{(r,0)}$  a (16.145) skalárszorzat és  $\|\cdot\|_{(r,0)}$  a hozzátartozó norma.  $(A_h y_h, y_h)_{(r,0)}$  explicit alakjából vezessük le azt, hogy

$$(A_h y_h, y_h)_{(r,0)} \leq \frac{4}{h^2} \|y_h\|_{(r,0)}^2,$$

felhasználva az  $r_{1/2} = 2r_0$  relációt is!

b) Próbáljuk meg az  $\|A_h\|_{(r,0)}$  becslését a következőképpen, a  $(\cdot, \cdot)_{(0,h)}$   $L_2(\omega_h)$ -skalárszorzatot és  $\|\cdot\|_{(0,h)}$  hozzátartozó normát is használva, az  $R_h := \text{diag}(r_i)$  mátrixot és a  $v_h = R_h^{1/2} y_h$  vektort bevezetve :

$$\begin{aligned} (A_h y_h, y_h)_{(r,0)} &= (R_h A_h y_h, y_h)_{(0,h)} = (R_h^{1/2} A_h R_h^{-1/2} v_h, v_h)_{(0,h)} \\ &\leq \rho(R_h^{1/2} A_h R_h^{-1/2}) \|v_h\|_{(0,h)}^2 = \rho(A_h) \|v_h\|_{(0,h)}^2 \\ &= \rho(A_h) (R_h y_h, y_h)_{(0,h)} = \rho(A_h) \|y_h\|_{(r,0)}^2 \leq \|A_h\|_{(\infty)} \|y_h\|_{(r,0)}^2, \end{aligned}$$

tehát  $\|A_h\|_r \leq \|A_h\|_{\infty}$  (és  $\|A_h\|_{\infty} = \frac{4}{h^2}$ ). Melyik lépés nem volt indokolt?

34. feladat. a) Vizsgáljuk az  $y_N = a$ ,  $A_r y_h + f = 0$ ,  $x \in \omega_h^{(\rho)}$  séma approximációját az  $\mathcal{L}u + f = 0$ ,  $0 < r \leq 1$ ,  $u'(0) = 0$ ,  $u(1) = a$  egyenlet  $C^6[0, 1]$ -beli megoldásain! Itt  $\mathcal{L}$  a (16.140) operátor,

$$\begin{aligned} (A_r y_h)_i &:= - \left( \frac{1}{r} ((\bar{r} y_r)_{\bar{r}}) \right)_i = - \frac{1}{r_i h} ((\bar{r} y_r)_{i+1/2} - (\bar{r} y_r)_{i-1/2}), \\ \bar{r}_{i+1/2} &:= \left( r - \frac{h^2}{12r} \right)_{i+1/2}, \end{aligned}$$

és  $\omega_h^{(\rho)}$  a (16.142) rács.

b) Határozzuk meg a rács  $\rho$  konstansát úgy, hogy

$$(A_r y_h)_0 = -\frac{1}{r_0 h} (\bar{T}_{1/2} y_{r,0})$$

ugyanazzal a negyedrendű approximációval rendelkezzen, mint  $(A_r y_h)_i$ ,  $i \geq 1$ !

35. feladat. Lássuk be, hogy teljesülnek a (16.174) relációk és (16.175) becslések, ha  $y, v$  tetszőleges, a (16.142),  $\rho = 1$  rácson definiált olyan függvények, amelyeknek  $r_N$ -beli értéke nulla.

36. feladat. (16.165) alapján mutassuk meg, hogy a (16.171) séma pontossága  $O(\tau^2 + h^4)$ , amennyiben a (16.161)-(16.163) feladat  $u$  megoldása  $C^{6,3}([0, 1] \times [0, T])$ -ből való és

$$\varphi = \left( f + \frac{h^2}{12} f_{\bar{r}r} \right)^{j+1/2}, \quad \sigma = \frac{1}{2} - \frac{h^2}{12\tau}.$$

37. feladat. a) Mutassuk meg, hogy igaz

$$\|y_h\|_{C(\omega_h^{(1)})} \leq h^{-1/2} \|y_h\|_{(r^2, 1)},$$

ha az  $y$  rácsfüggvény eltűnik  $r = r_N$ -nél! Ennek alapján következtessünk a (16.171) séma maximum normabeli konvergenciájára!

b) Bizonyítsuk be közvetlenül a (16.171),  $\sigma = 1$  séma stabilitását a maximum normában!

[Útmutatás. Mutassuk meg, hogy a (16.164)-gyel, (16.170)-nel definiált mátrix M-típusú. Alkalmassal használjuk az  $c_0(1-r)$  rácsfüggvényt majoráns függvényként a (16.171) feladat  $y$  megoldás becslésére!]

38. feladat. Felhasználva a (16.179) skalárszorzatokat, bizonyítsuk be az összes a 16.5.4. pontban, a (16.178) harmadfajú peremfeltétellel kapcsolatban kimondott állítást!

39. feladat. Milyen következtetést (a séma stabilitására és konvergenciájára nézve) lehet levonni a Szamarszkij-féle stabilitási eredményekből abban az esetben, ha a (16.28) súlyozott sémát változó lépéstávolsággal alkalmazzuk?

40. feladat. a) Írjuk ki a (16.197)-(16.199) differencia-egyenleteket részletesen, programozásra alkalmas alakban, jobbra rendezve az ismert mennyiségeket, balra az ismeretleneket.

b) A 18. fejezet \*\*\* 18.3-18.5. tételeit alkalmazzuk a sémára és írjuk fel az adódó becsléseket!

41. feladat. a) Legyen  $\omega_h$  a  $h_1, h_2$  lépéstávolságú ekvidisztáns rács a  $\{0 \leq x_1 \leq a, 0 \leq x_2 \leq b\}$  téglalapban és (16.227) a (16.225) differencia-approximációja. Mutassuk meg, hogy érvényes (16.224), ha (16.226) teljesül!
- b) Lássuk be, hogy a (16.224) tulajdonság a megfelelően megváltoztatott (16.225) feltétel és (16.226) mellett továbbra is érvényes, ha

$$\mathcal{L}_\ell u = \frac{\partial}{\partial x_\ell} \left( k_\ell \frac{\partial u}{\partial x_\ell} \right) - q_\ell u, \quad \ell = 1, 2.$$

c) Arra az esetre, hogy a tartomány L-alakú és a rácsnak 5 belső pontja van (ld. a 15.36. ábrát 15.8.5-ben), írjuk fel a Poisson-egyenletnek megfelelő  $A_1, A_2$  mátrixokat és lássuk be, hogy (16.224) nem teljesül!

42. feladat. a) A (16.206), (16.207) egyenletekből zárjuk ki az  $y^{j+1/2}$  közbülső megoldást és hozzuk a (16.236),  $d = 2$  faktorizált séma alakjára. Ehhez  $y^{j+1/2}$  számára milyen peremfeltételt kell választani a  $\gamma_h^{(1)}$  peremen?

b) Ezután, homogén peremfeltételek és a (16.224) feltételek mellett, alakítsuk át a kapott (16.236) sémát a (16.45) kétréteges sémába, és bizonyítsuk be, hogy ez stabil a kezdetiértékek és a jobboldal szerint a  $\Lambda_1, \Lambda_2$  operátorokhoz tartozó  $A_h = A_{1,h} + A_{2,h}$  mátrixszal definiált  $\|\cdot\|_A$ -normában!

[Útmutatás. Alkalmazzuk a 16.5-16.7 tételeket!]

43. feladat. Tekintsük a (16.210)-(16.212),  $z^{j+1/2} = 0, x \in \gamma^{(1)}$  feladatot abban az esetben, amikor  $\Lambda_1, \Lambda_2$  az időtől is függnék. Legyen

$$\Lambda_1 z^{j+1/2} = \Lambda_1(t_{j+1/2}) z^{j+1/2}, \quad \Lambda_2 z^j = \Lambda_2(t_j) z^j, \quad \Lambda_2 z^{j+1} = \Lambda_2(t_{j+1}) z^{j+1}.$$

Tegyük fel, hogy ezek az  $\Lambda_1, \Lambda_2$  operátorok mind valóban pozitív definiték. Bizonyítsuk be a következő stabilitási becslést :

$$\|z^j\|_{(2,t_j)} \leq \|z^0\|_{(2,t_0)} + \sum_{k=0}^{j-1} \tau \left( \|\psi^{k+1/2}\|_{(0,h)} + \|\psi^{k+1}\|_{(0,h)} \right),$$

ahol

$$\|z\|_{(2,t_j)} := \left\| \left( I + \frac{\tau}{2} \Lambda_2(t_j) \right) z \right\|_{(0,h)} \geq \|z\|_{(0,h)}.$$

[Útmutatás. Használjuk a következő lemmát! (Ld. Szamarszkij 1971-es könyvét, 366. o.) Ha  $\Lambda$  valóban pozitív definit,  $\tau > 0$  és  $\sigma \geq \frac{1}{2}$ , akkor érvényes

$$\|(I - (1 - \sigma)\tau\Lambda)v\|_{(0,h)} \leq \|(I + \sigma\tau\Lambda)v\|_{(0,h)}.$$

44. feladat. Mint az 5. feladat általánosítását tekintsük a (16.200), (16.201) feladatot abban az esetben, amikor  $\mathcal{L}$  a Laplace-operátor, a rács



négyszetes és a (16.236) differenciaséma paramétereire érvényes (16.280). Ekkor vezessünk le olyan kifejezést  $\varphi$ -re, hogy a (16.200), (16.201) feladat megoldásának megfelelő simasága esetén a séma  $O(h^6)$ -rendű legyen!

45. feladat. A 16.26. tétel (16.267) hibabecslésénél (16.254)-ből indultunk ki, de feltételeink  $V$ -ben is engednek meg egy becslést. Hozzuk létre ezt, a (16.255)-höz vezető gondolatmenetet (16.263)-ra alkalmazva.

[Kommentár. Ekkor (16.262) helyett – lineáris végeselemek és  $r = 2$  esetén – hivatkozhatunk (16.261)-re, tehát elsőrendű a konvergencia.]

46. feladat. a) Bizonyítsuk be, hogy a Crank–Nicolson–Galjorkin módszer esetén, tehát

$$\left( \frac{u_h^{j+1} - u_h^j}{\tau}, v_h \right)_0 + a \left( \frac{u_h^{j+1} + u_h^j}{2}, v_h \right) = (f^{j+1/2}, v_h)_0 \quad j = 0, 1, \dots,$$

$$u_h^0 = u_{h0}$$

esetén van olyan  $C$  konstans, amellyel igaz a következő becslés :

$$\|u_h^j - u(t_j)\|_0 \leq \|u_0 - u_h^0\|_0 + C \left\{ h^2 \left[ \|u_0\|_2 + \int_0^{t_j} \|u'\|_2 \, d\eta \right] + \tau^2 \int_0^{t_j} [\|u'''\|_0 + \|\Delta u''\|_0] \, d\eta \right\}.$$

b) (16.268) helyett írjuk fel azt a variációs feladatot, amely (16.248)-(16.249)-ből keletkezik az időderivált másodrendű retrográd differencia-approximációjával, majd vizsgáljuk ezen módszer stabilitását!

47. feladat. Javasoljunk megoldási módszert a következő feladathoz : megoldandó a kétdimenziós diffúzió-konvekció egyenlet egy olyan tartomány esetén, amely téglalaphból áll, valamint annak egyik sarkából induló vonalból, és ezen a vonalon ugyancsak megoldandó a diffúzió-konvekció egyenlet!

[Kommentár. A feladatnak nyilvánvaló környezetvédelmi háttere van : számítandó a szennyeződés egy folyón a hozzátartozó tóval együtt. A főprobléma – az approximáció és a megoldási algoritmus szempontjából – a csatlakozási pont kezelése.]



# Fejezet 17

## A Navier–Stokes egyenletek numerikus megoldása

### 17.1 Bevezetés: fizikai háttér

A Navier–Stokes egyenletek az impulzus megmaradását egy áramló közegben írják le :

$$\rho \left( \frac{\partial u_i}{\partial t} + \sum_{j=1}^d u_j \frac{\partial u_i}{\partial x_j} \right) = (\operatorname{div} \sigma)_i + \rho f_i, \quad i = 1, \dots, d, \quad (17.1)$$

ahol  $\vec{u} = (u_1, \dots, u_d)^T$  a sebesség vektora,  $d$  a vizsgált tartomány dimenziója ( $d = 2, 3$ ),  $\rho$  a sűrűség (amit pozitív konstansnak tételezünk fel),  $\rho \vec{f} = (f_1, \dots, f_d)^T$  a térfogati erők sűrűsége és  $\sigma$  a feszültségi tenzor :

$$\sigma = -P \cdot I + 2\mu D(\vec{u}). \quad (17.2)$$

Itt  $P$  a nyomás,  $\mu = \text{const} > 0$  a viszkozitás és

$$D_{ij}(\vec{u}) = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right).$$

Konstans  $\rho$  esetén szokás a következő „kinematikus” mennyiségeket bevezetni  $P$  és  $\mu$  helyett :

$$p := P/\rho, \quad \nu := \mu/\rho. \quad (17.3)$$

Kiszámítva  $\operatorname{div} \sigma$ -t azt kapjuk, hogy

$$\frac{1}{\rho} \operatorname{div} \sigma = -\operatorname{grad} p + \nu \Delta \vec{u},$$

ugyanis

$$\sum_{j=1}^d \frac{\partial}{\partial x_j} \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \frac{1}{2} \sum_{j=1}^d \left( \frac{\partial^2 u_i}{\partial x_j^2} + \frac{\partial^2 u_j}{\partial x_i \partial x_j} \right) = \frac{1}{2} \Delta u_i, \quad (17.4)$$

ahol felhasználtuk azt is, hogy a  $\rho = \text{const}$  feltevésnek megfelelően

$$\text{div } \vec{u} = 0. \quad (17.5)$$

Ez az egyenlet a tömegmegmaradást írja le. Teljes alakja

$$\frac{\partial \rho}{\partial t} + \text{div}(\rho \vec{u}) = 0, \quad (17.6)$$

amiből a konstans sűrűség miatt közvetlenül adódik (17.5). (17.5)-ből egyébként rögtön látható, hogy nem érdekes a  $d = 1$  eset. Ekkor ugyanis csak egy helykoordináta és egy sebességi komponens lenne és utóbbinak konstansnak kellene lennie.

Figyelembe véve a (17.1), (17.4), (17.5) egyenleteket, megkapjuk a teljes “összenyomhatatlan” Navier–Stokes-féle egyenletrendszert :

$$\frac{\partial \vec{u}}{\partial t} + \sum_{j=1}^d u_j \frac{\partial \vec{u}}{\partial x_j} + \text{grad } p = \nu \Delta \vec{u} + \vec{f}, \quad (17.7)$$

$$\text{div } \vec{u} = 0. \quad (17.8)$$

Ezen rendszernek a „divergens” alakja – amely a véges térfogat módszerrel történő numerikus megoldás kiindulási pontja – úgy adódik (17.7)-(17.8)-ból, hogy a  $\sum_{j=1}^d u_j \frac{\partial \vec{u}}{\partial x_j}$  részt a  $\text{div } \vec{u} = 0$  feltétel segítségével átírjuk :

$$\begin{aligned} \sum_{j=1}^d u_j \frac{\partial u_i}{\partial x_j} &= \sum_{j=1}^d u_j \frac{\partial u_i}{\partial x_j} + u_i \sum_{j=1}^d \frac{\partial u_j}{\partial x_j} \\ &= \sum_{j=1}^d \frac{\partial (u_i u_j)}{\partial x_j} = \text{div}(u_i \vec{u}) =: \mathbf{D}(\vec{u})u_i, \end{aligned}$$

tehát

$$\frac{\partial \vec{u}}{\partial t} + \mathbf{D}(\vec{u})\vec{u} + \text{grad } p = \nu \Delta \vec{u} + \vec{f}, \quad (17.9)$$

$$\text{div } \vec{u} = 0. \quad (17.10)$$

Ahhoz, hogy ezek az egyenletek jól írják le az áramló közeg állapotát, nemcsak az kell, hogy a közeg összenyomhatatlan legyen (amely feltétel víz esetén

igen jól teljesül : még 10 km-es mélységben sűrűsége is közel van 1-hez. Levegő esetén elég jó az összenyomhatatlanság feltevése, míg az áramlási sebesség a hangsebesség ötödénél nem nagyobb). Hanem az is kell, hogy a közeg „Newton-féle folyadék” legyen. Ezt (17.2) fejezi ki egyszerűsített formában :  $\sigma$  és  $D(\vec{u})$  között lineáris összefüggés létezik ilyenkor. A „nem-Newtoni” folyadékot gyakran úgy jellemzik, hogy  $\mu\Delta\vec{u}$  helyett  $\text{div } \mu \text{grad } \vec{u}$  szerepel, ahol  $\mu$  függ  $\text{grad } \vec{u}$  hosszától (vagy pedig  $\sigma$ -t egy további differenciálegyenlet definiálja, úgy hogy érdemes  $\sigma$ -t is keresett mennyiségnek tekinteni).

A kétdimenziós Navier–Stokes egyenletek komponensenkénti alakját úgy írjuk fel, hogy  $x_1, x_2$  helyett az  $x, y$  jelölést használjuk (és  $u_1, u_2$  helyett  $u, v$ -t) :

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + \frac{\partial p}{\partial x} = \nu \Delta u + f_1, \quad (17.11)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + \frac{\partial p}{\partial y} = \nu \Delta v + f_2, \quad (17.12)$$

$$\text{div } \vec{u} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0. \quad (17.13)$$

Legyen  $L$  egy jellemző hosszúság,  $u_0$  egy jellemző sebesség (tehát  $L/u_0$  egy jellemző időegység). Ekkor az

$$(u, v) \rightarrow (u_0 \bar{u}, u_0 \bar{v}), \quad (x, y) \rightarrow L(\bar{x}, \bar{y}), \quad t \rightarrow \frac{L}{u_0} \bar{t}, \quad p \rightarrow u_0^2 \bar{p}, \quad f \rightarrow \frac{u_0^2}{L} \bar{f}$$

transzformációkkal kapjuk a „dimenziótlan” alakot,

$$\frac{\partial}{\partial \bar{t}} \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} + \bar{u} \frac{\partial}{\partial \bar{x}} \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} + \bar{v} \frac{\partial}{\partial \bar{y}} \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} + \overline{\text{grad}} \bar{p} = \frac{1}{\text{Re}} \overline{\Delta} \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} + \bar{f}, \quad (17.14)$$

$$\overline{\text{div}} \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} = \frac{\partial \bar{u}}{\partial \bar{x}} + \frac{\partial \bar{v}}{\partial \bar{y}} = 0, \quad (17.15)$$

ahol  $\text{Re} := \frac{u_0 L}{\nu}$  a Reynolds-szám. Ezután szokás a felülvonást mindenütt (17.14)-(17.15)-ben elhagyni.

Ahogy már 16.3-ban rámutattunk, a fenti *dimenziótlanítás* olyan művelet, amellyel minden gyakorlati feladat megoldását érdemes kezdeni, mert így az derül ki, hogy mely paraméter-kombinációktól függ valóban a vizsgált jelenség. Jelen esetünkben tehát a Reynolds-szám lényeges.

A következő táblázatban néhány, a Reynolds-számmal kapcsolatos adatot gyűjtöttünk össze.

Reynolds-szám	jellemző jelenség
4	henger körüli áramlás esetén : a henger mögött két örvény jelenik meg, az áramlás többé nem szimmetrikus
Re < 30	csőben : a Hagen–Poiseuille áramlás tartománya
Re < 40	stacionárius és aszimptotikusan stabil áramlás a henger körül
40 < Re	instacionáriussá válik a henger mögötti áramlás
40 < Re < 120	instacionárius, periodikus áramlás a henger mögött
Re ≤ 130	stacionárius a gömb körüli áramlás
55–200	a Kármán-féle örvényút tartománya
200–400	henger körüli áramlás : átmenet a turbulens áramláshoz
2000–3000	csőbeli áramlás : átmenet a turbulens áramláshoz
5000	a vér áramlása a főütőérben
3 · 10 <sup>5</sup> –3 · 10 <sup>6</sup>	turbulens lesz az áramlás a hengerhez tapadó határretegben is
7.4 · 10 <sup>6</sup>	4 m hosszú, 100 km/h sebességgel haladó személygépkocsi
10 <sup>9</sup>	áramlás modern repülőgép törse mentén

A Hagen–Poiseuille-féle áramláshoz megjegyezzük, hogy ez – kísérletek szerint – a feltüntetett kicsi Re értékekre csövekben jó közelítéssel valósul meg. A Hagen–Poiseuille áramlásnak a Navier–Stokes egyenletek kevés ismert *egzakt megoldásai* egyike felel meg :

Stacionáriusan áramlik viszkózus, összenyomhatatlan folyadék egy  $R$  sugarú,  $x_1$ -irányban fekvő csőben.  $x_1 = 0$ -nál  $p_1$  a nyomás,  $x_1 = 1$ -nél  $p_2 < p_1$  a nyomás. Keressük azt a megoldást, amelynek sebességvektora  $\vec{u} = (u_1(r), 0, 0)^T$  alakú, ahol  $u_1(R) = 0$ , külső erők nincsenek!

Ekkor a  $0 = \operatorname{div} \vec{u}$  feltétel teljesül, az  $u_2$ - és  $u_3$ -egyenletekből következik  $p = p(x_1)$ . Továbbá,

$$\begin{aligned} \operatorname{div}_{x_1 x_2 x_3} \vec{u} &= \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} = 0 \\ &= \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial r} \cos \varphi - \frac{\partial u_2}{\partial \varphi} \frac{\sin \varphi}{r} + \frac{\partial u_3}{\partial r} \sin \varphi + \frac{\partial u_3}{\partial \varphi} \frac{\cos \varphi}{r} = \operatorname{div}_{x_1 r \varphi} \vec{u}. \end{aligned}$$

Itt  $u_1$  helyett  $u$ -t írva és  $x_1$  helyett  $x$ -et, az  $u_1$ -egyenletből kapjuk, hogy :

$$\begin{aligned} \frac{\partial p}{\partial x} = \frac{1}{\operatorname{Re}} \Delta_{x r \varphi} u &= \frac{1}{\operatorname{Re}} \left( \frac{\partial^2 u}{\partial x^2} + \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} \right) \\ &= \frac{1}{\operatorname{Re} \cdot r} \frac{\partial}{\partial r} \left( r \frac{\partial u}{\partial r} \right), \end{aligned}$$

és mivel a bal oldal csak  $x$ -től, a jobb oldal csak  $r$ -től függ, így

$$\frac{\partial p}{\partial x} = \text{const}, \quad p = p_1 + (p_2 - p_1)x, \quad (17.16)$$

$$\text{Re} \cdot r \cdot (p_2 - p_1) = (ru')', \quad u = \frac{\text{Re}}{4}(p_1 - p_2)R^2 \left(1 - \left(\frac{r}{R}\right)^2\right), \quad (17.17)$$

az  $u$ -ra és  $p$ -re felírt peremfeltételeket is figyelembe véve.

Innen egyébként az következik, hogy

$$q = 2\pi \int_0^R ur \, dr = \frac{\pi}{8} R^4 \text{Re}(p_1 - p_2)$$

az átáramló térfogat per időegység. Ez Ohm-féle törvénynek értelmezhető, ld. I. 33. o. is, tehát  $\frac{8}{\pi R^4 \text{Re}}$  a cső áramlási ellenállása. (További példákat egzaktul megoldható feladatokra tartalmaz az 1. feladat.)

Fenti példánkban szerepelnek *peremfeltételek* a nyomásra vonatkozólag. Ilyenek tipikusan akkor fordulnak elő, ha szabad felülete van az áramló közegnek. Egyébként viszont csak valamelyik pontban egy vonatkoztatási nyomást kell megadni.

A sebességre az  $\Omega$  áramlási tartomány falán peremfeltételeket kell megadni, amelyek két részből állnak fizikailag :

a) A fal nem áteresztő :  $\vec{u} \cdot \vec{n} = \vec{U} \cdot \vec{n}$ , ahol  $\vec{U}$  a fal sebessége és  $\vec{n}$  a külső normálvektor;

b) a fal érintősíkjában megegyezik a fal és a folyadék sebessége:  $\vec{u} \times \vec{n} = \vec{U} \times \vec{n}$ .

Ez utóbbi peremfeltétel a viszkozitás hatása miatt érvényes (és jó megegyezéssel igazolható kísérletileg).

Végül, időtől függő feladatban a sebességekre kezdeti feltételeket is szükséges megadni.

A Navier–Stokes egyenletek megoldására vonatkozó *egzisztencia-eredmények* főként Leray-re és E. Hopf-ra mennek vissza, ld. Ladüzsenszkaja és Temam, valamint Galdi és Sohr könyveit.

Ezek az eredmények biztosítják a gyenge megoldás létezését nemcsak elég általános háromdimenziós tartomány esetére is, hanem  $n > 3$ -ra is. Ami jelenleg még nem teljesen megoldott kérdés, az a gyenge megoldás regularitása. A klasszikus megoldás létezéséről szóló eredmények részben olyan feltételt is tartalmaznak, amely a kezdeti függvény kicsi voltát követeli meg.

A Navier–Stokes egyenletek energiamegmaradási tétele matematikai és fizikai szempontból egyaránt fontos. Ennek a levezetését az alábbiakban megmutatjuk azért, mert egészen hasonló lesz majd a differencia-approximáció stabilitási tételnek a bizonyítása.

**Megmaradási tétel.** Legyen (17.9)-(17.10)-ben  $d = 3$  és

$$E(t) := \frac{\rho}{2} \int_{\Omega} (u_1^2(x, t) + u_2^2(x, t) + u_3^2(x, t)) \, dx$$

az áramló közegnek a kinetikus energiája. Ez zárt edényben csak csökkenhet.

Ennek belátásához a (17.9) rendszer  $i$ -edik egyenletét  $\rho u_i$ -vel szorozzuk meg, majd figyelembe véve a (17.3) relációkat és azt, hogy  $\rho \sum_i u_i \frac{\partial u_i}{\partial t} = \frac{\rho}{2} \frac{\partial}{\partial t} (u_1^2 + u_2^2 + u_3^2)$ , a kapott egyenleteket  $i = 1, 2, 3$ -re adjuk össze, és ezután  $\Omega$  felett integrálunk :

$$\begin{aligned} \frac{dE}{dt} + \rho \int_{\Omega} \sum_i \operatorname{div}(u_i \vec{u}) u_i \, dx + \int_{\Omega} \vec{u} \cdot \operatorname{grad} P \, dx \\ = \mu \int_{\Omega} \sum_i u_i \Delta u_i \, dx + \int_{\Omega} \rho \vec{f} \cdot \vec{u} \, dx. \end{aligned}$$

Alakítsuk át az egyes integrálokat! Ehhez használjuk a parciális integrálást (v.ö. a 15.7.1. pontban a (15.194)-(15.196) relációkkal) :

$$\int_{\Omega} \operatorname{div}(u_i \vec{u}) u_i \, dx = \int_{\Gamma} u_i^2 \vec{u} \cdot \vec{n} \, ds - \int_{\Omega} u_i \vec{u} \cdot \operatorname{grad} u_i \, dx,$$

valamint (15.195) következő alakját is :

$$\int_{\Omega} u_i \vec{u} \cdot \operatorname{grad} u_i \, dx = \int_{\Omega} \vec{u} \cdot \operatorname{grad} \left( \frac{1}{2} u_i^2 \right) \, dx = \int_{\Gamma} \vec{u} \cdot \vec{n} \left( \frac{1}{2} u_i^2 \right) \, ds - \int_{\Omega} \frac{1}{2} u_i^2 \operatorname{div} \vec{u} \, dx.$$

Ha tehát  $(\vec{u} \cdot \vec{n})|_{\Gamma} = 0$ , akkor (17.10) miatt

$$\int_{\Omega} \operatorname{div}(u_i \vec{u}) u_i \, dx = 0.$$

A konvekciós tagok ezek szerint nem változtatják még a kinetikus energia összegét  $\Omega$ -ban emellett a peremfeltétel mellett. Továbbá

$$\int_{\Omega} \operatorname{grad} P \cdot \vec{u} \, dx = \int_{\Gamma} P \vec{u} \cdot \vec{n} \, ds - \int_{\Omega} P \operatorname{div} \vec{u} \, dx = 0,$$

ugyancsak akkor, ha  $(\vec{u} \cdot \vec{n})|_{\Gamma} = 0$ . Végülis (ld. (15.196))

$$\int_{\Omega} u_i \Delta u_i \, dx = \int_{\Gamma} u_i (\operatorname{grad} u_i) \cdot \vec{n} \, ds - \int_{\Omega} |\operatorname{grad} u_i|^2 \, dx.$$

Itt a peremtag elesik, ha vagy  $u_i|_{\Gamma} = 0$  vagy  $\left( \frac{\partial u_i}{\partial n} \right) = 0$ .



Ha nincsenek térfogati erők vagy azok merőlegesek  $\vec{u}$ -ra, vagy ha potenciálisak (mint a nyomás), amilyen pl. a gravitációs erőtér vagy az  $\vec{f} = c_0 \text{grad } T$  formában figyelembe vett felhajtó erő :

$$\vec{f} = \text{grad } \varphi,$$

akkor

$$\int_{\Omega} \vec{f} \cdot \vec{u} \, dx = \int_{\Omega} \text{grad } \varphi \cdot \vec{u} \, dx = \int_{\Gamma} \varphi \vec{u} \cdot \vec{n} \, ds - \int_{\Omega} \varphi \text{div } \vec{u} \, dx = 0,$$

$(\vec{u} \cdot \vec{n})|_{\Gamma} = 0$  esetén.

Amennyiben pl.  $\vec{u}|_{\Gamma} = 0$ , akkor minden peremtag eltűnik és ekkor

$$\frac{dE}{dt} = -\mu \int_{\Omega} \sum_{i=1}^3 |\text{grad } u_i|^2 \, dx \leq 0 \quad (17.18)$$

következik, tehát a  $\mu \Delta \vec{u}$  tag disszipálja az energiát, a konvekció és a nyomástag az energiát csak átrendezi a (zárt) edényen belül.  $\square$

## 17.2 Alapvető numerikus problémák, az áramfüggvény

A (17.11)-(17.13) Navier–Stokes egyenletek numerikus megoldása során a következő nehézségekkel találkozunk :

1) A  $\text{div } \vec{u} = 0$  feltétel (ha mellékfeltételként tudjuk kezelni, akkor parabolikus egyenletekkel állunk szemben – csak hogy ez nem egyszerű);

2) A nyomásra nincsen külön egyenletünk. A helyes nyomás az, amely mellett a sebesség (mint az impulzusegyenlet megoldása) teljesíti a tömegmegmaradási egyenletet is;

3) A konvekciós-diffúziós probléma (műszakilag érdekes áramlásoknál a Reynolds-szám gyakran nagy, ld. a táblázatot. Ekkor kicsi a kinematikus viszkozitás a feladat hosszúsági méretéhez és tipikus sebességéhez képest. A konvekciós-diffúziós probléma miatt v.ö. II. 11.4.10-zel);

4) A  $\sum_j u_j \frac{\partial \vec{u}}{\partial x_j}$  tag nemlinearitása (erre nyilván iterációkat fogunk alkalmazni, pl.  $u_j \frac{\partial u_i}{\partial x_j}$  helyett  $u_j^{(k)} \frac{\partial u_i^{(k+1)}}{\partial x_j}$ -t véve lineárizáljuk az egyenleteket. Itt  $k$  az iteráció számláló).

Az első két problémával azonnal foglalkozunk, a 3.-ra és 4.-re \*\*\* 17.4-ben térünk vissza.

A kétdimenziós esetben a numerikus szempontból kellemetlen (17.5) mellékfeltételt a  $\psi$  áramfüggvény segítségével elégíthetjük ki :

$$u =: \frac{\partial \psi}{\partial y}, \quad v =: -\frac{\partial \psi}{\partial x}, \quad \text{összefoglalva : } \begin{pmatrix} u \\ v \end{pmatrix} = \text{curl } \psi. \quad (17.19)$$

Hogy ezt pontosan mikor tehetjük, arra lejjebb térünk ki. A (17.19) definícióból látjuk, hogy egyrészt

$$\text{div} \begin{pmatrix} u \\ v \end{pmatrix} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = \frac{\partial}{\partial x} \left( \frac{\partial \psi}{\partial y} \right) + \frac{\partial}{\partial y} \left( -\frac{\partial \psi}{\partial x} \right) = 0, \quad (17.20)$$

a (17.10) összenyomhatatlansági feltétel tehát automatikusan teljesül. Másrészt

$$\text{grad } \psi \perp \vec{u}, \quad \text{hiszen } \frac{\partial \psi}{\partial x} u + \frac{\partial \psi}{\partial y} v = 0.$$

A sebesség vektora tehát  $\psi$  izovonalainak az érintő vektora, másszóval  $\psi = \text{const}$  az áramvonalakon.

A (17.19) áramfüggvény mellett vezessük be még az *örvényességet* is :

$$\omega := \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = (\text{rot } \vec{u})_3 = -\frac{\partial^2 \psi}{\partial y^2} - \frac{\partial^2 \psi}{\partial x^2} = -\Delta \psi, \quad (17.21)$$

ld. (17.19). Amennyiben az áramló közeg örvényessége nulla, akkor tehát

$$0 = \Delta \psi. \quad (17.22)$$

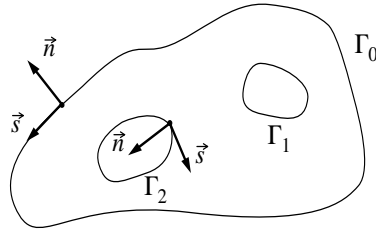
A (17.20), (17.21),  $\omega = 0$  egyenletek (azaz  $\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$ ,  $\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} = 0$ ) jól ismertek mint a *Cauchy–Riemann egyenletrendszer*. Itt azt látjuk, hogy ez a rendszer a stacionárius sűrűlódásmentes (nemviszkózus) örvénymentes, összenyomhatatlan áramlás leírását adja. Nemcsak a  $\psi$  áramfüggvény, hanem – mint a Cauchy–Riemann rendszerből következik – az  $u$  és  $v$  sebesség komponensek is eleget tesznek a Laplace–egyenletnek :

$$\Delta u = 0, \quad \Delta v = 0.$$

A klasszikus függvénytan éppen a felsorolt tulajdonságokkal rendelkező áramlás elméletét adja.

A  $\rho = \text{const}$  sűrűségű áramló közeg összes kinetikus energiáját megadó

$$J(\psi) := \frac{\rho}{2} \int \int_{\Omega} \left[ \left( \frac{\partial \psi}{\partial x} \right)^2 + \left( \frac{\partial \psi}{\partial y} \right)^2 \right] dx dy = \frac{\rho}{2} \int \int_{\Omega} [u^2 + v^2] dx dy$$



ábra 17.1: Többszörösen összefüggő tartomány, jelölések

integrál (v.ö. (17.19)-cel) a (17.22) egyenlet elméletében is fontos szerepet játszik és – eltekintve a  $\frac{L}{2}$  szorzótól – az egyenlethez tartozó bilineáris formával is kapcsolatos.

A háromdimenziós esetben a  $\psi$  áramfüggvény helyett a  $\vec{\psi}$  úgynevezett vektorpotenciált vezethetjük be :

$$\vec{u} := \text{rot } \vec{\psi},$$

ahol (az  $\vec{i}, \vec{j}, \vec{k}$  koordináta egységvektorokkal)

$$\text{rot } \vec{\psi} := \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ \psi_1 & \psi_2 & \psi_3 \end{vmatrix} = \vec{i} \left( \frac{\partial \psi_3}{\partial x_2} - \frac{\partial \psi_2}{\partial x_3} \right) + \vec{j} \left( \frac{\partial \psi_1}{\partial x_3} - \frac{\partial \psi_3}{\partial x_1} \right) + \vec{k} \left( \frac{\partial \psi_2}{\partial x_1} - \frac{\partial \psi_1}{\partial x_2} \right).$$

### 17.2.1 Az áramfüggvény peremértékei és létezése

Tisztázzuk először, hogyan kell a  $\psi$ -re a kétdimenziós esetben peremfeltételeket előírni!

1. Át nem eresztő, nyugvó falon érvényes  $u = v = 0$ . Ebből következik  $\text{grad } \psi = 0$ , tehát

$$\frac{\partial \psi}{\partial n} = \text{grad } \psi \cdot \vec{n} = 0 \quad \text{és} \quad \frac{\partial \psi}{\partial s} = \text{grad } \psi \cdot \vec{s} = 0.$$

Itt  $\vec{n} = (n_1, n_2)^T$  a perem külső normálvektora,  $\vec{s} = (-n_2, n_1)^T$  az érintővektora. A  $\frac{\partial \psi}{\partial s} = 0$  tulajdonságból következik, hogy zárt  $\Gamma$  peremvonal esetén csakis konstans lehet  $\psi$  a peremen. Ha az  $\Omega$  áramlási tartomány többszörösen összefüggő, pl.  $p + 1$ -szeresen :

$$\Gamma = \bigcup_{i=0}^p \Gamma_i,$$

ahol  $\Gamma_0$  a külső pereme, akkor minden egyes  $\Gamma_i$  komponensen az áramfüggvény konstans. Mivel  $\psi$  a deriváltjain keresztül van definiálva, csak egy

konstans erejéig egyértelműen meg van határozva, így pl.  $\Gamma_0$ -n  $\psi$ -t rögzíthetjük le nullára. A perem többi komponensén az áramfüggvény ismeretlen, de konstans értéket vesz fel,

$$\psi|_{\Gamma_0} = 0; \quad \psi|_{\Gamma_i} = \text{const}_i, \quad 1 \leq i \leq p.$$

Ha a fal át nem eresztő, de mozog (adott  $\vec{U}$  vektor sebességgel), akkor ott az áramló közeg sebességére az  $\vec{u} = \vec{U}$  feltételt írjuk elő.

**2.** Nyugvó, áteresztő falon pl. a sebesség normálirányú komponensét írjuk elő :

$$\vec{u} \cdot \vec{n} = a(x, y).$$

A definíció alapján ekkor

$$a(x, y) = \vec{u} \cdot \vec{n} = n_1 \frac{\partial \psi}{\partial y} - n_2 \frac{\partial \psi}{\partial x} = \text{grad } \psi \cdot \vec{s} = \frac{\partial \psi}{\partial s}.$$

Ha a peremnek az ilyen  $\Gamma_i$  komponense zárt, akkor nyilván kell, hogy legyen

$$\oint_{\Gamma_i} \vec{u} \cdot \vec{n} \, ds = \oint \frac{\partial \psi}{\partial s} \, ds = 0,$$

tehát ilyen mellékfeltétele van  $\vec{u} \cdot \vec{n}$  megadásának, mert egyébként a  $\psi$ -t nem lehetne egyértelműen bevezetni. Az, hogy  $\vec{u} \cdot \vec{n}$  integrálja egy  $\gamma$  peremszakaszon nem nulla, az viszont hasznos akkor, amikor rajta keresztül ki- vagy beáramlás történik. Ekkor

$$q = \int_{\gamma} \vec{u} \cdot \vec{n} \, ds = \int_{\gamma} \frac{\partial \psi}{\partial s} \, ds$$

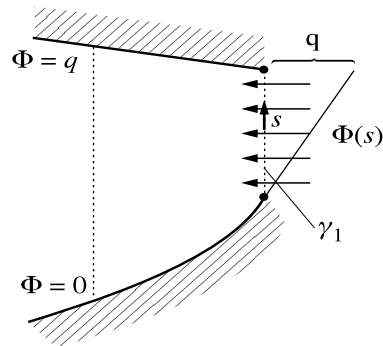
az átáramló térfogat per időegység ezen a szakaszon. Legyen  $\psi = 0$  a  $\gamma$  kezdeti pontjában (az át nem eresztő falszakasz végén), akkor, ha ismert az  $a(x, y)$  normálirányú sebesség  $\gamma$ -nak minden  $(x(s), y(s))$  pontjában,

$$\psi(x(s), y(s)) = \int_0^s a(x(\tau), y(\tau)) \, d\tau.$$

A  $\gamma$  szakasz végpontjában ( $s = L$ ) a

$$\psi(x(L), y(L)) = \int_0^L a(x(L), y(L)) \, ds = q$$

értéket érjük el, és ezzel, mint konstans értékkel, folytatódik  $\psi$  a  $\gamma$  másik végpontjánál kezdődő át nem eresztő szakaszon. Kívánatos, hogy ilyen peremfeltételt a beáramlási szakaszokon meg tudjunk adni.



ábra 17.2: Beáramlási szakasz

**3.** Ha csak a kiáramló térfogat per időegység (a volumenáramlás) ismert egy kiáramló szakaszon, akkor bevált az, hogy a  $\gamma$  szakaszon mint peremfeltételt előírjuk, hogy a kiáramlás az  $\vec{n}$  normálvektor irányába történjen, vagyis ortogonális a perem  $\vec{s}$  érintővektorára :

$$0 = \vec{u} \cdot \vec{s} = \frac{\partial \psi}{\partial y} (-n_2) + \left( -\frac{\partial \psi}{\partial x} \right) n_1 = -\text{grad } \psi \cdot \vec{n} = -\frac{\partial \psi}{\partial n}.$$

Ezt a Neumann-féle feltételt írjuk elő mint  $\psi$  peremfeltételét. Ha pl.  $x = 1$ ,  $0 < y < 1$ -nél kiáramlás történik  $x > 1$  irányába, akkor  $\vec{n} = (1, 0)^T$  és

$$\frac{\partial \psi}{\partial x} = 0$$

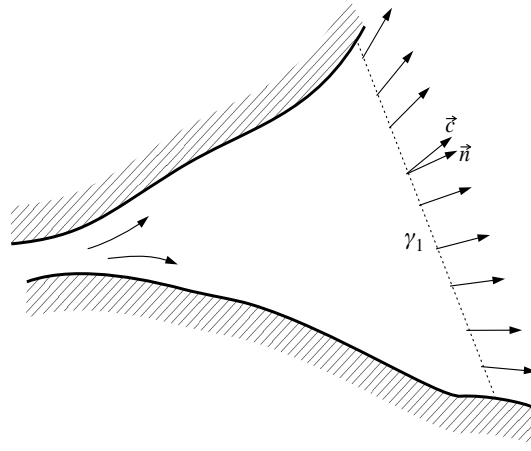
a peremfeltétel.

**4.** Előfordulhat, hogy  $\gamma$ -n csak az  $u$  iránya ismert (vagy ismertnek tételezzük fel), pl. ha egy tengeröbölre akarunk áramlási számításokat végezni és valahol a tengeren, az öbötől nem túl nagy távolságra, le kell zárni a számítást, de ott túl durva lenne előírni azt, hogy  $\vec{u}$  a normálvektor irányába mutat. Ekkor az előző ötlet általánosítását alkalmazhatjuk – ha a megfelelő információval rendelkezünk : ha  $\vec{u}$  iránya megfelel  $\Gamma_i$  mentén (ahol az  $s$  koordinátát használjuk) a  $\vec{c}(s) = (c_1(s), c_2(s))^T$  egységvektorénak, akkor a következő *ferdeirányú peremfeltételt* kapjuk :

$$0 = -c_2 u + c_1 v = -\left( c_1 \frac{\partial \psi}{\partial x} + c_2 \frac{\partial \psi}{\partial y} \right) = -\text{grad } \psi \cdot \vec{c} = \frac{\partial \psi}{\partial c}.$$

Ilyen peremfeltétel nem okoz problémát, ha  $\vec{c}$  és  $\vec{n}$  hegyesszöget zárnak be.

Vegyük észre, hogy ezek a gyakorlati megoldást segítő ötletek két tény nem tudnak eltakarni :



ábra 17.3: Ferdeirányú peremfeltétel

- a) két peremfeltétel érvényes valójában  $\psi$ -re, pl.  $\psi = 0$   $\Gamma_0$ -n és ugyanekkor  $\frac{\partial \psi}{\partial n} = 0$  ( $\Gamma_i$  legyen nyugvó, át nem eresztő),  
 b) nem minden megengedhető peremfeltételre tudunk  $\psi$ -t egyértelműen meghatározni, pl. ha a  $\Gamma_i$  zárt falon érvényes  $\oint_{\Gamma_i} \vec{u} \cdot \vec{n} \, ds \neq 0$ . A Gauss–Osztrogradskij tétel segítségével csupán a

$$\int_{\Gamma} \vec{u} \cdot \vec{n} \, ds = 0, \quad \Gamma = \bigcup_{i=0}^p \Gamma_i,$$

feltétel következik, nem pedig  $\int_{\Gamma_i} \vec{u} \cdot \vec{n} \, ds = 0$  minden  $i$ -re. Ilyen esetekben mit lehet tenni?

Az ilyen áramlási mezőt nem (kizárólagosan) áramfüggvény segítségével kell előállítani.

Először tisztázzuk, hogy divergenciamentes áramlási mezőt pontosan mikor lehet áramfüggvény segítségével leírni.

**17.1. Tétel** (áramfüggvény létezése és unicitása). Legyen  $\Gamma = \bigcup_{i=0}^p \Gamma_i$  Lipschitz-folytonos pereme a korlátos, összefüggő  $\Omega \in \mathbb{R}^2$  tartománynak,  $\Gamma_0$  a külső pereme. Tetszőleges  $\vec{v} \in V := \{\vec{u} \in (H_0^1(\Omega))^2, \operatorname{div} \vec{u} = 0\}$  vektorfüggvényhez pontosan egy

$$\psi \in H_{00}^2 := \{y \in H^2(\Omega), y|_{\Gamma_0} = 0, y|_{\Gamma_i} = \text{const}, \frac{\partial y}{\partial n} \Big|_{\Gamma} = 0\}$$

áramfüggvény létezik (tehát érvényes  $\vec{v} = \operatorname{curl} \psi$ ), és megfordítva.  $\square$

Ezen tétel bizonyítása Temam könyvében található (I. fejezet, 4.4. pont).

**Megjegyzések. 1.** Az előzőek szerint, ha minden fal át nem eresztő, akkor  $\vec{v}|_{\Gamma} = 0$  mindenütt és így, ha  $\vec{v} \in H^1$  és  $\operatorname{div} \vec{v} = 0$ , az áramfüggvény létezik és egyértelmű.

**2.** Hasonló tétel Girault és Raviart könyvében is szerepel, ahol az  $\vec{u} \in (L^2(\Omega))^2$  esetét vizsgálják. Ehhez pontosan akkor létezik egy (17.19)-et kielégítő  $\psi \in H^1(\Omega)$ , ha igaz (disztribúciók értelmében), hogy

$$\operatorname{div} \vec{u} = 0, \quad \int_{\Gamma_i} \vec{u} \cdot \vec{n} \, ds = 0, \quad 1 \leq i \leq p. \quad (17.23)$$

**3.** Ha a  $\psi \in H^1$  áramfüggvény létezik, akkor  $\vec{u} = \operatorname{curl} \psi$  teljesíti a (17.23) feltételt, ugyanis  $\operatorname{div} \operatorname{curl} \psi = 0$  és

$$\vec{n} \cdot \operatorname{curl} \psi = \frac{\partial \psi}{\partial x_2} n_1 - \frac{\partial \psi}{\partial x_1} n_2 = \operatorname{grad} \psi \cdot \vec{s} = \frac{\partial \psi}{\partial s}$$

ahol az  $\vec{s} := (-n_1, n_2)^T$  vektor  $\vec{n}$ -re ortogonális. Ha  $\vec{n}$  a  $\Gamma_i$  peremszakasz normálvektora, akkor

$$\int_{\Gamma_i} \vec{n} \cdot \operatorname{curl} \psi \, ds = \int_{\Gamma_i} \frac{\partial \psi}{\partial s} \, ds = 0,$$

mivel  $\Gamma_i$  zárt.

**4.** Megfelelő változtatással (curl helyett rot szerepel) a háromdimenziós esetben is igaz a tétel, tehát – az első megjegyzés értelmében – (17.23) mellett van olyan  $\vec{\psi}$ , amellyel  $\vec{u} = \operatorname{rot} \vec{\psi}$ , sőt a  $\vec{\psi}$  vektorpotenciálra még  $\operatorname{div} \vec{\psi} = 0$  is teljesíthető. Vegyük észre, hogy a kétdimenziós eset itt is a háromdimenziós eset speciálizálódásával adódik:  $\vec{u} = (u_1(x_1, x_2), u_2(x_1, x_2), 0)^T$  és

$$\vec{\psi} = \begin{pmatrix} 0 \\ 0 \\ \psi(x_1, x_2) \end{pmatrix}, \quad \operatorname{rot} \vec{\psi} = \begin{pmatrix} \frac{\partial \psi}{\partial x_2} \\ -\frac{\partial \psi}{\partial x_1} \\ 0 \end{pmatrix} = \begin{pmatrix} \operatorname{curl} \psi \\ 0 \end{pmatrix}, \quad \operatorname{div} \vec{\psi} = \frac{\partial \psi}{\partial x_3} = 0. \quad \square$$

Az áramfüggvény vizsgálata után már könnyebb felelni arra a kérdésre, hogyan írható fel az a  $\vec{v}$  divergenciamentes sebességmező, amely csupán a

$$\operatorname{div} \vec{v} = 0, \quad x \in \Omega; \quad \vec{v}|_{\Gamma} = \vec{a}(s), \quad \int_{\Gamma} \vec{a} \cdot \vec{n} \, ds = 0 \quad (17.24)$$

feltételeknek tesz eleget: a következő két ötletet kell megfelelően kombinálni, amellyel lehet (speciális) divergenciamentes áramlási mezőt előállítani:

a) Oldjuk meg a  $\Delta \varphi = \operatorname{div}(\operatorname{grad} \varphi) = 0$  egyenletet (megfelelő peremfeltételek mellett), majd vegyük  $\varphi$  gradiensét!

b) Vegyük valamelyik  $\vec{\psi}$  vektorfüggvény rotációját!

**17.2. Tétel** (divergenciamentes vektormező ortogonális felbontása). Legyen  $\vec{v} \in (L_2(\Omega))^3$  és teljesüljenek a (17.24) feltételek. Ekkor van olyan  $\varphi$  és  $\vec{\psi}$ , hogy

$$\vec{v} = \text{grad } \varphi + \text{rot } \vec{\psi}; \quad \vec{\psi} \in (H^1(\Omega))^3, \quad (\vec{n} \cdot \text{rot } \vec{\psi})|_{\Gamma} = 0; \quad (17.25)$$

$$\varphi \in H^1(\Omega), \quad \Delta \varphi = 0, \quad x \in \Omega, \quad \left. \frac{\partial \varphi}{\partial n} \right|_{\Gamma} = \vec{a} \cdot \vec{n}. \quad (17.26)$$

A (17.25) felbontás ortogonális.

**Bizonyítás.** A (17.26) feladatnak van megoldása (általánosított értelemben:  $H^1$ -ben), mert a  $\int_{\Gamma} \frac{\partial \varphi}{\partial n} ds = 0$  kompatibilitási feltétel teljesül. Ezután legyen

$$\vec{w} := \vec{v} - \text{grad } \varphi \in (L_2(\Omega))^3.$$

Ez a vektorfüggvény a következő tulajdonságokkal rendelkezik:

$$\begin{aligned} \text{div } \vec{w} &= 0, \quad x \in \Omega \quad (\text{disztribúciók értelmében}), \\ \vec{w} \cdot \vec{n} &= \vec{a} \cdot \vec{n} - \frac{\partial \varphi}{\partial n} = 0, \quad x \in \Gamma. \end{aligned}$$

Az előző tétel alapján  $\vec{w} = \text{rot } \vec{\psi}$ , és itt  $\vec{\psi} \in (H^1(\Omega))^3$  teljesíti még a  $\text{div } \vec{\psi} = 0$  feltételt is.  $\square$

**17.3. Következmény** (tetszőleges vektormező ortogonális felbontása). Tetszőleges  $\vec{v} \in (L_2(\Omega))^3$  is felbontható a (17.25) alakban (csak (17.26) módosul), és ez a felbontás ortogonális. Legyen ugyanis

$$\Delta \psi = \text{div } \vec{v}, \quad x \in \Omega; \quad \psi = 0, \quad x \in \Gamma.$$

Ezen feladat megoldása létezik és teljesíti az

$$\int_{\Gamma} \frac{\partial \psi}{\partial n} ds = \int_{\Omega} \Delta \psi dx = \int_{\Omega} \text{div } \vec{v} dx = \int_{\Gamma} \vec{v} \cdot \vec{n} ds$$

feltételt. Vezessük be a következő jelölést:

$$\vec{w} := \vec{v} - \text{grad } \psi,$$

akkor

$$\text{div } \vec{w} = 0, \quad \int_{\Gamma} \vec{w} \cdot \vec{n} ds = \int_{\Gamma} \left( \vec{v} \cdot \vec{n} - \frac{\partial \psi}{\partial n} \right) ds = 0.$$

Most alkalmazzuk az előző tételt:

$$\vec{w} = \text{grad } \varphi + \text{rot } \vec{\Phi},$$



ahol

$$\Delta\varphi = 0, \quad x \in \Omega, \quad \frac{\partial\varphi}{\partial n} = (\vec{w} \cdot \vec{n})|_{\Gamma}, \quad (\vec{n} \cdot \text{rot } \vec{\Phi})|_{\Gamma} = 0.$$

Így kaptuk a

$$\vec{v} = \text{grad}(\varphi + \psi) + \text{rot } \vec{\Phi}$$

felbontást. Ez ortogonális : mivel

$$\begin{aligned} \text{div} \left( (\varphi + \psi) \text{rot } \vec{\Phi} \right) &= \text{grad}(\varphi + \psi) \cdot \text{rot } \vec{\Phi} + (\varphi + \psi) \text{div rot } \vec{\Phi} \\ &= \text{grad}(\varphi + \psi) \cdot \text{rot } \vec{\Phi}, \end{aligned}$$

így

$$\begin{aligned} \int_{\Omega} \text{grad}(\varphi + \psi) \cdot \text{rot } \vec{\Phi} \, dx &= \int_{\Omega} \text{div} \left( (\varphi + \psi) \text{rot } \vec{\Phi} \right) \, dx \\ &= \int_{\Gamma} (\varphi + \psi) \vec{n} \cdot \text{rot } \vec{\Phi} \, dx = 0. \end{aligned}$$

Ennek alapján a felbontás egyértelmű is.  $\square$

**Megjegyzés.** A bizonyítás során levezetett

$$\vec{v} = \text{grad } \varphi + \text{grad } \psi + \text{rot } \vec{\Phi}$$

felbontás is ortogonális. Ennek belátásához csak azt kell még bebizonyítani, hogy  $\text{grad } \varphi$  és  $\text{grad } \psi$  ortogonálisak :

$$\int_{\Omega} \text{grad } \varphi \cdot \text{grad } \psi \, dx = \int_{\Gamma} \psi \frac{\partial\varphi}{\partial n} \, ds - \int \psi \Delta\varphi \, dx = 0. \quad \square$$

Az ortogonális felbontás birtokában tisztázhatjuk a nyomás szerepét :

Az előzőek szerint tetszőleges  $\vec{v} \in (L_2(\Omega))^3$  vektor ortogonálisan felbontható. Ezt a felbontást most

$$\vec{v} = P\vec{v} + Q\vec{v}$$

alakban írjuk fel, ahol

$$P\vec{v} = \text{grad}(\varphi + \psi), \quad Q\vec{v} = (I - P)\vec{v} = \text{rot } \vec{\Phi}$$

az ortogonális projektorok (ld. a 17.3. következményt).

Legyenek (17.9)-ben most az erők potenciálosak :  $\vec{f} = \text{grad } \phi$ , akkor azt kapjuk meg, hogy

$$\text{grad}(p - \phi) = \nu \Delta \vec{u} - \mathbf{D}(\vec{u})\vec{u} - \frac{\partial \vec{u}}{\partial t}.$$

Ha  $\operatorname{div} \vec{u} = 0$  és  $(\vec{u} \cdot \vec{n})|_{\Gamma} = 0$ , akkor  $\vec{u} = Q\vec{u}$  és  $P\vec{u} = 0$ , továbbá  $P \frac{\partial \vec{u}}{\partial t} = \frac{\partial}{\partial t} P\vec{u} = 0$ . Ennek alapján, és mivel  $P$  projektor,

$$\begin{aligned} P \operatorname{grad}(p - \phi) &= \operatorname{grad}(p - \phi) = P \left( \nu \Delta \vec{u} - \mathbf{D}(\vec{u})\vec{u} - \frac{\partial \vec{u}}{\partial t} \right) \\ &= P(\nu \Delta \vec{u} - \mathbf{D}(\vec{u})\vec{u}). \end{aligned}$$

Tehát a nyomás kiegyensúlyozza a disszipációt, ill. konvekció miatt fellépő nem potenciális erőket.

### 17.2.2 Az $\omega - \psi$ -rendszer

Tekintsük a Navier-Stokes egyenleteket a kétdimenziós esetben, ld. (17.11)-(17.13). Ha nincsenek szabad felületek, akkor célszerű az  $\omega - \psi$ -rendszerhez átmenni, ahol  $\omega$  a (17.21) örvényesség és  $\psi$  a (17.19) áramfüggvény. Ehhez feltesszük, hogy a tartomány egyszerűen összefüggő (vagy minden fal át nem eresztő), tehát az áramfüggvény bevezethető. (17.11)-re és (17.12)-re alkalmazzuk a következő két műveletet :

a) vegyük először a rotációt (harmadik komponensét), ld. (17.21),  
 $-\frac{\partial}{\partial y}(17.11) + \frac{\partial}{\partial x}(17.12)$ .

Felcserélve a differenciálások sorrendjét, az következik, hogy  $\omega$ -ra teljesül

$$\begin{aligned} \frac{\partial \omega}{\partial t} + u \frac{\partial \omega}{\partial x} - \frac{\partial u}{\partial y} \frac{\partial u}{\partial x} + \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \\ + v \frac{\partial \omega}{\partial y} - \frac{\partial v}{\partial y} \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} = \nu \Delta \omega + \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}, \end{aligned}$$

a nyomás tehát kiesett, és az egyenlet tovább egyszerűsíthető :

$$\frac{\partial \omega}{\partial t} + u \frac{\partial \omega}{\partial x} + v \frac{\partial \omega}{\partial y} = \nu \Delta \omega + g, \quad (17.27)$$

ahol  $g := \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} = (\operatorname{rot} f)_3$ . Itt azt használtuk fel, hogy

$$-\frac{\partial u}{\partial y} \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \frac{\partial u}{\partial y} = -\frac{\partial u}{\partial y} \operatorname{div} \vec{u} = 0, \quad \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + \frac{\partial v}{\partial x} \frac{\partial v}{\partial y} = \frac{\partial v}{\partial x} \operatorname{div} \vec{u} = 0.$$

b) vegyük a divergenciát (legyen  $D := \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}$ )  
 $\frac{\partial}{\partial x}(17.11) + \frac{\partial}{\partial y}(17.12)$  :

$$\begin{aligned} \frac{\partial D}{\partial t} + u \frac{\partial D}{\partial x} + \left( \frac{\partial u}{\partial x} \right)^2 + \frac{\partial u}{\partial y} \frac{\partial v}{\partial x} \\ + v \frac{\partial D}{\partial x} + \frac{\partial v}{\partial x} \frac{\partial u}{\partial y} + \left( \frac{\partial v}{\partial y} \right)^2 + \Delta p = \nu \Delta D + \operatorname{div} \vec{f}. \end{aligned}$$

Itt a legtöbb tag  $D = 0$  miatt eltűnik, és a visszamaradó

$$\Delta p = - \left[ \left( \frac{\partial u}{\partial x} \right)^2 + 2 \frac{\partial u}{\partial y} \frac{\partial v}{\partial x} + \left( \frac{\partial v}{\partial y} \right)^2 \right] + \operatorname{div} \vec{f}$$

egyenlet

$$0 = D^2 = \left( \frac{\partial u}{\partial x} \right)^2 + 2 \frac{\partial u}{\partial x} \frac{\partial v}{\partial y} + \left( \frac{\partial v}{\partial y} \right)^2$$

alapján a

$$\Delta p = 2 \left[ \frac{\partial u}{\partial x} \frac{\partial v}{\partial y} - \frac{\partial u}{\partial y} \frac{\partial v}{\partial x} \right] + \operatorname{div} \vec{f}$$

alakra hozható, ill. (17.19) segítségével:

$$\Delta p = 2 \left[ \frac{\partial^2 \psi}{\partial x^2} \frac{\partial^2 \psi}{\partial y^2} - \left( \frac{\partial^2 \psi}{\partial x \partial y} \right)^2 \right] + \operatorname{div} \vec{f}. \quad (17.28)$$

(17.27)-et hasonlóan írhatjuk át :

$$\frac{\partial \omega}{\partial t} + \frac{\partial \psi}{\partial y} \frac{\partial \omega}{\partial x} - \frac{\partial \psi}{\partial x} \frac{\partial \omega}{\partial y} = \nu \Delta \omega + g, \quad (17.29)$$

amely egyenletet (17.21)-gyel egészítjük ki :

$$\Delta \psi = -\omega. \quad (17.30)$$

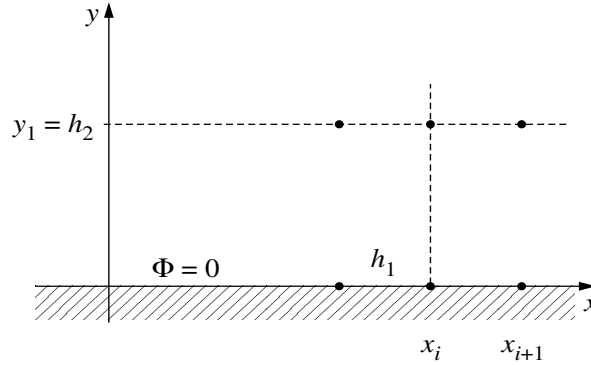
A (17.29)-(17.30) egyenletek előnye, hogy már csak két ismeretlent tartalmaznak,  $\omega$ -t és  $\psi$ -t, míg a (17.28) az előbbieket birtokában utólagosan is megoldható, ha erre szükségünk van.

A háromdimenziós esetben hasonlóan járhatunk el, de kisebb ennek előnye :  $\vec{\omega} = \operatorname{rot} \vec{u}$  is vektor,  $\vec{\psi}$  is vektor ( $\vec{u} = \operatorname{rot} \vec{\psi}$ ), így a 4 ismeretlen helyett  $6 + 1$  lett – csak a nyomást sikerült leválasztani. Ezenkívül itt még nagyobb probléma a peremfeltételek megadása mint a kétdimenziós esetben (ehhez ld. a következő pontot).

Ha behelyettesítjük (17.30)-at (17.29)-be, akkor a

$$- \left( \frac{\partial}{\partial t} \Delta \psi + \frac{\partial \psi}{\partial y} \frac{\partial}{\partial x} \Delta \psi - \frac{\partial \psi}{\partial x} \frac{\partial}{\partial y} \Delta \psi \right) = -\nu \Delta^2 \psi + g \quad (17.31)$$

negyedrendű egyenlethez jutunk – amelyhez valóban kell a 17.2.1-ben az át nem eresztő, nyugvó fal esetére említett két peremfeltétel  $\Gamma$ -n :  $\frac{\partial \psi}{\partial n}|_{\Gamma} = 0$ ,  $\psi|_{\Gamma} = 0$ .



ábra 17.4: Thom-féle peremfeltétel levezetése

Mivel a negyedrendű feladat megoldása viszont lényegesen bonyolultabb, azért a klasszikus (mérnöki) út az, hogy a másodrendű (17.29), (17.30) egyenleteknél maradjunk, a  $\psi$  peremfeltételeiből csak az elsőfajukat választjuk, és ezenkívül  $\omega$  részére még szerzünk peremfeltételeket (ld. a következő pontot), valamint  $p$  részére is, ha a nyomás valóban érdekes.

A kezdeti feltételek nem jelentenek problémát :  $\omega$  és  $\psi$  kezdetiértéke  $u$ -ból és  $v$ -ből előállítható, ezzel  $\Delta\psi$  is megvan – ami elegendő (17.29) vagy akár (17.31) megoldásához.

### 17.2.3 $\omega$ és $p$ peremfeltételei

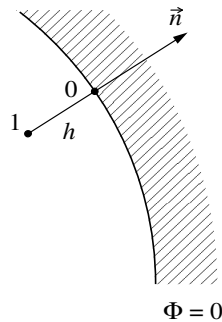
Az  $\omega$  örvényességre vonatkozó peremfeltételek eredetileg nem adottak, levezetésükkel most foglalkozunk (P.J. Roache könyvét követve). Fizikailag arról van szó, hogy külső források nélkül az örvényesség csak az olyan peremszakaszokon keletkezhet, ahol a közeg odatapad a falhoz.

**1.** Tekintsünk egy egyenes, át nem eresztő falszakaszt, amely éppen rajta van pl. az  $x$ -tengelyen, a tartomány belsejében már bevezettünk egy rácsot. Annak  $(x_i, 0)$  pontjában fejtük sorba a többször folytonosan differenciálhatónak feltételezett áramfüggényt :

$$\psi_{i1} = \psi_{i0} + \frac{\partial\psi}{\partial y}h + \frac{h^2}{2} \frac{\partial^2\psi}{\partial y^2} + O(h^3). \quad (17.32)$$

Mivel a peremen  $\psi = \text{const}$ , ha az áramló közeg odatapad, úgy

$$\frac{\partial\psi}{\partial x} = \frac{\partial^2\psi}{\partial x^2} = 0, \quad -\omega = \Delta\psi = \frac{\partial^2\psi}{\partial y^2}, \quad y = 0,$$



ábra 17.5: Thom-féle peremfeltétel, általános irányítás

továbbá  $\frac{\partial \psi}{\partial y}(x, 0) = u(x, 0) = 0$ . Ezért (17.32)-ből következik

$$\psi_{i1} \approx \psi_{i0} + \frac{h^2}{2}(-\omega_{i0}).$$

Innen látjuk, hogy

$$\omega_{i0} = -\frac{2}{h^2}(\psi_{i1} - \psi_{i0}) = -\frac{2}{h}\psi_{y,0}$$

használható  $\omega$  megadására a peremen, ha  $\psi$  ismert. Ez az úgynevezett *Thom-féle peremfeltétel*.

A levezetés szerint, ha  $\psi$  háromszor folytonosan differenciálható  $y$  szerint, ez a peremfeltétel elsőrendű.

Hasonló módon másodrendű peremfeltételt is lehet felírni, ld. a 2. feladatot, de a számítási tapasztalatok szerint a Thom-féle peremfeltétel előnye, hogy stabilabbá teszi a számítást, a másodrendű peremfeltételhez képest. A Thom-féle peremfeltétel alakja a fal irányításától függetlenül felírva ( $\vec{n}$  legyen a külső normálvektor) :

$$\omega_0 = \frac{2}{h}\psi_{\vec{n}}, \quad (17.33)$$

ahol  $\psi_{\vec{n}}$  a  $\frac{\partial \psi}{\partial n}$  normálderivált differencia-approximációja :

$$\psi_{\vec{n}} := \frac{\psi_0 - \psi_1}{h}. \quad (17.34)$$

Másodrendű a következő peremfeltétel is :

$$\omega_0 = \frac{1}{2h^2}(7\psi_0 - 8\psi_1 + \psi_2),$$

ugyanis ismert, hogy  $\frac{\partial\psi}{\partial n}(0) = 0$ , így

$$\begin{aligned}\psi_1 &= \psi_0 + \frac{h^2}{2}(-\omega_0) + \frac{h^3}{6} \frac{\partial^3\psi}{\partial n^3}(0) + O(h^4), \\ \psi_2 &= \psi_0 + 2h^2(-\omega_0) + \frac{4h^3}{3} \frac{\partial^3\psi}{\partial n^3}(0) + O(h^4),\end{aligned}$$

ebből  $\frac{\partial^3\psi}{\partial n^3}(0)$  kifejezhető és  $\psi_1$  egyenletébe behelyettesíthető. Egy másik helyzethez tartoznak a 3. és 4. feladatban konstruálandó peremfeltételek.

**2.** Ha a fal mozog  $U_0$  sebességgel az  $\vec{s}$  érintővektora irányában, akkor

$$\frac{\partial\psi}{\partial y} = U_0,$$

hiszen a folyadék most is tapad a falhoz. Innen, valamint (17.32)-ből következik (17.33) helyett

$$\omega_0 = \frac{2}{h}(\psi_{\vec{n}} + U_0).$$

**3.** Szimmetria vonalon (pl.  $x$ -irányban) érvényes  $v = 0$ ,  $\frac{\partial v}{\partial x} = 0$ . Ugyanakkor  $\frac{\partial u}{\partial y} = 0$ , így  $\omega = 0$ .

**4.** Beáramlásnál adott  $(u, v)_{\Gamma}$ , de általánosabban, mint fent 1. alatt (ha a perem iránya megint az  $x$ -tengely és  $\vec{n} = (0, -1)$ ) most  $v|_{\Gamma} > 0$ , és  $u_0 := u|_{\Gamma} \neq 0$  lehetséges. Ekkor (17.32)-ből következik (17.19) és (17.21) segítségével

$$\psi_{i1} = \psi_{i0} + u_0 h + \frac{h^2}{2} \left( -\omega_0 - \frac{\partial^2\psi}{\partial x^2} \right) + O(h^3) = \psi_{i0} + u_0 h + \frac{h^2}{2} \left( -\omega_0 + \frac{\partial v}{\partial x} \right) + O(h^3).$$

Így, figyelembe véve a (17.34) jelölést és azt, hogy  $\frac{\partial v}{\partial x} = \frac{\partial v}{\partial s}$ , vehetjük az

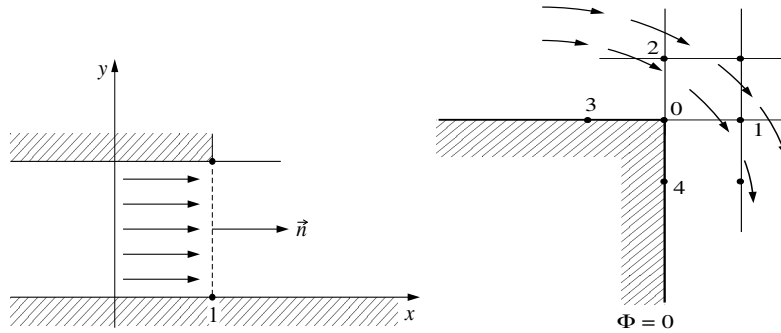
$$\omega_0 = \frac{2}{h}(\psi_{\vec{n}} + u_0) + \left( \frac{\partial v}{\partial s} \right)_0 = \frac{2}{h}[\psi_{\vec{n}} + (\vec{u} \cdot \vec{s})_0] - \left( \frac{\partial(\vec{u} \cdot \vec{n})}{\partial s} \right)_0 \quad (17.35)$$

peremfeltételt.

**5.** Kiáramlásnál :  $\omega_{\vec{n}} = 0$ , Neumann-féle feltétel.

**6.** 90°-os sarokpont: Ha ez belső szöget jelent, legyen  $\omega = 0$ . Ha külső szöget, akkor használhatjuk a (17.30) összefüggés ötponos differencia-approximációját, amelyben  $\psi_3 = \psi_4 = \psi_0$ , mivel az át nem eresztő fal a  $\psi$  izovonala. Így

$$\omega_0 = - \left[ \frac{\psi_1 - \psi_0}{h_1^2} + \frac{\psi_2 - \psi_0}{h_2^2} \right].$$



ábra 17.6: Kiáramlási szakasz és beugró sarok

7. Utolsónak még azt kell tisztáznunk, hogy milyen peremfeltétel írható elő a nyomás számára, amikor azt – az  $\omega - \psi$ -rendszer megoldása után – (17.28)-ből kívánjuk kiszámítani.

Az eredeti egyenletrendszerrel már feltűnt, hogy az a nyomás legfeljebb egy konstans erejéig meghatározható. Ezt a konstanszt rögzíthetjük azzal, hogy valamelyik pontban előírjuk a nyomást. Ezután ebből a pontból kiindulva kiszámíthatjuk a nyomást az eredeti egyenletek segítségével, feltéve, hogy a sebességek már rendelkezésünkre állnak, hiszen ekkor  $\frac{\partial p}{\partial x}$  és  $\frac{\partial p}{\partial y}$  ismert. Ez az eljárás csak akkor ajánlható, ha valamilyen vonal mentén keresett a nyomás, tipikusan egy felület mentén.

Ehhez viszont célszerű az  $\omega - \psi$ -rendszert alapul venni. A Thom-féle peremfeltételhez hasonlóan járunk el.

Mivel az áramló közeg odatapad a felülethez, annak egy rögzített pontjában (ha a felület párhuzamos az  $x$ -tengellyel)

$$u = v = 0, \quad \frac{\partial u}{\partial t} = 0, \quad \frac{\partial u}{\partial x} = \frac{\partial^2 u}{\partial x^2} = 0,$$

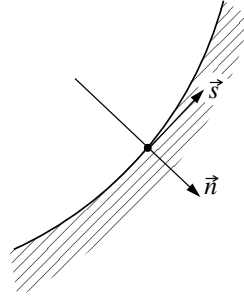
így a (17.21) definíció szerint  $\omega = -\frac{\partial u}{\partial y}$ , és most (17.11) alapján

$$\frac{\partial p}{\partial x} = \nu \frac{\partial^2 u}{\partial y^2} + f_1 = -\nu \frac{\partial \omega}{\partial y} + f_1.$$

Általános irányítás esetén

$$\frac{\partial p}{\partial s} = +\nu \frac{\partial \omega}{\partial n} + \vec{f} \cdot \vec{s}$$

(ill. a (17.14) dimenziótlan alakból :  $\frac{\partial p}{\partial s} = \frac{1}{\text{Re}} \frac{\partial \omega}{\partial n} + \vec{f} \cdot \vec{s}$ ).



ábra 17.7: Normál- és érintővektor a peremen

Ez a képlet eléggé pontatlanul adja a nyomást sarokpontok környezetében. Ekkor a számítási tapasztalatok szerint a következő jobban válik be. Mivel

$$\frac{\partial \omega}{\partial x} = -\frac{\partial^2 u}{\partial y \partial x} = \frac{\partial^2 v}{\partial y^2}, \text{ hiszen } \frac{\partial}{\partial y} \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = 0,$$

a (17.12) egyenletből a fentiekhez hasonló módon kapjuk azt, hogy

$$\frac{\partial p}{\partial y} = \nu \frac{\partial^2 v}{\partial y^2} + f_2 = \nu \frac{\partial \omega}{\partial x} + f_2,$$

ill., a fal irányításától függetlenül,

$$-\frac{\partial p}{\partial n} = \nu \frac{\partial \omega}{\partial s} + \vec{f} \cdot \vec{n},$$

és dimenziótlan alakban

$$\frac{\partial p}{\partial n} = -\frac{1}{\text{Re}} \frac{\partial \omega}{\partial s} - \vec{f} \cdot \vec{n}.$$

Ezzel a Neumann-féle peremfeltétellel egészítjük ki a (17.28) nyomásegyenletet, a perem egy pontjában előírva a nyomás értékét. A most levezetett peremfeltétel helyett gyakran homogén Neumann-feltételt vesznek a nyomás számára, ami az előzőek szerint az  $\vec{f} = 0$ ,  $\text{Re} = \infty$  esetnek felel meg.

**8.** Külső feladatok esetén (testek körüli áramlásoknál) még egy fajta peremfeltétel szükséges:  $u$ ,  $v$  és  $p$  megadása a végtelenben.

A szokásos eljárás abból áll, hogy ezeket a mennyiségeket a számítási rács külső peremén adjuk meg. Kísérletekkel és analitikus megoldásokkal való összehasonlításból kiderült, hogy elfogadható pontosság (pl. a testen fellépő nyomásban) csak akkor érhető el, ha a külső peremet az elvárásnál jóval távolabbra tesszük a testtől, pl. annak legnagyobb méretének a százszoros



távolságra. Továbbá azt is figyelembe kell venni, hogy az elmélet szerint a kétdimenziós esetben a végtelenben valójában nem lehet akármilyen értéket előírni  $\vec{u}$ -ra, hanem csak azt szabad megkövetelni, hogy  $\vec{u}$  legyen véges a végtelenben. Az ott felvett érték függ a térfogati erőktől, a testtől és a rajta előírt sebesség-eloszlástól (ld. O. A. Ladüzsenszkaja könyvét a 60. és 61. oldalon.).

### 17.3 Lassú áramlások számítása

Vizsgáljuk a (17.27) vagy (17.29) egyenletek legegyszerűbb véges differencia-approximációját! Ennek legkézenfekvőbb formája a következő, ha az ekvidisztáns  $\{(x_i, y_j, t_k) = (ih_1, jh_2, k\tau)\}$  rácson (a  $h_1$  és  $h_2$  helybeli lépésközzel, valamint a  $\tau$  időlépéssel) az  $\omega(x, y, t)$  approximációját ugyancsak  $\omega$ -val jelöljük, és ha  $\omega(x_i, y_j, t_k) = \omega_{ij}^k$ :

$$\omega_t + \Lambda_{1/2}(u, v)\omega = g^{k+1/2}, \quad (17.36)$$

ahol ( $\sigma = 1/2$ -del)

$$\Lambda_\sigma(u, v)\omega := \left[ -\nu(\omega_{\bar{x}x} + \omega_{\bar{y}y}) + u\omega_{\bar{x}} + v\omega_{\bar{y}} \right]^{(\sigma)}.$$

Itt, a 16.4.1. pont (16.32) jelölésének megfelelően, a felső indexszel ellátott rácspontok értékeit konvex kombináció segítségével definiáljuk:

$$g^{(\sigma)} := \sigma g^{k+1} + (1 - \sigma)g^k.$$

Vegyük észre, hogy a (17.36) differenciaséma nem tartalmazza az  $(u\omega)_{ij}$  vagy  $(v\omega)_{ij}$  szorzatokat! (Ebben az értelemben azt mondhatjuk, hogy (17.36) „lokálisan lineáris”). Hasonló igaz az  $\omega - \psi$ -fogalmazásban, hiszen az  $u \frac{\partial \omega}{\partial x}$ , ill.  $v \frac{\partial \omega}{\partial y}$  kifejezéseknek megfelelnek (17.19) szerint a

$$\psi_y \omega_{\bar{x}} \quad \text{ill.} \quad -\psi_x \omega_{\bar{y}}$$

differencia-képletek, amelyekkel (17.36)-ból kapjuk

$$\begin{aligned} \omega_t + \Lambda_{1/2}(\psi)\omega &= g^{k+1/2}, & (17.37) \\ \Lambda_\sigma(\psi)\omega &:= \left[ -\nu(\omega_{\bar{x}x} + \omega_{\bar{y}y}) + \psi_y \omega_{\bar{x}} - \psi_x \omega_{\bar{y}} \right]^{(\sigma)}. \end{aligned}$$

Ehhez társul (17.30) kézenfekvő approximációja (ld. (15.16)):

$$-\Delta_h \psi^{k+1} - \omega^{k+1} = 0. \quad (17.38)$$

Az implicit (17.37)-(17.38) rendszer megoldását többrácsos módszerrel képzelhetjük el leginkább. Ekkor az A. Brandt-ra visszamenő, és a számítási tapasztalat által igazolt következő általános észrevétel lényeges, mivel differenciálegyenletek *rendszerét* meg akarunk megoldani :

A simító iteráció lehet a Gauss–Seidel eljárás, de blokkformában, olymódon, hogy *az egy diszkretizációs ponthoz tartozó összes ismeretlen egyszerre kezeljük*. Jelen esetünkben tehát az  $\omega_{ij}$  és  $\psi_{ij}$  értékeit egy Gauss–Seidel blokk-lépéssel egyszerre fel kell újítanunk.

Esetünkben tekintsük az  $(x_i, y_i, t_{k+1})$  pontot! Ha csak a  $t_{k+1}$  időszinthez tartozó kifejezéseket írjuk fel részletesen és minden  $t_k$ -hoz tartozó, azaz ismert mennyiséget  $F_{ij}$  alatt vonjunk össze a jobb oldalon, akkor (17.37) részletes alakja :

$$\begin{aligned} \left( \nu \left( \frac{1}{h_1^2} + \frac{1}{h_2^2} \right) + \frac{1}{\tau} \right) \omega_{ij}^{k+1} - (\omega_{i+1,j} + \omega_{i-1,j})^{k+1} a_{ij}^{(k+1)} \\ - (\omega_{i,j+1} + \omega_{i,j-1})^{k+1} b_{ij}^{(k+1)} = F_{ij}. \end{aligned} \quad (17.39)$$

Itt

$$\begin{aligned} F_{ij} &:= g^{k+1/2} - \left( \nu \left( \frac{1}{h_1^2} + \frac{1}{h_2^2} \right) - \frac{1}{\tau} \right) \omega_{ij}^k \\ &\quad + (\omega_{i+1,j} + \omega_{i-1,j})^k a_{ij}^{(k)} + (\omega_{i,j+1} + \omega_{i,j-1})^k b_{ij}^{(k)}, \\ a_{ij} &= \frac{1}{2} \left( \frac{\nu}{h_1^2} - \frac{\psi_{i,j+1} - \psi_{i,j-1}}{2h_2} \frac{1}{2h_1} \right), \quad b_{ij} = \frac{1}{2} \left( \frac{\nu}{h_2^2} + \frac{\psi_{i+1,j} - \psi_{i-1,j}}{2h_1} \frac{1}{2h_2} \right). \end{aligned} \quad (17.40)$$

(17.38) részletes alakja :

$$\begin{aligned} \left( \frac{2}{h_1^2} + \frac{2}{h_2^2} \right) \psi_{ij}^{k+1} - \frac{1}{h_1^2} (\psi_{i-1,j} + \psi_{i+1,j})^{(k+1)} \\ - \frac{1}{h_2^2} (\psi_{i,j+1} + \psi_{i,j-1})^{(k+1)} - \omega_{ij}^{k+1} = 0. \end{aligned} \quad (17.41)$$

Ezek szerint a Gauss–Seidel eljárás szempontjából az  $(x_i, y_i, t_{k+1})$  rácspontban a következő (lineáris!) egyenletekkel állunk szemben :

$$\left( \nu \left( \frac{1}{h_1^2} + \frac{1}{h_2^2} \right) + \frac{1}{\tau} \right) \omega_{ij}^{k+1} = f_{ij}, \quad (17.42)$$

$$\left( \frac{2}{h_1^2} + \frac{2}{h_2^2} \right) \psi_{ij}^{k+1} - \omega_{ij}^{k+1} = g_{ij}, \quad (17.43)$$

ahol az  $f_{ij}$ , ill.  $g_{ij}$  mennyiségek a (17.39) ill. (17.41) jobboldalán kívül még az  $\omega^{k+1}$  és  $\psi^{k+1}$  értékeit az  $(x_i, y_i)$ -től különböző helyeken is tartalmazzák.

Vegyük először a (17.42), majd a (17.43) egyenletet szokásos Gauss–Seidel eljárással dolgozva, mert ha ilyen sorrendben iteráljuk az egyenleteket, akkor mindig a (17.42), (17.43) lokális rendszerek pontos megoldását állítjuk elő, vagyis ugyanazt kapjuk, mint a  $2 \times 2$ -es blokkokkal működő blokk-Gauss–Seidel módszerrel (amelynek használatára itt szükség van, ahogyan azt fent említettük).

A perempontokban nincs különbség ilyen szempontból, mert ott (17.42), (17.43) helyett csak

$$\omega_{ij}^{k+1} = f_{ij} \quad (17.44)$$

alakú egyenletekkel találkozunk : vagy  $\psi_{ij}$  adott a peremen (át nem eresztő fal, beáramlási szakasz), és ekkor a Thom-féle peremfeltétel vagy a másodrendű peremfeltételek érvényesek  $\omega$ -ra vagy (17.35) – és ezek mind a Gauss–Seidel eljárás szempontjából csakis (17.44) alakúak – kivéve azt az egy esetet, amikor kiáramlási szakaszon vagyunk.

Csupán itt van szükség arra, hogy  $\psi$  értékét is iteráljuk, de ez ilyenkor független az  $\omega$  peremfeltételétől. A sorrend lehet  $\omega \rightarrow \psi$  mint ezelőtt, és így végig megyünk az összes ponton.

Az  $\omega - \psi$ -rendszer megoldása így világosnak tűnik – de csak addig az, míg a Reynolds-szám kicsi. Ilyen feladatok is vannak (kenőcsök, polimérek vagy olvasztott üveg áramlása – ekkor bonyolult peremfeltételek is felléphetnek, amelyek a felületi feszültség hatását írják le – valamint környezetvédelmi problémák : éppen kevés légmozgás esetén veszélyes koncentrációk képződhetnek). De az ipari áramlások nagyobbik része nagy Reynolds-számmal kapcsolatos, ld. a táblázatot 17.1-ben.

A II. 11.4.10. pontban az  $au'' + bu' + f = 0$ ,  $0 < x < 1$ ,  $u(0) = u(1) = 0$  mintafeladatot elemeztük. Amennyiben ott igaz  $|b| \gg a$  (és ez megfelel a nagy Reynolds-számnak) akkor kimutattuk, hogy egyáltalán nem megfelelő a  $bu'$  tag központi differencia-approximációja – amire viszont (17.37)-ben és (17.41)-ben támaszkodtunk.



## Fejezet 18

# Kiegészítő információk, jelölések III

Legyen  $\Omega$  véges tartomány az  $\mathbb{R}^d$ -ben ( $1 \leq d$ ), tekintsük azokat az  $\Omega$ -ban definiált, négyzetben integrálható függvényeket, amelyeknek  $\int_{\Omega} u^2(x) dx$  integrálja véges. Ezek az  $L_2(\Omega)$  függvényteret alkotják, amely Hilbert-tér skalárszorzata

$$(u, v)_0 := \int_{\Omega} u(x)v(x) dx,$$

normája :

$$\|u\|_0^2 := \int_{\Omega} u^2(x) dx.$$

Ezenkívül az alapvető téren kívül a szövegben különféle Szoboljev-térek is használunk (ezekhez ld. részletesebben Simon L. és Baderko vagy Adams könyvét). Számunkra fontos mindenekelőtt a  $H^1(\Omega)$ -val jelölt Hilbert-tér. Ennek elemei azon  $\Omega$ -ban definiált, deriváltjaival együtt négyzetben integrálható függvények, amelyeknek  $\int_{\Omega} \left[ \sum_{i=1}^d \left( \frac{\partial u}{\partial x_i} \right)^2 + u^2 \right] dx$  integrálja véges. A tér skalárszorzata

$$(u, v)_1 := \int_{\Omega} \left[ \sum_{i=1}^d \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} + uv \right] dx,$$

normája :

$$\|u\|_1^2 := \int_{\Omega} \left[ \sum_{i=1}^d \left( \frac{\partial u}{\partial x_i} \right)^2 + u^2 \right] dx = \int_{\Omega} [|\text{grad } u|^2 + u^2] dx,$$

ahol  $|\text{grad } u|$  jelöli a  $\text{grad } u$  vektor euklideszi normáját.

A  $H_0^1(\Omega)$  tér normája viszont

$$\|u\|_1^2 := \int_{\Omega} |\text{grad } u|^2 dx,$$

ez  $H^1(\Omega)$  félnormája. Hasonló a  $H^2(\Omega)$  tér definíciója, amelynek normája :

$$\|u\|_2^2 := \int_{\Omega} \sum_{i,j=1}^d \left| \frac{\partial^2 u}{\partial x_i \partial x_j} \right|^2 dx + \|u\|_1^2.$$

A  $H^k$  tér felső indexe tehát azon deriváltak számát adja meg, amellyel függvényei (általánosított értelemben) rendelkeznek. Ha  $k$  tört szám, akkor definícióját II-ben lehet találni (280. o.). Ilyen terek itt főként a peremértékek miatt érdekesek, amikor  $k + 1/2$  egész. Az  $\Omega$  peremét  $\Gamma$ -val jelöljük.

A folytonos beágyazást a  $\hookrightarrow$  nyíllal jelöljük, így pl. igaz a  $H^2 \hookrightarrow C$  beágyazás, ha  $1 \leq d \leq 3$ , vagyis : ekkor a  $H^2$ -beli függvények (durván fogalmazva) egyben folytonosak is, és van olyan (csak  $\Omega$ -tól függő)  $c_0$  konstans, amellyel igaz a következő becslés :

$$\|v\|_C \leq c_0 \|v\|_{H^2} \text{ minden } v \in H^2\text{-re.}$$

A következő táblázat azokat a feltételeket adja meg, hogy mikor hová beágyazható a  $H^k(\Omega)$ , amikor  $\Omega \in \mathbb{R}^d$  és  $\Gamma_s$  az  $\Omega$ -nak egy sima,  $s$ -dimenziós felülete.

feltétel	$d - 2k \leq s$	$d - 2k < s$	$d - 2k < 0$
beágyazás $L_2(\Gamma_s)$ -be:	folytonos	kompakt	beágyazás $C(\Omega)$ -ba

A kompakt beágyazás pl. a sajátérték feladattal kapcsolatban érdekes.

Ha  $u$   $k$ -szor folytonosan deriválható függvény  $\Omega$ -ban, akkor  $M_k$ -val jelöljük az összes  $k$ -adik deriváltjainak maximumát  $\Omega$ -n.

$\bar{\omega}_h = \omega_h \cup \gamma_h$  jelöli a téglalap alakú rácsot,  $\gamma_h$  az  $\bar{\omega}_h$  rács perempontjait, ld. 15.3.1., ahol ezen halmazok definíciója található abban az esetben, amikor  $\Omega$  téglalap. Ekkor  $\gamma_0$ -val jelöljük a sarokpontok halmazát. Az  $\bar{\omega}_h$ -n definiált függvényeket rácsfüggvényeknek hívjuk. Skalárszorzatainak definíciójánál azt a konvenciót tartjuk be, hogy az összegek mindig az összes olyan indexre térnek ki, amelyre az összegzett kifejezések definiáltak.

A differenciahányadosok jelöléseit az időderiváltakat approximáló differenciahányadosok példáján mutatjuk meg itt (ld. részletesebben II-ben, 302-303. o.) :

$$\begin{aligned} y_t &= y_{t,i} = y_{i,i}^j := (y_i^{j+1} - y_i^j)/\tau, & y_{\bar{t}} &:= y_{\bar{t},i}^j := (y_i^j - y_i^{j-1})/\tau, \\ y_t^o &:= y_{t,i}^j := (y_i^{j+1} - y_i^{j-1})/(2\tau), \\ y_{\bar{t}\bar{t}}^j &= y_{\bar{t}\bar{t}} := \frac{1}{\tau^2}(\hat{y} - 2y + \check{y})_i, & \hat{y} &:= y_i^{j+1}, \quad y := y_i^j, \quad \check{y} := y_i^{j-1}. \end{aligned}$$

A helykoordinátákra vonatkozó differenciahányadosokkal kapcsolatban a következő, már II-ben is használt szabályt tartjuk be :

„egész” pontokban ( $x_i$ -ben) definiált rácsfüggvény differenciahányadosát  $x$ -szel jelöljük, az eredményt „feles” ponthoz ( $x_{i\pm 1/2}$ -hez) rendeljük hozzá.

Megfordítva, „feles” pontokban definiált rácsfüggvény differenciahányadosát  $\tilde{x}$ -mal jelöljük, az eredményt „egész” ponthoz rendeljük hozzá. Így például a változó együttható és lépéstávolság esetén használjuk a  $\frac{d}{dx} \left( k(x) \frac{du}{dx} \right)$  következő approximációját :

$$\begin{aligned} (ky_x)_{\tilde{x}} &:= \frac{2}{h_{i+1/2} + h_{i-1/2}} \left( (ky_x)_{i+1/2} - (ky_x)_{i-1/2} \right) \\ &= \frac{2}{h_{i+1/2} + h_{i-1/2}} \left( k(x_{i+1/2}) \frac{y(x_{i+1}) - y(x_i)}{h_{i+1/2}} \right. \\ &\quad \left. - k(x_{i-1/2}) \frac{y(x_i) - y(x_{i-1})}{h_{i-1/2}} \right), \end{aligned}$$

ahol

$$h_{i-1/2} := x_i - x_{i-1}, \quad h_{i+1/2} := x_{i+1} - x_i, \quad x_{i\pm 1/2} := \frac{x_{i\pm 1} + x_i}{2}.$$

Kétdimenziós egyenletek approximációja téglalapos ekvidisztáns rácson gyakran úgy írható fel áttekinthetőbb és rövidebb módon, hogy a kifejezésben változatlanul szereplő  $j$  második, ill.  $i$  első indexet csak az utolsó zárjel mögött írjuk fel. Erre itt álljon mint példa a Laplace-operátor standard differencia-approximációja :

$$\begin{aligned} (\Delta_h u)_{ij} &:= (u_{x_1 \tilde{x}_1} + u_{x_2 \tilde{x}_2})_{i,j} \\ &= \frac{1}{h_1} [(u_{x_1})_{i+1/2} - (u_{x_1})_{i-1/2}]_j + \frac{1}{h_2} [(u_{x_2})_{j+1/2} - (u_{x_2})_{j-1/2}]_i \\ &= \left( \frac{u_{i+1} - 2u_i + u_{i-1}}{h_1^2} \right)_j + \left( \frac{u_{j+1} - 2u_j + u_{j-1}}{h_2^2} \right)_i \\ &= \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h_1^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h_2^2} \\ &= (u_{\tilde{x}_1 x_1} + u_{\tilde{x}_2 x_2})_{i,j}. \end{aligned}$$

A megfelelő peremfeltételek mellett kitűzött peremérték feladat (klasszikus)  $u$  megoldásának  $(x_{1i}, x_{2j})$ -beli  $u_{ij}$  értékeinek vektorát  $\vec{u}_h$ -val jelöljük (az  $u \rightarrow \vec{u}_h$  egy leképezés  $C(\bar{\Omega})$ -ből bizonyos  $R^n$ -be). A differencia egyenletek megoldását, az  $u_{ij}$ -ket approximáló értékeket gyakran  $y_{ij}$ -vel jelöljük.

Ezeknek vektora  $y_h$ , ill., ha azt akarjuk hangsúlyozni, hogy az elsőfajú peremfeltételek még nincsenek eliminálva és hozzátartoznak ehhez a vektorhoz, akkor  $\bar{y}_h$  a jelölés. A differencia egyenleteket összefoglaló lineáris rendszer mátrixa gyakran  $A_h$  (maga a rendszer pl.  $A_h y_h = \varphi_h$ ), ill.  $\bar{A}_h$ , ha az elsőfajú peremfeltételeknek megfelelő triviális egyenletek is hozzátartoznak a rendszerhez.

Ugyanakkor az  $y_{ij}$ -ket tekinthetjük, mint az  $y$  rácsfüggvény értékeit :  $y_{ij} = y(x_{1i}, x_{2j})$ . Differencia kifejezésekben a rácsfüggvény szerinti értelmezés előnyesebb, ott megtakarítjuk a  $h$  indexet. Mátrixokkal együtt vagy normákban inkább az  $y_h$  vektort írjuk (ezzel rámutatva, hogy itt valójában  $h$ -val indexelt seregekről van szó), de pl. instacionárius feladatoknál (amikor  $y_h$   $t$ -től is függ :  $y_h(t_j) = y_h^j = y_h$ ) vegyes jelölés is előfordul :

$$y_t + A_h y_h = \varphi_h, \text{ ahelyett, hogy } \frac{1}{\tau} (y_h^{j+1} - y_h^j) + A_h y_h^j = \varphi_h^j.$$

A végeselem módszer által előállított közelítő megoldás  $u_h = \sum_{j=1}^N y_j w_j$ , ahol a  $w_j$ -k a bázisfüggvények és  $y_j$  azoknak együtthatói. Az  $y_j$ -k vektora  $y$ , amely az  $Ay = b$  rendszer megoldása, ahol az  $A$  mátrix a bilineáris  $a$  formából adódik :  $A = (a_{k\ell})$ ,  $a_{k\ell} = a(w_\ell, w_k)$ . Ha a peremérték feladat  $d$ -dimenziós, akkor  $A$  és  $h^d A_h$  összehasonlíthatóak, ahol  $A_h$  egy megfelelő differencia approximáció mátrixa. Speciális esetben  $A = h^d A_h$  is előfordul.

$\mathbb{R}^d$ -ben a  $k$ -adfokú polinomok  $P_k(\Omega_i)$  lineáris terét (amely tipikus a háromszöges végeselemekre) a  $Q_k(\Omega_i)$  polinomtértől különböztetjük meg, amely tipikus a négyszöges elemekre. Az előbbi azon  $x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_d^{\alpha_d}$  alakú tagok lineáris kombinációja, melyekre  $0 \leq \alpha_1 + \dots + \alpha_d \leq k$ , míg az utóbbinál  $0 \leq \alpha_\ell \leq k$  a feltétel ( $\ell = 1, \dots, d$ ).



# Fejezet 19

## Irodalom III

*M. Adams*, Evaluation of three unstructured multigrid methods on 3D finite element problems in solid mechanics. *Int. J. Numer. Methods Eng.* 55, 5 (2002), 519–534.

*R.A. Adams*, Sobolev Spaces. Academic Press, New York 1975.

*M. Ainsworth, J.T. Oden*, A Posteriori Error Estimation in Finite Element Analysis. Wiley-Interscience, New York 2000.

*W.F. Ames*, Nonlinear Partial Differential Equations in Engineering, v. I, II. Academic Press, New York 1965, 1972.

*A.B. Andreev*, The lumped mass finite element method for some elliptic eigenvalue problems. *Proc. Conf. Optimal Algorithms* (Bl. Sendov, ed.), Sofia 1986, 6–14.

*W. Anzinger*, A quantitative discrete  $H^2$  regularity estimate for the Shortley-Weller scheme in convex domains. *Numer. Math.* 52 (1988), 523–537.

*A. Arakawa, F. Mesinger*, Numerical Methods Used in Atmospheric Models. GARP Public. Series 17, WMO 1978.

*D.N. Arnold, D. Boffi, R.S. Falk*, Approximation by quadrilateral finite elements. *Math. Comp.* 71, 239 (2002), 909–922.

*D.N. Arnold, F. Brezzi, M. Fortin*, A stable finite element for the Navier–Stokes equations. *Calcolo* 23 (4), (1984), 337–344.

*J.-P. Aubin*, Approximation of Elliptic Boundary Value Problems. Wiley, New York 1972.

*O. Axelsson, V.A. Barker*, Finite Element Solution of Boundary Value Problems : Theory and Computation. Academic Press, New York 1984.

*O. Axelsson, S.V. Gololobov*, Stability and error estimates for the  $\theta$ -method for strongly monotone and infinitely stiff evolution equations. *Numer. Math.* 89,1 (2001), 31–48.

*I. Babuška*, Error bound for finite element method. Numer. Math. 16 (1971), 322–333.

*I. Babuška et al. (eds.)*, Accuracy Estimates and Adaptive Refinements in Finite Element Computations. J. Wiley, Chichester 1986.

*I. Babuška, J. E. Osborn*, Estimates for the Errors in Eigenvalue and Eigenvector Approximation by Galerkin Methods, with Particular Attention to the Case of Multiple Eigenvalues. SIAM Journal on Numerical Analysis, Vol. 24, No. 6 (1987), 1249–1276.

*I. Babuška*, Advances in the  $p$  and  $h - p$  versions of the finite element method. A survey. In : Numerical Mathematics Singapore 1988 (R.P. Agarwal, Y.M. Chow, S.J. Wilson, eds.), Birkhäuser, Basel 1988, 31–46.

*I. Babuška, J.E. Osborn*, Finite element–Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems. Math. Comp. 52, (1989), 275–297.

*D. Bahlmann, U. Langer*, A fast solver for the first biharmonic boundary value problem. Numer. Math. 63 (1992), 297–313.

*N.Sz. Bahvalov*, Egy relaxációs módszer konvergenciájáról, az elliptikus operátorra nézve természetes mellékfeltételek esetén. Zs. Vücs. Matem. i Matem. Fiz. 6, 5 (1966), 861–883 (oroszul).

*G.A. Baker, J.H. Bramble*, Semidiscrete and single step full discrete approximations for second order hyperbolic equations. RAIRO Anal. Numér. 13 (1979), 75–100.

*R. Balder, C. Zenger*, The solution of multidimensional real Helmholtz equations on sparse grids. SIAM J. Sci. Comput. 17, No.3 (1996), 631–646.

*R.E. Bank, A. Sherman, A. Weiser*, Refinement algorithms and data structures for regular local mesh refinement. In : Stepleman R. (ed.) Scientific Computing. North-Holland, Amsterdam 1983, 3–17.

*R.E. Bank, A. Weiser*, Some a-posteriori error estimators for elliptic partial differential equations. Math. Comp. 44 (1985), 283–301.

*R.E. Bank*, Analysis of a local a posteriori error estimate for elliptic equations. Chapter 7 in : I. Babuška et al. Accuracy Estimates and Adaptive Refinements in Finite Element Computations. J. Wiley, New York 1986.

*R.E. Bank, T.F. Chan*, PLTMGC: a multi-grid continuation program for parameterized nonlinear elliptic systems. SIAM J. Sci. Stat. Comput. 7,2 (1986), 540–559.

*R.E. Bank, T.F. Dupont, H. Yserentant*, The hierarchical basis multigrid method. Numer. Math. 52 (1988), 427–458.

*J.B. Bell, C.N. Dawson, G.R. Shubin*, An unsplit, higher order Godunov method for scalar conservation laws in multiple dimensions. *J. Comp. Phys.* 74 (1988), 1–24.

*G. Ben-yu, Z. Xiao-yong*, Spectral method for differential equations of degenerate type on unbounded domains by using generalized Laguerre functions. *App. Numer. Math.* 57 (2007), 455–471.

*J. Bey*, Simplicial grid refinement: on Freudenthal’s algorithm and the optimal number of congruence classes. *Numer. Math.* 85 (2000), 1–29.

*J.J. Blair*, Bounds for the change in the solutions of second order elliptic partial differential equations. *SIAM J. Appl. Math.* 24,3 (1973), 277–285.

*V.S. Borisov*, On discrete maximum principles for linear equation systems and monotonicity of difference schemes. *SIAM J. Matrix Anal. Appls.* 24, 4 (2003), 1110–1135.

*A. Bowyer*, Computing Dirichlet tessellations. *Comput. J.* 24 (1981), 162–166.

*D. Braess, W. Hackbusch*, A new convergence proof for the multigrid method including the V-cycle. *SIAM J. Numer. Anal.* 20,5 (1983), 967–975.

*D. Braess*, Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics. 2nd. ed., Cambridge Univ. Press 2001.

*D. Braess, R. Sarazin*, An efficient smoother for the Stokes problem. *Appl. Numer. Math.* 23 (1997), 3–19.

*J.H. Bramble, J.E. Pasciak, A.H. Schatz*, The construction of preconditioners for elliptic problems by substructuring, I–IV. *Math. Comp.* 47 (1986), 103–134; 49 (1987), 1–16; 51 (1988), 415–430; 53 (1989), 1–24.

*A. Brandt*, Guide to Multigrid Programming. Lecture Notes in Math. 960, Springer 1982.

*A. Brandt*, Guide to CFD Multigrid Programming. GMD-Studie, St. Augustin 1984.

*A. Brandt, S. McCormick, J. Ruge*, Multigrid methods for differential eigenproblems. *SIAM J. Sci. Stat. Comput.* 4 (1983), 244–260.

*S.C. Brenner, L.R. Scott*, The Mathematical Theory of Finite Element Methods. Springer-Verlag, New York 1994.

*F. Brezzi, M. Fortin*, Mixed and Hybrid Finite Element Methods. Springer, New York 1991.

*W.L. Briggs, E. van Henson, S.F. McCormick*, A Multigrid Tutorial. 2nd. ed., SIAM, Philadelphia 2000.

*H-J Bungartz, M. Griebel*, Sparse grids. *Acta Numerica* 13 (2004), 147–269.

*C. Canuto, M.Y. Hussaini, A. Quarteroni*, Spectral Methods in Fluid Dynamics. Springer, Berlin 1987.

*F. Chatelin*, Spectral Approximations of Linear Operators. Academic Press, New York 1983.

*P.G. Ciarlet*, The Finite Element Method for Elliptic Problems. Amsterdam, North-Holland 1978.

*P.G. Ciarlet*, Discrete maximum principle for finite-difference operators. *Aequat. Math.* 4 (1970), 338–352.

*P.G. Ciarlet, P.A. Raviart*, Maximum principle and uniform convergence for the finite element method. *Comp. Meth. Appl. Mech. Eng.* 2 (1973), 17–31.

*Ph. Clément*, Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.* 9 (1975), 77–84.

*P. Colella*, Multidimensional upwind methods for hyperbolic conservation laws. *J. Comp. Phys.* 87 (1990), 171–200.

*M. Crouzeix, P.A. Raviart*, Conforming and non-conforming finite elements for solving the stationary Stokes equations. *RAIRO Anal. Numér. Method.* 7 (1973), 33–76.

*R. Courant, K.O. Friedrichs, H. Lewy*, Über die Differenzengleichungen der Mathematischen Physik. *Math. Annalen*, 100 (1928), 32–74.

*R. Courant, D. Hilbert*, Methoden der Mathematischen Physik II, 2. Auflage, Springer-Verlag, Berlin 1968.

*J.K. Cullum, R.A. Willoughby*, Lanczos Algorithms for Large Symmetric Eigenvalue Computations, v. I (Theory), II (Programs). Birkhäuser, Basel 1985.

*J.W. Demmel*, Applied Numerical Linear Algebra. SIAM, Philadelphia 1997.

*G. Denk*, A new numerical method for the integration of highly oscillatory second-order ordinary differential equations. *APNUM* 13 (1993), 57–67.

*P. Destuynder, B. Métivet*, Explicit error bounds in a conforming finite element method. *Math. Comp.* 68, 228 (1999), 1379–1396.

*J.G. Djakonov*, Többdimenziós instacionárius feladatok megoldását szolgáló, faktorizált operátorú differenciálsémákról. *Zs. Vücs. Mat. i Mat. Fiz.* 2,4 (1962), 549–568 (oroszul).

*J.G. Djakonov*, A Stokes–operátorú peremérték feladatok megoldásához szükséges aritmetikai műveletek számának becslése. *Izv. Vuzov, Matematika*, No. 7 (1983), 46–58 (oroszul).

*J.G. Djakonov*, A numerikus műveletek minimalizálása. Elliptikus egyenleteket megoldó, aszimptotikusan optimális algoritmusok. Nauka, Moszkva (1989) (oroszul).

*M. Dryja*, Metoda Galerkina przemiennych kierunków dla quasiliniowych równan parabolicznych. *Roczniki Polsk. Tow. Mat., Ser. III: Matematyka Strowana XV* (1979) 5–23.

*M. Dryja*, A capacitance matrix method for Dirichlet problems on polygonal regions. *Numer. Math.* 39 (1982), 51–64.

*A. Ecker, W. Zulehner*, On the smoothing property for the non-symmetric case. University Linz, Inst.-ber. Nr. 489, Linz 1995.

*N. Egidí, P. Maponi*, A robust direct variational approach for generation of quadrangular and triangular grids on planar domains. *Math. Comput. Simul.* 75, No. 5-6 (2007), 171–181.

*B. Einfeldt*, Ein schneller Algorithmus zur Lösung des Riemann-Problems. *Computing* 39 (1987), 77–86.

*B. Engquist, A. Majda*, Radiation boundary conditions for acoustic and elastic wave calculations. *Comm. Pure Appl. Math.* 32 (1979), 313–357.

*Faragó I., H. Hariton, Komáromi N., Pfeil T.*, A hővezetési egyenlet és numerikus megoldásának kvalitatív tulajdonságai. I. Az elsőfokú közelítések nemnegativitása. *Alk. Matem. Lapok* 17 (1993), 101–121; II. A másodfokú közelítés nemnegativitása, a maximumelv és az oszcillációmentesség. *Alk. Matem. Lapok* 17 (1993), 123–141.

*R.P. Fedorenko*, Diszkrét elliptikus egyenleteket megoldó relaxációs módszer. *Zs. Vücs. Matem. i Mat. Fiz.* 1,5 (1961), 922–927 (oroszul).

*R.P. Fedorenko*, Egy iterációs módszer konvergencia sebességéről. *Zs. Vücs. Matem. i Mat. Fiz.* 4,5 (1964), 559–564 (oroszul).

*R.P. Fedorenko*, Diszkrét elliptikus egyenleteket megoldó iterációs módszerekről. *Uszpehi Matem. Nauk* 28,2 (1973), 121–182 (oroszul).

*M. Feistauer*, Computational Fluids Dynamics. Longman Scientific & Technical, Harlow 1993.

*M. Fortin*, An analysis of the convergence of mixed finite element methods. *RAIRO Anal. Numér.* 11, 3 (1977), 341–354.

*M. Fortin*, Finite element solution of the Navier–Stokes equations. *Acta Numerica* 1993, 239–284.

*I. Fried, D.S. Malkus*, Finite element mass matrix lumping without convergence rate loss. *Int. J. Solids & Structures* 11 (1976), 461–466.

*A. Friedman*, *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs 1964.

*O. Friedrich*, A new method for generating inner points of triangulations in two dimensions. *Comp. Meth. Appl. Mech. Eng.*, 104 (1993), 77–86.

*I.V. Frjazinov*, A Poisson-egyenletet approximáló differenciasémákról, polár-, hengersizmetrikus és gömbszimmetrikus koordinátá-rendszerekben. *Zs. Vücs. Matem. i Matem. Fiz.* 11,5 (1971), 1219–1228 (oroszul).

*G.P. Galdi*, *An Introduction to the Navier-Stokes Equations I: Linearized Steady Problems, II: Nonlinear Steady Problems*. Springer, New York 1994.

*B.G. Galjorkin*, *Vesztnyk Inzsenera* No. 19 (1915), 897–908.

*Gáspár Cs., Józsa J.*, Two-dimensional Lagrangian flow simulation using fast quadtree-based adaptive multigrid solver. *Proc. 9th GAMM Conf.* 1991, Vieweg-Verlag 1992.

*Gáspár Cs.*, *Elliptikus peremérték feladatok perem típusú numerikus megoldási módszerei*. Thesis, Budapest 1993.

*Gáspár Cs., Józsa J., P. Simberowicz*, Új szemléletmód a numerikus hidraulikában. *Hidrológiai Közlöny* 3. és 4. sz. (1994), 1. és 2. sz. (1995).

*U. Ghia, K.N. Ghia, C.T. Shin*, High-Re solutions for incompressible flow using the Navier–Stokes equations and a multigrid method. *J. Comp. Phys.* 48 (1982), 387–411.

*V. Girault, P.A. Raviart*, *Finite Element Methods for Navier–Stokes Equations*, Springer 1986.

*R. Glowinski, J. Periaux*, *Numerical methods for nonlinear problems in fluid dynamics*. *Proc. Int. Seminar on Sci. Supercomputers*. North-Holland, Amsterdam 1987.

*S.K. Godunov*, Differencia-módszer a gázdinamikai egyenletek numerikus megoldására. *Mat. Zbornik* 47 (89), 3 (1959), 271–306 (oroszul).

*S.K. Godunov, G.P. Prokopov*, A konform leképezések kiszámításáról és a differencia rácsok szerkesztéséről. *Zs. Vücs. Matem. i Matem. Fiz.* 7,5 (1967), 1031–1059 (oroszul).

*Sz.K. Godunov, V.Sz. Rjabenyki*, *Differenciasémák*. Nauka, Moszkva 1977 (oroszul).

*Sz.K. Godunov*, *A Matematikai Fizika Egyenletei*. Nauka, Moszkva 1979 (oroszul).

*H. Goering, H.-G. Roos, L. Tobiska*, Die Finite-Element-Methode. Akademie-Verlag, Berlin 1988.

*V.Ja. Golgyin, N.N. Kalitkin, T.V. Sisova*, Nemlineáris differenciasémák hiperbolikus egyenletek megoldására. Zs. Vücs. Matem. i Matem. Fiz. 5, 5 (1965), 938–944 (oroszul).

*A.L. Goncsarov, I.V. Frjazinov*, A Navier–Stokes egyenleteket megoldó differenciasémák kilencpontos keresztalakú differenciacsillagon. Zs. Vücs. Matem. i Matem. Fiz. 28, 6 (1988), 867–878 (oroszul).

*P.M. Gresho, R.L. Sani*, Incompressible Flow and the Finite Element Method. I: Advection - Diffusion, Isothermal Laminar Flow. J. Wiley, Chichester 1998.

*Ch. Großmann, H.-G. Roos*, Numerik partieller Differentialgleichungen. Teubner, Stuttgart 1992. A 3. német kiadás angol fordítása : *Ch. Grossmann, H.-G. Roos, M. Stynes*, Numerical treatment of partial differential equations. Universitext. Springer Verlag, Berlin 2007.

*A.V. Gulin*, Néhány nemönadjungált háromréteges séma stabilitási kritériumai. Diff. Urav. XVI,7 (1980), 1205–1210 (oroszul).

*A.V. Gulin, A.A. Szamarszkij*, A differenciasémák stabilitási elméletének néhány eredményéről és problémájáról. Mat. Zborn. 99 (141), (1976), 290–330 (oroszul).

*A.V. Gulin, I.V. Frjazinov*, A hővezetési egyenletet megoldó, váltakozó irányok módszereinek pontosságáról, általános tartomány esetén. Diff. Uravnyenija XII, 10 (1976), 1906–1914 (oroszul).

*Qi Guo-Ben*, The  $h - p$  version of the finite element method for elliptic equations of order  $2m$ . Numer. Math. 53, (1988), 1-2, 199-224.

*V.A. Guscsin*, Henger körüli, leváló viszkózus áramlások numerikus vizsgálata. Szovjet Akadémia Számítóközpontja, Moszkva 1985 (oroszul).

*W. Hackbusch*, On the computation of approximate eigenvalues and eigenfunctions of elliptic operators by means of a multi-grid method. SIAM J. Numer. Anal. 16 (1979), 201–215.

*W. Hackbusch, U. Trottenberg (eds.)*, Multigrid Methods. Lecture Notes in Mathematics 960. Springer, Berlin 1982.

*W. Hackbusch*, Multi-Grid Methods and Applications. Springer, Berlin 1985.

*W. Hackbusch, U. Trottenberg (eds.)*, Multigrid Methods II. LNM 1228. Springer, Berlin 1986.

*W. Hackbusch*, Theorie und Numerik elliptischer Differentialgleichungen. Teubner, Stuttgart 1986.

*W. Hackbusch*, On first and second order box schemes. *Computing* 41 (1989), 277–296.

*W. Hackbusch, A. Reusken*, Analysis of a damped nonlinear multilevel method. *Numer. Math.* 55,2 (1989), 225–246.

*G. Haase, U. Langer, A. Meyer*, The approximate Dirichlet domain decomposition method. I.: An algebraic approach. *Computing* 47 (1991), 137–151; II: Applications to 2nd order elliptic boundary value problems. *Computing* 47 (1991), 153–157.

*G. Haase, U. Langer, A. Meyer, S.V. Nepomnyaschih*, Hierarchical extension and local multigrid methods in domain decomposition preconditioners. Preprint SPC 94–8, TH Chemnitz 1994.

*F.H. Harlow, J.E. Welch*, Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Phys. Fluids* 8 (1965), 2183–2189.

*Hegedűs Cs.*, Conjugate gradient methods for general matrices. Diszertáció, Magyar Tudományos Akadémia 1995.

*B. Heinrich*, Finite Difference Methods On Irregular Networks. Akademie-Verlag, Berlin 1987.

*R.L. Higdon*, Numerical absorbing boundary conditions for the wave equation. *Math. Comp.* 49,179 (1987), 65–90.

*R.W. Hockney*, A fast direct solution of Poisson's equation using Fourier analysis. *J. ACM* 12 (1965), 95–113.

*P. Hood*, Frontal solution program for unsymmetric matrices. *Int. J. Num. Meth. Eng.* 10 (1976), 379–398.

*E. Hopf*, Elementare Bemerkungen über die Lösungen partieller Differentialgleichungen vom elliptischen Typus. *Sitzungsberichte d. Preußischen Akad. Wiss. Berlin*, 19 (1927), 147–152.

*E. Hopf*, Über die Anfangswertaufgabe für die hydrodynamischen Grundgleichungen. *Math. Nachr.* 4 (1951), 213–231.

*T.J.R. Hughes*, The Finite Element Method: Linear Static and Dynamic Finite Element Analysis. Prentice-Hall, Englewood Cliffs 1987.

*A. Iserles, G. Strang*, The optimal accuracy of difference schemes. *Trans. Amer. Math. Soc.* 277,2 (1983), 779–803.

*Istvánffy E.*, Tápvonalak, antennák, hullámterjedés. Tankönyvkiadó, Budapest 1967.

*N.N. Janenko*, Egy, a többdimenziós hővezetési egyenletet megoldó differenciasémáról, *Doklady AN SSSR*, 125, 6 (1959) 1207–1210 (oroszul).



*V. John, P. Knobloch, G. Matthies, L. Tobiska*, Non-nested multi-level solvers for finite element discretisations of mixed problems. *Computing* 68, 4 (2002), 313–341.

*V. John, L. Tobiska*, A coupled multigrid method for nonconforming finite element discretizations of the 2D-Stokes equation. *Computing* 64 (2000), 307–321.

*V. John*, Higher order finite element methods and multigrid solvers in a benchmark problem for the 3D-Navier-Stokes equations. *Int. J. Numer. Meth. Fluids* 40 (2002), 775–798.

*B.S. Jovanović, L.D. Ivanović, E.E. Süli*, Convergence of finite-difference schemes for elliptic equations with variable coefficients. *IMA. J. Numer. Anal.* 7 (1978), 301–305.

*J. Kadlec*, A Poisson-egyenlet megoldásának regularitásáról olyan tartomány esetén, amelynek pereme lokálisan konvex tartomány peremére hasonló. *Czechosl. Math. J.* 14 (1964), 386–393. (oroszul).

*Kiss B., Krebsz A.*, On the  $H^{1/2}$  norm based Schur complement preconditioners. In: *Advances in Computational Structures Technology* (B.H.V. Topping, ed.) Civil-Comp Press, Edinburgh 1996, 343–349.

*Kiss B., Krebsz A., Molnárka Gy.*, On the separability of the  $H^{1/2}$  norm in finite element spaces. Preprint 1996, SZIF Győr.

*V.G. Kornyejev*, Magasrendű végeselem módszerek. Állami Leningradi Egyetem 1977 (oroszul).

*V.G. Kornyejev*, Az egzakt differencia-approximációkról. *Zs. Vücs. Matem. i Matem. Fiz.* 22, 3 (1982), 646–654.

*H.O. Kreiss*, Über die Stabilitätsdefinition für Differenzgleichungen, die partielle Differentialgleichungen approximieren. *BIT* 2 (1962), 153–181.

*T. Kröger, T. Preusser*, Stability of the 8-tetrahedra-shortest-interior-edge partitioning method. *Numer. Math.* 109 (2008), 435–457.

*J.R. Kuttler*, Direct methods for computing eigenvalues of the finite difference Laplacian. *SIAM J. Numer. Anal.* 11,4 (1974), 732–740.

*O.A. Ladüzsenszkaja, V.A. Szolonnikov, N.N. Uralceva*, Lineáris és kvázi-lineáris parabolikus egyenletek. Nauka, Moszkva 1967 (oroszul, angol fordítás: AMS, Providence 1968).

*O.A. Ladüzsenszkaja*, A vizskózis összenyomhatatlan folyadék matematikai kérdései. Nauka, Moszkva 1970 (2. kiadás, oroszul).

*H. Lamb*, *Hydrodynamics*. 3rd. ed., Cambridge, University Press 1906.

*Lánczos C.*, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bur. Standards* 45, (1950), 255–282.

*E. Landau*, Vorlesungen über Zahlentheorie. Hirzel, Leipzig 1927 (Chelsea Publ. Co. 1946).

*K.H. Law*, A parallel finite element solution method. *Computers and Structures* 23, 6 (1989), 855–858.

*P.D. Lax, L. Nirenberg*, On stability for difference schemes: a sharp form of Gårding's inequality. *Commun. Pure Appl. Math.* 19,4 (1966), 437–492.

*P.D. Lax, B. Wendroff*, Systems of conservation laws. *Comm. Pure Appl. Math.* 13 (1960), 217–237.

*P.D. Lax, B. Wendroff*, Difference schemes for hyperbolic equations with high order of accuracy. *Comm. Pure Appl. Math.* 17 (1964), 384–398.

*W.J. Layton, P.J. Rabier*, Peaceman–Rachford procedure and domain decomposition for finite element problems. *Numer. Lin. Algebra Appls.*, 2,4 (1995), 363–393.

*V.I. Lebedev*, Difference analogons of orthogonal decompositions, of the basic differential operators and of some boundary value problems of Mathematical Physics. *Soviet J. Numer. Math. and Math. Physics*, part I: 4, 3 (1964), 449–465; part II: 4, 4 (1964), 649–659 (oroszul).

*M. Lees*, A linear three-level difference scheme for quasilinear parabolic equations. *Math. Comp.* 20 (1966), 516–522.

*H.J.W. Lenferink, M.N. Spijker*, On a generalization of the resolvent condition in the Kreiss matrix theorem. *Math. Comp.* 57,195 (1991), 211–220.

*Leonard B.P.*, A stable and accurate convective modelling procedure based on quadratic upstream interpolation. *Comp. Meth. Appl. Mech. & Eng.* 19 (1979), 59–98.

*J. Leray*, Sur le mouvement d'un liquid visqueux emplissant l'espace. *Acta Math.* 63 (1934), 193–248.

*R.J. LeVeque*, Finite Volume Methods for Hyperbolic Problems. Cambridge Univ. Press, Cambridge UK 2002.

*J.L. Lions, E. Magenes*, Non-homogeneous boundary value problems and applications I. Springer 1972.

*D.V. Liseikin*, Grid Generation Methods, Springer Verlag, New York 1999.

*D.V. Liseikin*, A computational differential geometry approach to grid generation. Springer Verlag, Berlin 2004.

*A. Liu, B. Joe*, On the shape of tetrahedra from bisection. *Math. Comp.* 63 (1994), 141–154.

*Z. Lozina*, A comparison of harmonic acceleration method with other commonly used methods for calculation of dynamic transient response. *Computers & Structures* 29,2 (1988), 227–240.

*D.S. Malkus*, Eigenproblems associated with the discrete LBB-condition for incompressible finite elements. *Int. J. Eng. Sci.* 19 (1981), 1299–1310.

*G.I. Marcsuk, V.V. Sajdurov*, Raising the Accuracy of Solutions of Difference Schemes. Nauka, Moszkva 1979.

*A. Meister*, Zur zeitgenauen numerischen Simulation reibungsbehafteter, kompressibler, turbulenter Strömungsfelder mit einer impliziten Finite-Volumen-Methode vom Box-Typ. Dissertation Darmstadt 1996.

*A. Meyer*, Modern Algorithms for Large Sparse Eigenvalue Problems. *Mathematical Research* v. 35., Akademie-Verlag, Berlin 1987.

*J.H. Michael*, A general theory for linear elliptic partial differential equations. *J. Diff. Eqs.* 23 (1977), 1–29.

*W.F. Mitchell*, A comparison of adaptive refinement techniques for elliptic problems. *ACM TOMS* 15,4 (1989), 326–347.

*A.R. Mitchell, G. Fairweather*, Improved forms of the alternating direction methods of Douglas, Peaceman and Rachford for solving parabolic and elliptic equations. *Numer. Math.* 6 (1964) 285–292.

*A.R. Mitchell, D.F. Griffiths*, The Finite Difference Method in Partial Differential Equations. Wiley-Interscience, New York 1980.

*M. Mohr, R. Wienands*, Cell-centred multigrid revisited. *Comput. Vis. Sci.* 7, 3-4 (2004), 129–140.

*K. Morgan, J. Periaux, F. Thomasset (eds.)*, Analysis of Laminar Flow over a Backward Facing Step. A GAMM workshop. *Notes on Numerical Fluid Mechanics* v. 9, Vieweg, Braunschweig 1984.

*J. Nečas*, Equations aux dérivées partielles. Presse de l' Université de Montréal, 1965.

*J. Nečas*, Les méthodes directes en théorie des équationes elliptiques. Prága, Academia 1967.

*J. v. Neumann, R.D. Richtmyer*, A method for the numerical calculation of hydrodynamic shocks. *J. Appl. Phys.* 21,3 (1950), 232–243.

*N.M. Newmark*, A method of computation for structural dynamics, ASCE J. Engineering Mechanics Division, 85 (1959), 67–94.

*Nyers J., G. Stoyan*, A dynamical model adequate for controlling the evaporator of a heat pump. *Int. J. Refrig.* 17,2 (1994), 101–108.

*L.A. Oganeszjan, V.J. Rivkind, L.A. Ruhovec*, Elliptikus egyenletek megoldása variációs differencia módszerekkel. I, II. Differencialegyenletek és alkalmazásai, 5,8, Vilniusz 1973, 1974 (oroszul).

*L.A. Oganeszjan, L.A. Ruhovec*, Elliptikus egyenletek megoldása variációs differencia módszerekkel. Erevan, Örmény Tud. Akadémia 1979 (oroszul).

*J.M. Ortega*, Introduction to Parallel and Vector Solution of Linear Systems. Plenum Press, New York 1988.

*S.V. Patankar*, Numerical Heat Transfer. Academic Press, New York 1980.

*S.V. Patankar, D.B. Spalding*, A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows. Int. J. Heat and Mass Transfer 15 (1972), 1787–1806.

*D.W. Peaceman, H.H. Rachford*, The numerical solution of parabolic and elliptic differential equations. J. Soc. Ind. Appl. Math. 3,1 (1955), 28–41.

*A.G. Peano*, Hierarchies of conforming finite elements for plane elasticity and plate bending problems. Comp. Math. Appl. 2 (1975), 211–224.

*P.P. Pébay, T.J. Baker*, Analysis of triangle quality measures. Math. Comp. 72 (2003), 1817–1839.

*D. Pelletier, A. Fortin, R. Camarero*, Are FEM solutions of incompressible flows really incompressible? (Or how simple flows can cause headaches!) Int. J. Numer. Methods in Fluids 9 (1989), 99–112.

*M. Peric, R. Kessler, G. Scheuerer*, Comparison of finite-volume numerical methods with staggered and non-staggered grids. Computers and Fluids 16 (1988), 389–403.

*J.B. Perot*, An analysis of the fractional step method. J. Comp. Phys. 108 (1993), 51–58.

*V.G. Prikazcsikov, A.N. Himics*, Iterációs módszer lemezek és héjak stabilitási és rezgési feladatainak megoldására. Prikl. Mechanika 20, 1 (1984), 88–94 (oroszul).

*M.H. Protter, H.F. Weinberger*, Maximum Principles in Differential Equations. Prentice-Hall, Englewood Cliffs 1967.

*J.S. Przemieniecki*, Matrix structural analysis of substructures. AIAA J. 1 (1963), 138–147.

*C. Rajakumar*, Lanczos algorithm for the quadratic eigenvalue problem in engineering applications. Comp. Meth. Appl. Mech. & Eng. 105,1 (1993), 1–22.

*R. Rannacher*, Finite element solution of diffusion problems with irregular data. *Numer. Math.* 43 (1984), 309–327.

*A. Reusken*, A new lemma in multigrid convergence theory. Report RANA 91-07, Eindhoven 1991.

*R.D. Richtmyer, K.W. Morton*, Difference Methods for Initial-Value Problems. Wiley–Interscience, New York 1967.

*M.-C. Rivara*, Design and data structure of fully adaptive, multigrid, finite-element software. *ACM TOMS* 10,3 (1984), 242–264.

*P.J. Roache*, Computational Fluid Dynamics. Hermosa Publ., Albuquerque 1976.

*P.L. Roe, D. Sidilkover*, Optimum positive linear scheme for advection in two and three dimensions. *SIAM J. Numer. Anal.* 29 (1992), 1542–1568.

*J.B. Rosser*, Effect of discontinuous boundary conditions on finite-difference solutions. *ZAMP* 27 (1976), 249–272.

*Rózsa P.*, Lineáris algebra és alkalmazásai. 3. kiadás, Tankönyvkiadó, Budapest 1991.

*B.L. Rozsgyeesztvenszkij, N.N. Janenko*, Kvázilineáris rendszerek. Nauka, Moszkva 1978 (oroszul).

*V.V. Sajdurov*, Végeselem alapú többrácsos módszerek. Nauka, Moszkva, 1989 (oroszul). Multigrid Methods for Finite Elements. Kluwer Acad. Publ., Dordrecht 1995.

*V.V. Sajdurov*, On the convergence of the cascadic multigrid method. Proceedings of the Miskolc Conference on Numerical Methods, 1994. Mathematics and Computers with Applications, 1996.

*A.A. Samarskij*, Theorie der Differenzenverfahren. Teubner, Leipzig 1984. Fordítás oroszról: Nauka, Moszkva 1973.

*J.M. Sanz-Serna, C. Palencia*, A general equivalence theorem in the theory of discretization methods. *Math. Comp.* 45,171 (1985), 143–152.

*M. Schäfer, S. Turek*, Benchmark computations of laminar flow around cylinder. In: E.H. Hirschel (ed.), Flow Simulation with High-Performance Computers II. Notes in Numerical Fluid Mechanics, v. 52, Vieweg, Oldenburg 1996, 547–566.

*F. Schieweck*, Numerische Integration bei der Finite-Element-Diskretisierung singularär gestörter elliptischer Randwertaufgaben. *Wiss. Zt. TU Magdeburg* 31 (1987) 5, 95–101.

*F. Schieweck*, On the order of two nonconforming finite element approximations of upwind type for the Navier–Stokes equations. In : Numerical Methods for the Navier–Stokes Equations (F.K. Hebeker et al., eds.), Notes

on Numerical Fluid Mechanics v. 47, Vieweg, Braunschweig 1994, pp. 249–258.

*F. Schieweck, L. Tobiska*, A parallelization strategy for a Navier–Stokes multigrid solver based on macroelements. In : Computational Fluid Dynamics on Parallel Systems (S. Wagner, ed.), Notes on Numerical Fluid Mechanics v. 50, Vieweg, Braunschweig 1995, pp. 176–187.

*A. Schmidt, K.G. Siebert*, Design of Adaptive Finite Element Software. The Finite Element Toolbox Alberta. Springer, Berlin 2005.

*A. Schüller*, Mehrgitterverfahren für Schalenprobleme. GMD-Bericht 171. Oldenbourg Verlag, München 1988.

*I. Senjanovič*, Harmonic acceleration method for dynamic structural analysis. Computers & Structures, 18, 1 (1984), 71–80.

*G. Shortley, R. Weller*, The numerical solution of Laplace’s equation. J. Appl. Phys. 9 (1938), 334–345.

*Simon L., E.A. Baderko*, Másodrendű lineáris parciális differenciálegyenletek. Tankönyvkiadó, Budapest 1983.

*Sleijpen G.L.G.*, Optimal a priori error bounds for the Rayleigh–Ritz method. Math. Comp. 72 (2003), 677–684.

*G.A. Sod*, A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. J. Comp. Phys. 27 (1978), 1–31.

*H. Sohr*, The Navier-Stokes Equations. An Elementary Functional Analytic Approach. Birkhäuser, Basel 2001.

*S. Spekreijse*, Multigrid solution of monotone second order discretizations of hyperbolic conservation laws. Math. Comp. 49 (179), (1987), 135–155.

*R. Stenberg*, Analysis of mixed finite element methods for the Stokes problem: a unified approach. Math. Comp. 165 (1984), 9–23.

*R. Stevenson*, Discrete Sobolev spaces and regularity of elliptic difference schemes. M<sup>2</sup>AN 25,5 (1991), 607–640.

*G. Stoyan*, A parabolikus típusú, többdimenziós egyenletek peremérték feladatai megoldását szolgáló néhány gazdaságos sémáról. Zs. Vücs. Matem. i Matem. Fiz. 10,3 (1970), 644–653 (oroszul).

*G. Stoyan*, Az additív differenciasémák peremérték szerinti stabilitásáról. Zs. Vücs. Matem. i Matem. Fiz. 11,4 (1971), 934–947 (oroszul).

*G. Stoyan*, Néhány ekonomikus séma aszimptotikus stabilitásáról. Zs. Vücs. Matem. i Matem. Fiz. 20,2 (1980), 350–358 (oroszul).

*G. Stoyan*, On maximum principles for matrices, and on conservation of monotonicity. With applications to discretization methods. ZAMM 62 (1982), 375–381.

*G. Stoyan*, On maximum principles for monotone matrices. *Lin. Algebra and Its Appls.* 78 (1986), 147–161.

*G. Stoyan*, On the monotone approximation of a two-dimensional partial differential equation with nonnegative characteristic form. *Proceedings NUMDIFF-5*, Teubner, Leipzig 1991, 259–266.

*G. Strang, G.J. Fix*, *An Analysis of the Finite Element Method*. Prentice-Hall, Englewood Cliffs 1973.

*E. Süli*, Convergence and nonlinear stability of the Lagrange-Galerkin method for the Navier-Stokes equations. *Numer. Math.* 53 (1988), 459–483.

*A.A. Szamarszkij*, Bevezetés a differenciasémák elméletébe. Nauka, Moszkva 1971 (oroszul).

*A.A. Szamarszkij, V.B. Andrejev*, *Differencia módszerek elliptikus egyenletek megoldására*. Nauka, Moszkva 1976 (oroszul).

*A.A. Szamarszkij, A.V. Gulin*, *Differenciasémák stabilitása*. Nauka, Moszkva 1973 (oroszul).

*A.A. Szamarszkij, R.D. Lazarov, V.L. Makarov*, *Differenciasémák vizsgálata differenciálegyenletek általánosított megoldásain*. Vűszsaja skola, Moszkva 1987 (oroszul).

*A.A. Szamarszkij, E.Sz. Nyikolajev*, *Differencia egyenletek megoldási módszerei*. Nauka, Moszkva 1978 (oroszul).

*B.A. Szabó, I. Babuška*, *Finite Element Analysis*. J. Wiley, New York 1991.

*Sz.I. Szerdjukova*, A hővezetési egyenlet megoldását szolgáló, hatpontos magasabbrendű differenciaséma egyenletes stabilitásáról. *Zs. Vűcs. Matem. i Matem. Fiz.* 7,1 (1967), 214–218 (oroszul).

*Temam R.*, *Theory and Numerical Analysis for Navier–Stokes Equations*. North–Holland, Amsterdam 1977.

*A. Thom*, The flow past circular cylinders at low speeds. *Proc. Royal. Soc.*, A 141 (1933), 651–666.

*L.H. Thomas*, *Elliptic problems in linear difference equations over a network*. Watkins Sci. Lab., Columbia, New York 1949.

*V. Thomeé*, *Galerkin Finite Element Methods for Parabolic Problems*. Springer Verlag, New York 1997.

*D.J. Tritton*, *Physical Fluid Dynamics*. 2nd. ed. Clarendon Press, Oxford 1988.

*S. Turek*, *Efficient Solvers for Incompressible Flow Problems. An Algorithmic Approach*. Springer, Berlin 1999.

*R. Verfürth*, A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques. Wiley-Teubner, Chichester & Stuttgart 1996.

*H. Versteeg, W. Malalasekera*, An Introduction to Computational Fluid Dynamics: The Finite Volume Method, 2nd Edition. Prentice Hall, Englewood Cliffs 2007.

*R. Vichnevetsky, J.B. Bowles*, Fourier Analysis of Numerical Approximations of Hyperbolic Equations. SIAM, Philadelphia 1982.

*R. Vilsmeier, D. Hänel*, Adaptive methods on unstructured grids for Euler and Navier-Stokes equations. Computers & Fluids 22, 4/5 (1993), 485–499.

*V.R. Voller, C.R. Swaminathan, B.G. Thomas*, Fixed grid techniques for phase change problems : A review. Int. J. Num. Meth. Eng. 30 (1990), 875–898.

*H.A. van der Vorst*, A fast and smoothly converging variant of BICG for the solution of nonsymmetric linear systems. SIAM J. Sci. Stat. Comp. 13 (1990), 631–644.

*W. Weinelt*, Untersuchungen zur Konvergenzgeschwindigkeit bei Differenzenverfahren. Wiss. Z., H. Karl-Max-Stadt, 20 (1978), 763–769.

*G. Wittum*, Multi-grid methods for Stokes and Navier-Stokes equations. Numer. Math 54 (1989), 543–563.

*H. Yserentant*, Die maximale Konsistenzordnung von Differenzenapproximationen nichtnegativer Art. Numer. Math. 42 (1983), 119–123.

*H. Yserentant*, Über die Maximumnormkonvergenz der Methode der finiten Elemente bei geringsten Regularitätsvoraussetzungen. ZAMM 65 (1985), 91–100.

*E. Zeidler*, Nonlinear Functional Analysis and Its Applications I-VI, Springer-Verlag, New York 1990.

*Chr. Zenger*, Sparse grids. In: Parallel Algorithms for Partial Differential Equations. Proc. 6. GAMM Seminar, Kiel 1990 (W. Hackbusch, ed.). Notes on Numer. Fluid Mechanics, v. 31, Vieweg, Braunschweig 1991.

*S. Zhang*, Successive subdivisions of tetrahedra and multigrid methods on tetrahedral meshes. Houston J. Math. 21(3) (1995), 541–556.

*J.Z. Zhu, O.C. Zienkiewicz*, Adaptive techniques in the finite element method. Comm. Appl. Numer. Methods 4 (1988), 197–204.

*O.C. Zienkiewicz, J.Z. Zhu*, A simple error estimator and adaptive procedure for practical engineering analysis. Int. J. Num. Meth. Eng. 24 (1987), 337–357.



*W. Zulehner*, Analysis of iterative methods for saddle point problems. A unified approach. *Math. Comp.* 71, No. 238 (2001), 479–505.



# Fejezet 20

## Tárgymutató III

### 20.1 Címszavak jegyzéke

#### A

adaptív végeselem módszer15.7.10.

áramfüggvény17.2.

áramlási feladatok15.6.5., 16.4.11., 17. fejezet, 18.10.2.

aszimptotikus stabilitás16.4.9.

#### B

beágyazás, kompakt15.9.1., 19.

beágyazási tételek19.

diszkrét15.5.3., 15.11.: 33. feladat

Benard–cella, diszkrét17.5.3., 17.6.: 15. feladat

bilineáris elem15.7.4., 15.7.8., 15.9.2.

Black–Scholes egyenletld. egyenlet alatt

boxmódszer (véges térfogat módszer)15.6., 15.7.8., 17.4.1., 18.10.3.

Bramble–Hilbert lemma, bilineáris15.7.5.

buborékfüggvény17.5.4.

#### C

Cauchy–Riemann egyenletrendszer15.1., 17.2.

Clément-féle leképezés17.5.4.

Courant–elem15.7.4., 15.7.9., 15.8.1., 15.9.2., 17.5.3.

Courant–Friedrichs–Lewy feltétel18.3.3.

Courant-szám 16.4.3., 18.3.3.  
 Crank-Nicolson séma alatti séma alatt  
 Crouzeix-Raviart elem 17.5.4.  
 csomópontok 15.7.4.

nemkonform 15.7.10.  
 $C^1$ -interpoláció 15.7.4., 17.5.2.

## D

D'Alembert-féle képlet 18.2.3.  
 differenciacsillag 15.3.1., 16.4.1  
 differenciaséma 15.3.1.

5-pontos 15.3.1., 15.3.2.

9-pontos 15.3.2.

súlyozott 16.4.1., 16.4.11., 18.3.1.  
 dimenziócsökkentő eljárások 16.3.  
 „dimenziótlanítás” 16.3., 17.1.  
 Dirac-féle  $\delta$ -függvény 15.5.5., 16.1.1.  
 direkt átvitel 15.4.3., 15.4.7., 15.11.: 24. feladat  
 Dirichlet-perem 16.1.3.

diszkrét 16.4.3.  
 divergenciamentesség 17.5.3.

diszkrét 17.5.3.  
 Dryja-féle prekondicionálási mátrix 15.8.5.

## E

egyenesek módszere 16.4.1., 16.4.6.  
 egyenlet, Black-Scholes 16.1.2.

egyenletesen parabolikus 16.2.

elliptikus 15.2.

hiperbolikus 18.1.1.

- hővezetési 15.1., 16.1.3., 16.5., 16.6., 16.8., 18.4.3.
- hullám-18.1.1., 18.2.3, 18.2.2., 18.3.1., 18.3.2., 18.3.3., 18.4.3., 18.7.
- többdimenziós 18.1.2., 18.3.1., 18.4.3., 18.8., 18.9.
- konvekció-diffúzió 15.4.3., 15.7.6., 16.1.1., 16.1.3., 17.4.1., 18.12.: 11. feladat
- Laplace-15.1., 15.6.3., 15.7.10.
- Maxwell-18.1.2., 18.7., 18.8.
- Navier-Stokes 17.
- Poisson-15.1., 15.3.1., 15.3.2., 15.7.8., 16.7.2., 17.4.3.
- Poisson-, háromdimenziós 15.3.2.
- Schrödinger-18.4.3.
- szivárgási 15.1., 16.5.
- távíró 18.7.
- teljes potenciál-15.1.
- transzport-16.4.11., 17.4.1., 18.2.4., 18.3.3., 18.4.3., 18.5., 18.10.1.  
egyenletrendszer, akusztikai 18.1.1., 18.4.3., 18.12.: 10.d) feladat
- elsőrendű hiperbolikus 18.10.2.
- gázdinamikai 18.1.2., 18.10.2., 18.10.3.
- Navier–Stokes 17.1.
- elősimítás 15.4.3.
- energiamegmaradási tétel 18.2.2.
- Euler–Galjorkin módszer, implicit 16.8.

**F**

FAC-algoritmus\*\*\* 15.7.10  
 feladat, korrekt kitűzésű15.2., 16.2.  
 fizikai jelenségek szerinti felbontás16.7.3.  
 formafüggvények15.7.4.  
 Fourier–transzformáció, gyors15.3.3., 15.8.5.  
  
 diszkrét16.4.9., 18.4.2.  
 folytatásos módszer15.4.9.  
 Frjazinov–approximáció17.4.1., 17.4.2.  
 diszkrét Friedrichs-féle egyenlőtlenség15.11.: 33. feladat  
 Frobenius–Perron tétel15.9.2.  
 függőségi tartomány18.3.3.  
 függvény, majoráns15.3.1., 16.4.7.

**G**

Galjorkin–projekció15.8.1., 16.4.6., 16.8., 18.3.2.  
 Galjorkin–reláció15.4.10., 15.8.1.  
 Gårding-féle egyenlőtlenség15.11.: 38. feladat  
 Green-féle függvény, diszkrét15.3.5., 15.5.1., 15.5.2., 15.7.8.,

**Gy**

gyökfeltétel, egyenletes erősített18.4.2.

**H**

háromszögek stabil finomítása15.7.11  
 Hermite-típusú elem15.7.4.  
 hő-polinomok16.3.  
 hullámegyenletld. egyenlet alatt

**I**

inf-sup-feltétel17.5.2.  
  
 diszkrét17.5.4.  
 integrálás, numerikus15.7.6., 16.4.8.  
 interpolációs függvények15.7.4.  
 inverz egyenlőtlenség15.11.: 46. feladat, 17.5.4.

iteráció, Gauss–Seidel15.4.3.

áramvonal-követő17.5.4.

simító15.4.2.

zebra15.4.3.

izoparametrikus módszer15.7.9.

izovonalak15.7.9., 15.7.10.

## K

karakterisztika18.2.4., 18.10.1., 18.10.2.

karakterisztikák módszere18.10.1., 18.10.2.

képlethiba15.3.1., 16.4.1.

koercív15.9.1.

konvekció–diffúzió egyenletld. egyenlet alatt

konvergencia vizsgálat 15.3.2, 15.3.5., 15.4.5., 15.4.8., 15.7.5., 15.8.2,  
16.4.4.

Kreiss-féle mátrixtétel18.4.2.

Krülov–sorozat15.9.4.

## L

Ladüzenszkaja–Babuška–Brezzi-feltétel (LBB-feltétel)17.5.3.

Lagrange-típusú elem15.7.4.

Lánczos–módszer15.9.4.

Laplace–operátor15.2.

polárkoordinátákban15.2.

Lax–Milgram tétel15.7.3.

## M

mátrix, átmeneti18.4.2.

globális15.7.7.

lokális15.7.7.

merevségi15.7.7.

monoton15.3.2.

tömeg-15.7.8., 15.9.3., 16.4.6., 16.4.8.

mátrixsereg, egyenletesen diagonalizálható18.4.2.

spektrálisan ekvivalens mátrixok, ill. operátorok15.8.5., 15.9.3.

maximumelv15.2.

diszkrét15.3.2., 16.4.7.

Hopf-féle15.2.

szigorú15.3.2.

megmaradási tétel15.6.2., 18.2.2.

diszkrét16.4.5.

megoldás, klasszikus15.2., 16.2.

gyenge15.2., 16.2., 16.8., 17.5.2., 18.2.

megszámozás, globális15.7.7.

lokális15.7.7.

mérlegegyenlet15.6.2., 15.7.8.

minieleme17.5.4.

monotonitás megmaradása16.4.7., 18.10.1.

műveletigény15.3.3., 15.4.4.

## N

Navier-Stokes egyenletld. egyenlet alatt

Neumann-féle stabilitási feltétel18.4.2.

Neumann-féle stabilitási vizsgálat18.4.1, 18.4.2.

Newmark módszer18.7., (150)

Nitsche-fogás15.7.3., 15.8.1., 16.8., 17.5.3.

numerikus diffúzió17.4.1.

numerikus diszperzió18.10.1.

## O

operátor, egyenletesen elliptikus15.2.



leszűkítő15.4.1.

restrikciós15.4.1.

prolongációs15.4.1.

oszillációmentesség16.4.7., 17.4.2., 18.10.1.

## Ö

összehasonlítási tétel15.2., 16.2.

## P

párhuzamos számításra alkalmas módszerek15.8.3., 16.7.1., 16.7.3.

Patankar–Spalding-féle nyomáskorrekció17.4.3.

Peaceman–Rachford módszer16.7.1., 16.7.2.

perem, Lipschitz-folytonos15.2.

peremérték15.2.

peremérték feladat15.2.

peremfeltétel

ferdeirányú17.2.1.

harmadfajú15.3.4., 16.4.5., 16.5.4.

lényeges15.7.2.

másodfajú15.3.4., 16.4.5.

természetes15.7.2.

Thom-féle17.2.3.

peremfeltételekről15.2., 15.7.2., 17.2.3.

$P(k)$ -feltétel15.7.6.

Poisson-egyenletld. egyenlet alatt

polárkoordináták15.2., 15.6.3. 15.11.: 10. feladat, 16.5.3.

pozitivitástartás (nemnegativitás megőrzése)15.3.2., 15.6.4., 16.4.7.

Patankar-fogás16.4.7.

## Q

QUICKEST-algoritmus18.12.: 11.d) feladat  
 $Q(k)$ -feltétel15.7.6.

## R

rács15.3.1.  
 rács finomítása felezéssel15.4.4.  
 rácsrendszer, eltolt17.4.2.  
 rácsszerkesztés15.7.10.  
 Rayleigh–Ritz módszer15.9.4.  
 rend, végelemé15.7.4.  
 rendszer, akusztikai18.1.1., 18.4.3., 18.12.: 10.d) feladat

gázdinamikai18.1.2., 18.10.2., 18.10.3.  
 Reynolds-szám17.1.

cella-Reynolds-szám17.4.0.  
 Riemann-feladat18.10.2.  
 Riemann-féle invariánsok18.10.2.  
 Ritz-projekció16.8.  
 Rothe-módszer16.7.4.  
 Runge–Kutta módszerek18.4.3., 18.12.: 9. feladat

## S

sajátérték feladat15.9.1.  
 adjungált15.9.1.  
 sakktábla elrendeződés15.3.3., 15.4.3.  
 instabilitás17.4.3.  
 Schrödinger-egyenletld. egyenlet alatt  
 Schwarz-egyenlőtlenség, általánosított15.5.3.  
 Shortley–Weller approximáció15.3.5.  
 séma, explicit16.4.1.  
 faktorizált16.7.3.  
 gyengén monoton17.4.1.  
 háromréteges16.4.11., 16.6., 18.5.

implicit16.4.1.

kétréteges16.4.3.

magasabbrendű16.4.2.

Crank–Nicolson16.4.1., 16.4.3., 16.4.7., 16.4.9., 16.4.11., 17.4.3., 18.7., 18.10.1.

“splitting” módszer16.7.3

stabilitási vizsgálat15.3.2., 16.4.3., 16.4.9., 16.4.10., 17.5.2., 18.4.2., 18.5.

Stokes-féle feladat17.5.1.

## Sz

szabadsági fokok15.7.4.

Szamarszkij-féle stabilitási vizsgálat16.4.3., 18.5.

szárítási feladat16.10.: 4. feladat

szemidiszkrétizáció16.4.1., 16.4.6., 16.8., 18.3.2.

Sztyeklov-féle átlagolási operátorok15.5.4.

szuperpozíciós módszer18.8.

## T

tárigény15.3.3., 15.4.4.

tartomány közelítése15.7.9.

távíró egyenletld. egyenlet alatt

terhelési vektor15.7.7.

tetraéderek stabil finomítása15.7.11

többrácsos módszer15.4., 16.7.4.

approximációs tulajdonsága15.4.5.

sima15.4.3.

simítási tulajdonsága15.4.5.

teljes15.4.3.

tömegmátrixld. mátrix alatt

transzport egyenletld. egyenlet alatt

trianguláció15.6.2., 15.7.10.

egyenletesen reguláris15.7.5.

finomítása15.4.4., 15.7.10.

konform15.6.2., 15.7.4.

nemkonform15.6.2., 15.7.10.

## U

upwind-séma (avagy upstream-séma)16.4.10., 17.4.0., 17.4.1., 18.10.2.,  
18.12.: 11.c) feladat

utósimítás15.4.3.

## V

váltakozó irányok módszerei16.7.1., 16.7.4.

véges térfogat15.6.2.

véges térfogat módszerld. boxmódszer alatt

végeselem módszer  $p$ -verziója15.7.4.

végeselem séma16.4.6., 18.3.2.

végeselem térpár, stabil17.5.4.

vegyes deriváltak approximációja15.11.: 32. feladat

vegyes feladat16.1.3., 18.2.2.

vegyes végeselem módszer17.5.2.

## 20.2 Tételek, lemmák jegyzéke

15.1. Tétel (J.H. Michael; egyenletesen elliptikus elsőfajú peremérték feladat korrekt kitűzésű)

15.2. 15.2. Lemma (az ötpontos differenciaséma tulajdonságai, 1)15.3.2.

15.3. Tétel (az ötpontos differenciaséma stabilitása)15.3.2.

15.4. Tétel (az ötpontos differenciaséma konvergenciája)15.3.2.

15.5. Tétel (maximumelvvvel rendelkező mátrixosztály)15.3.2.

15.6. Tétel (szigorú maximumelvvvel rendelkező mátrixosztály)15.3.2.

15.7. Lemma (monoton mátrixú egyenletrendszer megoldásának becslése) 15.3.2.

15.8. Következmények (az ötpontos differenciaséma tulajdonságai, 2) 15.3.2.

- 15.9. Tétel (szigorú diszkrét Hopf-féle maximumelv)15.3.2.
- 15.10. Következmény (pozitív elemű inverz mátrix)15.3.2.
- 15.11. Tétel (harmadfajú peremérték feladat differencia approximációjának stabilitása és konvergenciája)15.3.4.
- 15.12. Tétel (Shortley–Weller séma stabilitása és pontossága)15.3.5.
- 15.13. Tétel (további, maximumelvvvel rendelkező mátrixosztály) 15.11.
- 15.14. Lemma (az egyszerű iteráció simító tulajdonsága)15.4.7.
- 15.15. Lemma (a szimmetrikus Gauss–Seidel iteráció simítási tulajdonsága)15.4.7.
- 15.16. Lemma (Reusken; Ecker–Zulehner; simítási tulajdonság a nem-szimmetrikus esetben)15.4.7.
- 15.17. Tétel (a W-ciklusos sima többrácsos módszer konvergenciája)15.4.8.
- 15.18. Tétel (Hackbusch; teljes többrácsos módszer konvergenciája)15.4.8.
- 15.19. Lemma (A Laplace-operátor diszkrét Green-féle függvényének becslése)15.5.1.
- 15.20. Tétel (önadjungált elliptikus egyenletet approximáló differenciaséma stabilitása)15.5.2.
- 15.21. Tétel (önadjungált elliptikus egyenlet diszkrét Green-függvényének becslése)15.5.2.
- 15.22. Lemma („tilos” beágyazási tétel)15.5.3.
- 15.23. Lemma (változó együtthatójú feladat megoldásának maximum-normabeli stabilitása)15.5.3.
- 15.24. Lemma (diszkrét beágyazási tétel)15.5.3.
- 15.25. Tétel (differenciaséma konvergenciája  $f \in L_2$  esetén) 15.5.4.
- 15.26. Tétel (Szamarszkij, Lazarov, Makarov; differenciaséma konvergenciája Dirac-féle  $\delta$ -függvény jelenlétében)15.5.5.
- 15.27. Lemma (a másodfajú és harmadfajú peremfeltétel természetes)15.7.2.
- 15.28. Lemma (Babuška; variációs feladat megoldhatósága)15.7.3.
- 15.29. Következmény (lineáris funkcionál által definiált operátor egyértelműsége)15.7.3.
- 15.30. Tétel (Kadlec; variációs megoldás  $H_0^2$ -ban)15.7.3.
- 15.31. Tétel ( $k$ -adfokú Lagrange-féle interpoláció hibája egyenletesen reguláris háromszöges trianguláción)15.7.5.
- 15.32. Következmény (Lagrange-féle konform végeselem módszer konvergenciája egyenletesen reguláris trianguláción)15.7.5.
- 15.33. Tétel (Ciarlet; közelítő bilineáris forma egyenletes ellipticitása)15.7.6.
- 15.34. Lemma (Strang; közelítő variációs feladat hibája)15.7.6.

- 15.35. Tétel (Ciarlet; végeelem konvergencia kvadratúra képletek használatákor)15.7.6.
- 15.36. Tétel (Hackbusch; a boxmódszer és a Courant–elemes végeelem módszer eltérése)15.7.8.
- 15.37. Lemma (végeelem alapú többrácsos módszer approximációs tulajdonsága)15.8.1.
- 15.38. Lemma (két norma ekvivalenciája a Courant–elemek esetén)15.8.1.
- 15.39. Tétel (Hackbusch; végeelem alapú többrácsos módszer konvergenciája)15.8.1.
- 15.40. Lemma (Hackbusch; a többrácsos iterációs mátrix tulajdonságai, 1)15.8.2.
- 15.41. Lemma (Braess, Hackbusch; többrácsos iterációs mátrix tulajdonságai, 2)15.8.2.
- 15.42. Tétel (Braess, Hackbusch; szimmetrikus többrácsos módszer konvergenciája)15.8.2.
- 15.43. Tétel (Braess, Hackbusch;  $V$ -ciklus konvergenciája szimmetrikus esetben)15.8.2.
- 15.44. Lemma (U. Langer és társai; önadjungált többrácsos operátorok spektrális ekvivalenciája)15.8.5.
- 15.45. Tétel (U. Langer és társai; Dirichlet–tartomány dekompozíciós prekondicioner becslése)15.8.5.
- 15.46. Lemma (Dryja; a Dryja-féle prekondicionálási mátrix becslése)15.8.5.
- 15.47. Tétel (Riesz–Schauder elmélet)15.9.1.
- 15.48. Tétel (egyszeres sajátérték végeelem approximációjának konvergenciája)15.9.1.
- 15.49. Lemma (bilineáris végeelem approximáció sajátértékei)15.9.2.
- 15.50. Lemma (Rayleigh–hányados tulajdonságai)15.9.3.
- 15.51. Tétel (Prikazcsikov–Himics; prekondicionált gradiens módszer konvergenciája)15.9.3.
- 16.1. Tétel (parabolikus egyenlet maximumelve)16.2.
- 16.2. Következmény (összehasonlítási tétel)16.2.
- 16.3. Tétel (Ladüzsenszkaja, Uralceva; gyenge megoldás tulajdonságai)16.2.
- 16.4. Lemma (az explicit séma maximumnormabeli stabilitása)16.4.3.
- 16.5. Tétel (Szamarszkij; kétréteges séma kezdetiértékek szerinti stabilitása)16.4.3.
- 16.6. Tétel (kezdetiérték szerinti és jobboldal szerinti stabilitás ekvivalenciája)16.4.3.

- 16.7. Tétel (Szamarszkij; kétréteges séma jobboldal szerinti stabilitása) 16.4.3.
- 16.8. Tétel (súlyozott differenciaséma konvergenciája) 16.4.4.
- 16.9. Tétel (másodfajú és harmadfajú peremfeltételekhez tartozó mátrix tulajdonságai) 16.4.5.
- 16.7'. Tétel (Kétréteges séma jobboldal szerinti  $L_2(\bar{\omega}_h)$ -stabilitása, ha  $\sigma \geq \frac{1}{2}$ ) 16.4.5.
- 16.9'. Következmény (konvergencia  $L_2(\bar{\omega}_h)$ -normában másod- és harmadfajú peremfeltételek esetén) 16.4.5.
- 16.10. Tétel (konvergencia  $L_2(\bar{\omega}_h)$ -normában másodfajú peremfeltételek esetén) 16.4.5.
- 16.11. Tétel (a végeselem séma stabilitása) 16.4.6.
- 16.12. Lemma (két súlyozott séma M-mátrix tulajdonsága) 16.4.7.
- 16.13. Tétel (két súlyozott séma pozitívstartása) 16.4.7.
- 16.14. Tétel (stabilitás a maximum normában) 16.4.7.
- 16.15. Tétel (a Crank–Nicolson eljárás aszimptotikus stabilitása) 16.4.9.
- 16.16. Tétel (stabilitás valósan pozitív definit mátrix esetén) 16.4.10.
- 16.17. Lemma (változó együtthatójú differencia-approximáció tulajdonságai) 16.5.1.
- 16.18. Következmény (változó együtthatójú, súlyozott differenciaséma stabilitása) 16.5.1.
- 16.19. Tétel (Szamarszkij; időtől függő kétréteges differenciaséma stabilitása) 16.5.2.
- 16.20. Lemma (speciális beágyazási tételek) 16.5.3.
- 16.21. Tétel (a súlyozott differenciaséma konvergenciája hengersizimetria esetén) 16.5.3.
- 16.22. Tétel (hengersizimetrikus eset: konvergencia a maximum normában) 16.5.3.
- 16.23. Tétel (Szamarszkij; Newton–módszer konvergenciája nemlineáris parabolikus egyenlet megoldásakor) 16.6.
- 16.24. Tétel (a Peaceman–Rachford eljárás stabilitása) 16.7.1.
- 16.25. Tétel (evolúciós egyenlet gyenge megoldása) 16.8.
- 16.26. Tétel (szemidiszkrétizáció hibabecslése) 16.8.
- 16.27. Tétel (az implicit Euler–Galjorkin módszer hibabecslése) 16.8.
- 17.1. Tétel (áramfüggvény létezése és unicitása) 17.2.1.
- 17.2. Tétel (divergenciamentes vektormező ortogonális felbontása) 17.2.1.
- 17.3. Következmény (tetszőleges vektormező ortogonális felbontása) 17.2.1.
- 17.4. Lemma (a Frjazinov–korrekció tulajdonságai) 17.4.1.
- 17.5. Lemma (a Frjazinov–approximáció gyenge monotonitása) 17.4.1.

- 17.6. Tétel (a konvekció-diffúzió egyenlet Frjazinov–approximációjának stabilitása)17.4.1.
- 17.7. Lemma (A diszkrét grad és div operátorok adjungáltsága) 17.4.2.
- 17.8. Tétel (Navier-Stokes egyenletek Frjazinov–approximációjának stabilitása)17.4.2.
- 17.9. Tétel (variációs feladat nyeregponjtjának jellemzése)17.5.1.
- 17.10. Tétel (vegyes variációs feladat megoldhatósága)17.5.2.
- 17.11. Tétel (Nečas, a divergencia-egyenlet stabil megoldása)17.5.2.
- 17.12. Tétel (vegyes variációs feladat megoldásának stabilitása) 17.5.2.
- 17.13. Lemma (diszkrét vegyes variációs feladat szingularitása) 17.5.3.
- 17.14. Tétel (vegyes végeelem megoldás tulajdonságai)17.5.3.
- 17.15. Tétel (a Céa–lemma általánosítása a vegyes végeelem módszerre) 17.5.3.
- 17.16. Tétel (vegyes végeelem megoldás konvergenciája)17.5.3.
- 17.17. Lemma (Fortin; az inf-sup-feltétel teljesülése)17.5.4.
- 17.18. Tétel (másodrendű Crouzeix–Raviart elem stabilitása)17.5.4.
- 17.19. Lemma (Verfürth–fogás)17.5.4.
- 17.20. Tétel (Arnold, Brezzi és Fortin; a minielem stabilitása) 17.5.4.
- 18.1. Tétel (hullámeqyenlet szemidiszkretizációjának hibája)18.3.2.
- 18.2. Tétel (Neumann-féle stabilitási feltétel érvényessége) 18.4.2.
- 18.3. Tétel (elégéses stabilitási feltétel, 1)18.4.2.
- 18.4. Tétel (elégéses stabilitási feltétel, 2)18.4.2.
- 18.5. Lemma (Szamarszkij, Gulin; speciális norma)18.5.
- 18.6. Tétel (Szamarszkij, Gulin; a súlyozott séma stabilitása kezdetiértékek szerint)18.5.
- 18.7. Tétel (Szamarszkij, Gulin; a súlyozott séma stabilitása a jobboldalra nézve)18.5.
- 18.8. Következmény (súlyozott séma stabilitási becslése)18.5.
- 18.9. Lemma (inhomogén hiperbolikus séma visszavezetése homogén egyenletek sorozatára)18.5.
- 18.10. Tétel (Szamarszkij; a hiperbolikus séma stabilitása a jobboldal szerint)18.5.
- 18.11. Tétel (Szamarszkij; háromréteges hiperbolikus séma konvergenciája)18.6.
- 18.12. Lemma (Godunov; lineáris homogén sémák monotonitása)18.10.2.
- 18.13. Tétel (Godunov; a homogén transzport-egyenlet sémáinak jellemzése)18.10.2.
- 18.12. Tétel (blokk-tridiagonális Gauss–elimináció stabilitása)18.10.3.



## 20.3 Pseudokódos algoritmusok jegyzéke

sima többrácsos módszer15.4.3.  
teljes többrácsos módszer15.4.3.  
többrácsos módszer nemrekurzív algoritmusa15.4.3.  
eljárás nemlineáris rendszer többrácsos megoldására15.4.9.  
merevségi mátrix és terhelési vektor összeállítása15.7.7.  
prekondicionált gradiens módszer (sajátérték feladat)15.9.3.  
blokk-tridiagonális Gauss-elimináció18.10.3.

## 20.4 Táblázatok jegyzéke

ötponos és kilencponos differenciaséma pontossága mintafeladaton  
15.3.2.  
fontosabb módszerek tár- és műveletigénye a kétdimenziós Poisson-egyen-  
let megoldásakor15.3.3.  
különböző végelemek rendje15.7.5.  
Courant-elemes módszer pontossága mintafeladaton15.7.8.  
az I. 1.6.1-beli táblázat folytatása ( $10^6$  ismeretlenű rendszer megoldása)  
15.8.  
beágyazási tételek feltételei Hilbert-terek esetén19.