

EÖTVÖS LORÁND TUDOMÁNYEGYETEM  
TERMÉSZETTUDOMÁNYI KAR

---

Szathmári Csaba

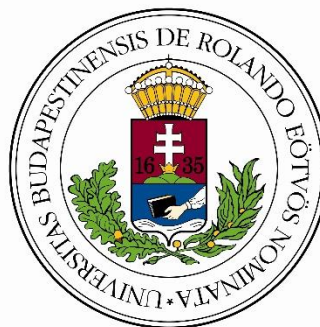
**PARCIÁLIS DIFFERENCIÁLEGYENLETEK  
NUMERIKUS MEGOLDÁSI MÓDSZEREI**

Szakedolgozat

Matematika BSc, alkalmazott matematikus szakirány

Témavezető:

Fábián Gábor



# Köszönetnyilvánítás

Köszönetet szeretnék mondani témavezetőmnek, Fábián Gábornak, hogy segített a szakirodalom kiválasztásában és részletes visszajelzéseket adott a szakdolgozatommal kapcsolatban.

Hálával tartozom továbbá családomnak és barátaimnak a rengeteg biztatásért és támogatásért amit kaptam illetve kapok tőlük.

## Tartalomjegyzék

1 Bevezetés .....	4
1.1 Motiváció .....	4
1.2 Egy alapvető példa.....	4
1.3 Definíciók .....	6
2 Véges differenciák módszere .....	7
2.1 Megoldási ötlet .....	7
2.2 Rácsfelosztás, differencia hányadosok .....	7
2.3 A módszer a Poisson-feladatra .....	8
2.4 Stabilitás és konvergencia.....	13
3 Végeselem módszer .....	15
3.1 Megoldási ötlet .....	15
3.2 A Ritz-Galjorkin módszer.....	15
3.3 A Ritz-Galjorkin módszer használata.....	19
3.4 Stabilitás és konvergencia.....	23
4 Véges térfogat módszer.....	24
4.1 Megoldási ötlet .....	24
4.2 A módszer bemutatása.....	24
4.3 Konvergencia .....	30
5 Összefoglalás .....	31

# 1 Bevezetés

## 1.1 Motiváció

Differenciálegyenletek és parciális differenciálegyenletek a tudomány minden területén előfordulnak, ahol valamilyen természetes folyamat modellezésére van szükség. Ilyen lehet például a hőterjedés, folyadékáramlás, egy populáció számának változása vagy a tőzsdei árfolyamok ingadozása. Az, hogy leírunk egy ilyen folyamatot általában azt jelenti, hogy megmondjuk melyik időpillanatban milyen állapotban van a rendszer amit modellezünk. Ezt azonban gyakran nem tudjuk megtenni, csak az időközben végbemenő változásokat ismerjük. Ezeket deriváltak segítségével fogalmazzuk meg, azaz felírunk egy egyenletet amiben szerepel(nek) a keresett függvény deriváltja(i). Az ilyen egyenleteket nevezzük differenciálegyenleteknek. Ha a függvényünk több változós, mondjuk egy térben is időben zajló eseménynél mint például a hőterjedés, akkor ebben a parciális deriváltak vannak és parciális differenciálegyenlet hívjuk. Azonban ez az egyenlet önmagában kevés a kérdés megválaszolására, szükségünk van további feltételekre, amiket mellékfeltételeknek hívunk. Gondoljunk csak a hőterjedésre, ahhoz hogy megmondjuk mikor és hol éppen milyen meleg van, legalább ismernünk kell a kiindulási (kezdeti) és határoló közeg (perem) hőmérsékletét. A rossz hír, hogy ha még létezik is a feladatnak egyértelmű megoldása, akkor is csak ritkán tudjuk analitikus módon megoldani, ezért szükség van numerikus megoldási módszerekre. A szakdolgozatomban megnézzük három alapvető technika, a véges differenciák-, a végelelem- és a véges térfogat-módszer alkalmazását egy-egy példán keresztül.

## 1.2 Egy alapvető példa

Ebben az alfejezetben egy példán keresztül, a hővezetési egyenleten, megnézzük, hogy milyen képet tudunk társítani egy parciális differenciálegyenlethez. Az egyenletet Joseph Fourier vezette le 1822-ben megjelent “Théorie analytique de la chaleur” című könyvében. Mi nem vezetni fogjuk, csupán a megértését próbáljuk segíteni. Az egyszerűség kedvéért tekintsünk a  $(0, l)$  szakaszt, gondolhatunk úgy rá, mint egy  $l$  hosszú (csak a két végén nem szigetelt) vékony rúdra. Tegyük fel, hogy ismerjük a szakasz minden pontjában a kezdeti hőmérsékletet.

Arra vagyunk kíváncsiak, hogy hogy fog változni a hőmérséklet, ahogy telik az idő. Azaz keresünk egy függvényt ami egy tér és egy idő koordinátához hozzárendeli az ezekhez tartozó hőmérsékletet. Fizikából tudjuk, hogy a magasabb hőmérséklet az alacsonyabb hőmérséklet felé áramlik. Ez azt jelenti, hogy minden pont hőmérséklete a környezetének a hőmérsékletéhez fog igazodni (ha hidegebb a környezet lehűl, ha melegebb akkor felmelegszik). Tapasztalatból tudjuk, hogy ennek a változásnak a mértéke a két érték különbségétől függ (például télen a forró tea a hóban hamarabb hűl mint a szobában). Nézzük meg hogyan tudjuk ezt a tudást felhasználni az egyenletünk felírásához. Legyenek  $x_1, x_2, x_3$  egymástól  $\Delta_x$  távolságra lévő pontok a szakaszon, úgy hogy  $x_1 = x_2 - \Delta_x$  és  $x_3 = x_2 + \Delta_x$ .  $T_1, T_2, T_3$  hőmérsékletekkel,  $T_2'$  pedig a  $x_2$ -beli hőmérséklet  $\Delta_t$  időegység múlva. Ezekkel a jelölésekkel az  $x_2$ -beli hőmérsékletváltozás  $T_2' - T_2$ . Az előbbieket szerint ez az átlagtól való eltéréssel arányos (vagy legalábbis annak valami szigorú monoton függvénye), azaz felírhatjuk, hogy

$$T_2' - T_2 = \alpha \left( \frac{T_1 + T_3}{2} - T_2 \right)$$

ahol  $\alpha$  egy fizikai konstans, ami jellemzi az anyag hővezetési képességét. Az egyenlőséget átrendezve

$$T_2' - T_2 = \frac{\alpha}{2} ((T_1 - T_2) - (T_2 - T_3))$$

A bal oldalon időben eltérő hőmérsékletek különbsége, a jobb oldalon helyben eltérő mennyiségek különbségének a különbsége áll. Azonban nem abszolút különbségekre vagyunk kíváncsiak, hanem relatívakra, ezért a megfelelő eltérésekkel leosztjuk a kifejezéseket:

$$\frac{T_2' - T_2}{\Delta_t} = \frac{\alpha}{2} \frac{\frac{T_1 - T_2}{\Delta_x} - \frac{T_2 - T_3}{\Delta_x}}{\Delta_x}$$

Erről a diszkrét példáról a folytonos esetre áttérve, ezeknek a hely és idő szerinti különbségeknek az analógja a megfelelő parciális deriváltak lesznek. Formálisan így néz ki: keresünk egy  $T(0, l) \times (0, \infty) \rightarrow \mathbb{R}$  függvényt, amire a  $T(x, 0)$  kezdeti értékeket ismerjük és igaz a következő összefüggés:

$$\partial_t T(t, x) = \alpha(x) \partial_x^2 T(t, x)$$

ahol  $\alpha$  a közeg hővezetését jellemzi. Továbbá adottak a szükséges mellékfeltételek. Ennek a több dimenzióra való általánosítását hívjuk hővezetési egyenletnek.

## 1.3 Definíciók

[3]-ban lévő definíciókat fogjuk használni.

Függvényekre  $f, g, h, u \dots$ , mátrixokra  $A, B, \dots$ , változókra  $x_1, x_2, t, \dots$ , vektorváltozóra  $\mathbf{x}$ , parciális deriváltakra a  $\partial_{x_1}, \partial_t, \dots$ , jelöléseket fogjuk használni, ha egyértelmű melyik változó szerint deriválunk, akkor az  $u'$  jelölést használjuk. A tartományt amelyen a megoldást keressük  $\Omega$ -val, a peremét  $\partial\Omega$ -val, a lezártját  $\bar{\Omega}$ -sal jelöljük.

**1.4.1. Definíció:** Legyen  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$  nemnegatív egész számokból álló vektor és  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  függvény. Ekkor

$$\partial^\alpha f(\mathbf{x}) := \partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n} f(\mathbf{x})$$

és  $\alpha$ -át multiindexnek nevezzük. Továbbá az

$$|\alpha| := \alpha_1 + \dots + \alpha_n$$

szám a parciális derivált rendje.  $|\alpha| = 0$ -ra  $\partial^\alpha f := f$ .

**1.4.2. Definíció:** Legyen  $\Omega \subset \mathbb{R}^n$  ( $n \in \mathbb{N}^+$ ) összefüggő nyílt halmaz (tartomány). Legyen  $\alpha$  multiindex, amire  $0 \leq |\alpha| \leq m$ , és jelölje  $N$  az ilyen multiindexek számát. Legyen  $F: \Omega \times G \rightarrow \mathbb{R}$  ( $n+N$ ) változós függvény. Keresünk egy  $u \in C^\infty(\Omega)$  függvényt (hogy konkrétan hányszor differenciálhatónak kell lenni az adott feladattól függ), amire teljesül a következő egyenlet:

$$F\left(\mathbf{x}, u(\mathbf{x}), \partial_{x_1} u(\mathbf{x}), \dots, \partial_{x_n} u(\mathbf{x}), \partial_{x_1} \partial_{x_2} u(\mathbf{x}), \dots, \partial_{x_1}^m u(\mathbf{x})\right) = 0 \quad (\mathbf{x} \in \Omega)$$

(itt  $F$  argumentumában szerepel az összes legfeljebb  $m$ -edrendű parciális deriváltja  $u$ -nak) Ezt az egyenletet nevezzük parciális differenciálegyenletnek, és azt kielégítő függvényeket hívjuk a parciális differenciálegyenlet klasszikus megoldásainak. Az  $m$  szám a parciális differenciálegyenlet rendje. Amennyiben a megoldás nem folytonosan differenciálható, akkor gyenge vagy általánosított megoldásról beszélünk.

# 2 Véges differenciák módszere

Ebben a fejezetben [2]-öt követjük.

## 2.1 Megoldási ötlet

Egzakt deriváltak helyett a deriváltakat differenciahányadosokkal közelítjük és ezekkel oldjuk meg a feladatot. Ehhez felosztjuk a tartományt diszkrét pontokra és ezeken a pontokon fogjuk a keresett függvény értékeit közelíteni. Így egy lineáris egyenletrendszert kapunk, aminek megoldására egyszerűsödött a parciális differenciálegyenletünk. Ebben a fejezetben megnézzük egy konkrét példán (a Poisson-egyenleten), hogy hogy is történik ennek az egyenletrendszernek a felírása, valamint becslést adunk az ebből adódó hibára.

## 2.2 Rácsfelosztás, differencia hányadosok

Az egyszerűség kedvéért a tartomány amin dolgozunk legyen egy "a" és "b" oldalhosszú nyílt téglalap.

$$\Omega := \{(x_1, x_2) \mid x_1 \in (0, a), x_2 \in (0, b), a, b \in \mathbb{R}\}.$$

Az  $x_1$  tengely mentén  $N_1 + 1$ , az  $x_2$  tengely mentén  $N_2 + 1$  ekvidisztáns osztópontot veszünk fel. Az pontok távolsága  $h_1$ , illetve  $h_2$ . A pontokat amiknek mindkét koordinátája ilyen osztópont a megfelelő tengelyek mentén, a rács csomópontjainak hívjuk. A belső csomópontok halmazát  $w_h$ -val jelöljük

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

a peremen lévöket pedig  $\gamma_h$ -val

$$\gamma_h := \{\mathbf{x}^{i,j} := (x_{1,i}, x_{2,j}) \in \partial\Omega\}$$

Az összes csomópont halmazára a  $\bar{w}_h := w_h \cup \gamma_h$  szimbólumot használjuk, azaz  $|\bar{w}_h| = N := (N_1 + 1)(N_2 + 1)$ . A rács finomsága a  $h := \max\{h_1, h_2\}$  szám. Ezen a rácson fogjuk felírni a differencia hányadosokat.

**2.2.1. Jelölés:** Egy  $u$  függvény  $\mathbf{x}^{i,j}$ -pontbani értékét  $y_{i,j}$ -vel jelöljük, azaz

$$u_{i,j} := u(\mathbf{x}^{i,j}) = u(x_{1,i}, x_{2,j})$$

**2.2.2. Definíció:** Az

$$u_{x_1, i, j} := \frac{u_{i+1, j} - u_{i, j}}{h_1}$$

hányadost az  $\mathbf{x}^{i, j}$  pont beli  $x_1$  szerinti haladó differenciasémának nevezzük. (Azaz az  $u$  függvény  $\mathbf{x}^{i, j}$  pont beli  $x_1$  szerinti deriváltját ezzel a hányadossal közelítjük.)

**2.2.3. Definíció:** Az

$$u_{\bar{x}_1, i, j} := \frac{u_{i, j} - u_{i-1, j}}{h_1}$$

hányadost az  $\mathbf{x}^{i, j}$  pont beli  $x_1$  szerinti retrogád differenciasémának nevezzük.

**2.2.4. Megjegyzés:** Ezekkel a közelítésekkel a Taylor-formulából következően  $O(h)$  hibát követünk el haladó és retrogád esetben. Például a haladóra egy változóra felírva, hogy ne bonyolódjanak el a jelölések ez innen látszik:

$$u_{i+1} = u_i + u'_i h + u''(\alpha) \frac{h^2}{2} \quad \alpha \in [x_i, x_{i+1}]$$

Átrendezve és  $h$ -val osztva

$$\frac{u_{i+1} - u_i}{h} = u'_i + u''(\alpha) \frac{h}{2} \quad \alpha \in [x_i, x_{i+1}]$$

Azaz korlátos második derivált esetén valóban  $h$ -val arányos a hiba. A retrogádra ugyanez az érvelés, csak egyfelől vannak csúsztatva az indexek.

## 2.3 A módszer a Poisson-feladatra

**2.3.1. Definíció:** Legyen  $\Omega \subset \mathbb{R}^n$  tartomány,  $u: \Omega \rightarrow \mathbb{R}$ ,  $u \in C^2(\Omega)$  függvény. Ekkor

$$\Delta u(\mathbf{x}) := \partial_{x_1} \partial_{x_1} u(\mathbf{x}) + \dots + \partial_{x_n} \partial_{x_n} u(\mathbf{x})$$

az úgynevezett Laplace-operátor. (ha időtől is függő függvényünk van, akkor az idő szerinti deriváltat általában nem vesszük be az összegbe)

Legyen  $u$  mint a 2.3.1. definícióban  $n = 2$ -vel. Legyen  $\Omega = (0, a) \times (0, b)$ ,  $f: \Omega \rightarrow \mathbb{R}$  függvény. A módszert a

$$\begin{aligned} (-\Delta u)(\mathbf{x}) &= f(\mathbf{x}) & \mathbf{x} &= (x_1, x_2) \in \Omega \\ u(\mathbf{x}) &= g(\mathbf{x}) & \mathbf{x} &\in \partial\Omega \end{aligned} \tag{2.1}$$



egyenleten fogjuk megnézni. Ezt az egyenletet Poisson-feladatnak hívjuk, az  $u(\mathbf{x}) = g(\mathbf{x}), \mathbf{x} \in \partial\Omega$  feltétel neve Dirichlet-peremfeltétel.

A csomópontok indexelésére bevezetünk egy új jelölést.  $\mathbf{x}_k := \mathbf{x}^{i,j} \quad k = 1 \dots N$ , azaz a  $k$ -adik indexű elem a bal alsó sarokból indulva sorfolytonosan haladva a  $k$ -adik csomópont.

**2.3.1. Definíció:** Ha  $u \in C(\bar{\Omega})$  akkor  $u^h \in \mathbb{R}^N$ ,  $(u^h)_k := u(\mathbf{x}_k) \quad k = 1 \dots N$  az  $u$  függvény  $h = (h_1, h_2)$  felosztáshoz tartozó csomóponti értékeiből álló  $N$  dimenziós vektor.

A Laplace-operátort a csomópontokban a következő módon fogjuk közelíteni:

$$\Delta u \approx u_{x_1, \bar{x}_1} + u_{x_2, \bar{x}_2} \quad (2.2)$$

**2.3.2. Megjegyzés:** Ezzel a közelítéssel  $O(h^2)$  hibát követünk el. (az egyszerűség kedvéért megint egy változóra írjuk fel, ez nem okoz problémát hiszen " $O(h^2) + O(h^2) = O(h^2)$ ") Ezt is a Taylor-formulával fogjuk belátni. Írjuk fel  $u_{i+1}$  és  $u_{i-1}$  sorfejtését  $x_i$  körül negyedrendű maradéktaggal

$$\begin{aligned} u_{i+1} &= u_i + u_i' h + u_i'' \frac{h^2}{2} + u_i''' \frac{h^3}{6} + u_i''''(\alpha) \frac{h^4}{24} & \alpha \in [x_i, x_{i+1}] \\ u_{i-1} &= u_i - u_i' h + u_i'' \frac{h^2}{2} - u_i''' \frac{h^3}{6} + u_i''''(\beta) \frac{h^4}{24} & \beta \in [x_{i-1}, x_i] \end{aligned}$$

A két egyenletet összeadva és átrendezve azt kapjuk, hogy

$$u_{i+1} - 2u_i + u_{i-1} = u_i'' h^2 + (u_i''''(\beta) + u_i''''(\alpha)) \frac{h^4}{24} \quad \alpha \in [x_i, x_{i+1}], \beta \in [x_{i-1}, x_i]$$

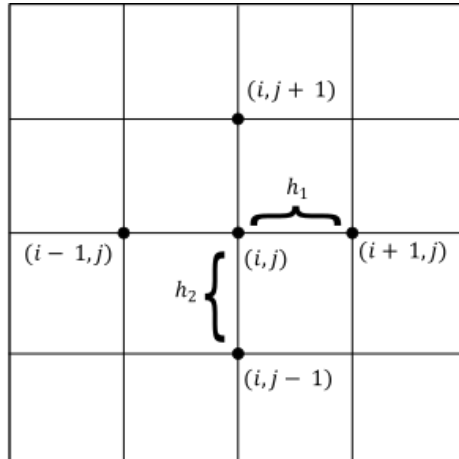
Folytonosság miatt  $u_i''''(\beta) + u_i''''(\alpha)$ -át kicserélhetjük  $2u_i''''(\gamma)$ -ra ahol  $\gamma \in [x_{i-1}, x_{i+1}]$ .  $h^2$ -el való osztás után

$$\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = u_i'' + u_i''''(\gamma) \frac{h^2}{12} \quad \gamma \in [x_{i-1}, x_{i+1}]$$

azaz valóban  $O(h^2)$  hibát követünk el ezzel a közelítéssel, ha  $u \in C^4(\Omega)$ .

Legyen tehát  $(\Delta_h u^h)$  közelítése  $\Omega$ -án

$$(\Delta_h u^h)_k := \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h_1^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h_2^2} \quad (2.3)$$



1.ábra.  $k = (i, j)$  indexen a közelítéshez felhasznált pontok

### 2.3.3. Megjegyzés:

- i. A  $k$  index itt nem 1-től  $N$ -ig megy, a peremen lévő pontokra nem írjuk fel ezt a közelítést. Ott egyszerűen a függvényértéknek kell az előre megadott  $(g(\mathbf{x}))$  számnak lennie.
- ii. Mivel a közelítést  $\Omega$ -án írtuk fel és legfeljebb eggyel kisebb vagy nagyobb indexű tagot használunk, így nem indexelünk ki  $\bar{\Omega}$ -ból.

A (2.3) differenciasémával a

$$(-\Delta u)(\mathbf{x}_k) = f(\mathbf{x}_k) \quad k = 1 \dots N$$

csomópontokon értelmezett egyenlőségek helyett a

$$-(\Delta_h u^h)_k = (f^h)_k \quad k = 1 \dots N$$

egyenlőségeket oldjuk meg, ahol már  $\Delta u$ -ra közelítést használunk.

### 2.3.4. Megjegyzés: $(f^h)_k := (g^h)_k$ ha $k$ a peremen lévő index.

Mivel  $(\Delta_h u^h)_k$  az  $u^h$  vektor koordinátáinak lineáris kombinációja, így fel tudjuk írni ezt az  $N$  darab egyenlőséget a következő lineáris egyenletrendszer alakban:

$$A_h u^h = f^h \tag{2.4}$$

Az  $A_h$  mátrix alakja a következő:



$B$  mellett lévő  $\frac{-1}{h^2}I$  részmátrixok tartoznak. A  $B$  előtti és utáni egy-egy 1-es az adott sorban lévő peremponti csomópontokhoz tartozik, melyekre nem használunk közelítést.

Ezzel a feladatunk ennek az egyenletrendszernek a megoldása. Ehhez bevezetjük az  $M$ -mátrix fogalmát.

**2.3.6. Definíció:** Egy  $A \in \mathbb{R}^{N \times N}$  mátrix  $M$ -mátrix, ha

- i.  $a_{i,j} \leq 0$  ha  $i \neq j$
- ii. van olyan  $v > 0$  vektor, melyre  $Av > 0$

**2.3.7. Állítás:** Ha  $A$   $M$ -mátrix akkor  $A$  reguláris és  $A^{-1} \geq 0$ . (az egyenlőtlenséget elemenként értjük)

**2.3.8. Állítás:** Ha  $A$   $M$ -mátrix és  $v > 0$  vektor, akkor  $\|A^{-1}\|_{\infty} \leq \frac{\max v}{\min Av}$ , ahol a maximumot és a minimumot koordinátánként értjük, ahol  $A \in \mathbb{R}^{n \times m}$ -re

$$\|A\|_{\infty} := \max_{i=1..m} \sum_{j=1}^n |a_{ij}| \quad (\text{sorösszegnorma})$$

**2.3.9. Állítás:**  $A_h$   $M$ -mátrix.

**Bizonyítás:**

- i.  $(\Delta_h u^h)_k$  definíciójában látszik, hogy amikor a  $k = (i, j)$ -edik  $\Delta_h u^h$  pontot közelítjük, akkor csak  $u_{i,j}$  együtthatója negatív, a többi pozitív.  $A_h$  elemei ezeknek a  $(-1)$ -szeresei, azaz  $a_{i,j} \leq 0$  ha  $i \neq j$ .
- ii. Legyen  $w: \Omega \rightarrow \mathbb{R}$ ,  $w(x_1, x_2) = x_1(a - x_1) + x_2(b - x_2)$  ( $\Omega = (0, a) \times (0, b)$ )  
Legyen  $v := w^h$  ( $h$  felosztáshoz tartozó csomóponti értékek vektora). Ekkor nyilván  $v > 0$ . Továbbá  $-\Delta w \equiv 4$ , azaz

$$(A_h v)_k = (-\Delta w^h)_k = -\Delta w(x_k) = 4 > 0 \quad k = 1 \dots N$$

Ahol  $x_k$  a csomóponti helyekből álló vektor  $k$ -adik koordinátája.

**2.3.10. Következmény:** Az  $A_h u^h = f^h$  lineáris algebrai egyenletrendszernek létezik egyértelmű megoldása.

**2.3.11. Állítás:** Az  $A_h$  mátrixra  $\|A_h^{-1}\|_{\infty} \leq \frac{a^2+b^2}{16}$ .

**Bizonyítás:** Legyen  $w: \Omega \rightarrow \mathbb{R}$ ,  $w(x_1, x_2) = x_1(a - x_1) + x_2(b - x_2)$  ( $\Omega = (0, a) \times (0, b)$ ), továbbá legyen  $v := w^h$ . Ekkor 2.3.8. állítást felhasználva:

$$\|A_h^{-1}\|_{\infty} \leq \frac{\max v}{\min Av} = \frac{\max_k (w^h)_k}{\min_k (-\Delta_h w^h)_k} = \frac{1}{4} \max_k (w(x_k)) \leq \frac{1}{4} \max_{x \in \Omega} (w(x)) = \frac{a^2 + b^2}{16}$$

**2.3.12. Megjegyzés:** Ha elsőfajú peremfeltételünk van, azaz az  $u$  függvényértékei vannak megadva a peremen, akkor az  $A_h$ -mátrixunk méretét tudjuk csökkenteni a következő kétlépéses módon:

- i. A peremen lévő csomópontok  $u_i$  értékeket ismerjük a feltétel miatt. Az  $i$ -edik oszlop  $u_i$ -szeresét levonjuk a jobboldali vektorból, és az  $i$ -edik oszlopot kinullázzuk. Ezt megteesszük az összes peremen lévő pontra.
- ii. A felesleges (csupa 0) sorokat elhagyjuk.

Az így kapott mátrix is M-mátrix, ez megfelel az eredetileg felírt  $A_h$  mátrixnak  $g \equiv 0$  peremfeltétellel.

## 2.4 Stabilitás és konvergencia

**2.4.1. Állítás:** Az  $A_h u^h = f^h$  lineáris egyenletrendszerre

- i.  $\|u^h\|_{\infty} \leq \frac{a^2 + b^2}{16} \|f^h\|_{\infty}$
- ii. az  $u^h$  megoldásvektor hibája konstanszorosa a jobboldali  $f^h$  hibájának.

**Bizonyítás:**

- i. 2.3.11. állítás és normatulajdonság alapján

$$\|u^h\|_{\infty} \leq \|A_h^{-1}\|_{\infty} \|f^h\|_{\infty} \leq \frac{a^2 + b^2}{16} \|f^h\|_{\infty}$$

- ii. legyenek  $f_1, f_2$  jobboldalak  $u_1, u_2$  megoldásokkal. Ekkor a linearitást kihasználva

$$\|u_1 - u_2\|_{\infty} \leq \|A_h^{-1}\|_{\infty} \|f_1 - f_2\|_{\infty} \leq \frac{a^2 + b^2}{16} \|f_1 - f_2\|_{\infty}$$

**2.4.2. Definíció:** Legyen  $u \in C^2(\bar{\Omega})$ . A fenti differenciaséma képlethibáján a

$$\psi^h := (\Delta u)^h - \Delta_h u^h$$

vektort értjük. (a kivonást koordinátánként értjük)

**2.4.3. Állítás:** Ha  $u \in C^4(\bar{\Omega})$ , akkor a képlethibára

$$\|\psi^h\|_\infty \leq \frac{1}{6} |u|_{C^4} h^2$$

ahol  $|u|_{C^4} := \max_{|\alpha|=4} \|\partial^\alpha u\|_\infty$ . (Az  $|\alpha|$  a parciális deriválás rendje.)

**2.4.4. Definíció:** Egy  $a \in \mathbb{R}^N$  vektor végtelen normáján a következő számot értjük:

$$\|a\|_\infty := \max_{i=1,\dots,N} |a_i|$$

**2.4.5. Jelölés:** Jelölje  $u_*^h \in C^2(\bar{\Omega})$  az eredeti egyenlet megoldását a csomópontokon,  $u^h$  a lineáris egyenletrendszer megoldását.

A konvergenciavizsgálathoz az

$$e^h := u_*^h - u^h$$

csomóponti hibavektort fogjuk becsülni végtelen normában.

**2.4.6. Definíció:** Azt mondjuk, hogy egy módszer  $p$ -edrendben konvergens ( $p \geq 1, p \in \mathbb{N}$ ), ha

$$\|e^h\|_\infty = O(h^p)$$

**2.4.7. Állítás:** Az  $A_h$  mátrix a csomóponti hibavektort a képlethibavektorba viszi, azaz

$$A_h e^h = \psi^h$$

**Bizonyítás:** Minden  $v \in \mathbb{R}^N$  vektorra  $A_h v = -\Delta_h v$ , így a pontos megoldás csomóponti vektorára is, azaz

$$A_h e^h = A_h (u_*^h - u^h) = -\Delta_h (u_*^h - u^h) = -\Delta_h u_*^h + \Delta_h u^h = -\Delta_h u_*^h + f^h$$

De mivel  $u_*$  pontos megoldás, így  $(-\Delta u)^h = f^h$ , azaz

$$A_h e^h = -\Delta_h u_*^h + f^h = -\Delta_h u_*^h - (-\Delta u)^h = (\Delta u)^h - \Delta_h u_*^h = \psi^h$$

Így használhatjuk a korábbi hibabecsléseinket.

**2.4.8. Állítás:** Ha  $u_* \in C^4(\bar{\Omega})$ , akkor a

$$(\Delta_h u^h)_k := \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h_1^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h_2^2}$$

differenciasémával a véges differencia módszer másodrendben konvergál, azaz

$$\|e^h\|_\infty \leq \frac{a^2 + b^2}{96} |u|_{C^4} h^2$$

**Bizonyítás:** 2.3.5. és 2.3.13. állításokat felhasználva

$$\|e^h\|_\infty \leq \|A_h^{-1}\|_\infty \|\psi^h\|_\infty \leq \frac{a^2+b^2}{16} \frac{1}{6} |u|_{C^4} h^2 = \frac{a^2+b^2}{96} |u|_{C^4} h^2$$

# 3 Végeelem módszer

Ebben a fejezetben [5]-öt követjük.

## 3.1 Megoldási ötlet

A második fejezetben bemutatott véges differenciák módszer egyik gyengéje, hogy csak speciális alakú egyenleteket tudunk vele megoldani. Például a Poisson-egyenlet megoldásában erősen kihasználtuk a módszer által létrejött lineáris egyenletrendszer  $M$ -mátrix tulajdonságát. További hátrány, hogy nem tudjuk tetszőleges tartományon használni a módszert. Ez motivációt ad más eszköz használatára. A végeelem módszernek a lényege, hogy a tartományt felbontjuk véges sok résztartományra (véges elemekre; ezek általában háromszögek vagy téglalapok, illetve ezek magasabb dimenziójú megfelelőik) és ezeken külön-külön polinomokkal közelítjük a megoldást. Közben figyelünk arra, hogy a tartományok határán folytonosak legyenek ezek a polinomok. A módszer alapja az úgynevezett Ritz-Galjorkin-módszer melynek lényege, hogy a végtelen dimenziós operátoregyenletet levetítjük annak egy véges dimenziós alterére. Így kapunk egy véges dimenziós feladatot, és ezt oldjuk meg valamilyen numerikus módszerrel.

## 3.2 A Ritz-Galjorkin módszer

Ebben az alfejezetben  $H$  valós Hilbert-teret fog jelölni  $\langle \cdot, \cdot \rangle$ -skalárszorzattal.

**3.2.1. Definíció:** Egy  $A: H \rightarrow H$  korlátos lineáris operátor

- Pozitív, ha  $\langle Ax, x \rangle \geq 0 \quad \forall x \in H$
- Szigorúan pozitív, ha  $\langle Ax, x \rangle > 0 \quad \forall x \in H, x \neq 0$
- Egyenletesen pozitív, ha  $\exists m > 0$ , hogy  $\langle Ax, x \rangle \geq m\|x\|^2 \quad \forall x \in H$

**3.2.2. Definíció:** Egy  $A: H \rightarrow H$  lineáris operátor szimmetrikus, ha

$$\langle Ax, y \rangle = \langle x, Ay \rangle \quad \forall x, y \in H$$

**3.2.3. Definíció:** Legyen  $A: H \rightarrow H$  szigorúan pozitív operátor. Az

$$\langle u, v \rangle_A := \langle Au, v \rangle$$

formát ( $H \times H \rightarrow \mathbb{R}$  leképezés) az  $A$ -hoz tartozó energia-skalárszorzatnak nevezzük. A

$$H_A := (H, \langle \cdot, \cdot \rangle_A)$$

Hilbert teret pedig energiaternek mondjuk.

**3.2.4. Tétel:** Legyen  $A: H \rightarrow H$  szimmetrikus és egyenletesen pozitív operátor. Legyen  $f \in H$ , és tekintsük az

$$Au = f \tag{3.1}$$

operátoregyenletet. Ennek létezik egyértelmű  $u^* \in H_A$  gyenge megoldása, azaz

$$\langle u^*, v \rangle_A = \langle f, v \rangle \quad \forall v \in H_A \tag{3.2}$$

Igaz továbbá, hogy  $u^*$  a

$$\phi: H_A \rightarrow \mathbb{R}, \quad \phi(u) := \|u\|_A^2 - 2\langle f, u \rangle$$

kvadratikus funkcionál egyetlen minimumhelye  $H$ -n.

**Bizonyítás:** Legyen  $\varphi: H_A \rightarrow \mathbb{R}$ ,  $\varphi(v) := \langle v, f \rangle$  lineáris funkcionál. Ez folytonos, mert

$$|\varphi(v)| = |\langle v, f \rangle| \leq \|v\| \|f\| \leq \frac{1}{\sqrt{m}} \|v\|_A \|f\|$$

így korlátos is, ezért a Riesz-reprezentációs tétel szerint egyértelműen létezik  $u^* \in H_A$ , hogy

$$\langle v, f \rangle = \varphi(v) = \langle v, u^* \rangle_A$$

A tétel második részéhez felhasználjuk, hogy mivel  $u^* \in H_A$  gyenge megoldás, ezért

$$\phi(u^*) = \|u^*\|_A^2 - 2\langle f, u^* \rangle = \langle u^*, u^* \rangle_A - 2\langle f, u^* \rangle = \langle Au^*, u^* \rangle - 2\langle Au^*, u^* \rangle = -\langle Au^*, u^* \rangle$$

Legyen  $u \in H_A$  tetszőleges, amire  $u^* \neq u$ . Ekkor

$$\begin{aligned} \phi(u) &= \langle Au, u \rangle - \langle f, u \rangle - \langle u, f \rangle = \langle Au, u \rangle - \langle Au^*, u \rangle - \langle u, Au^* \rangle = \\ &= \langle A(u - u^*), (u - u^*) \rangle - \langle Au^*, u^* \rangle = \langle A(u - u^*), (u - u^*) \rangle + \phi(u^*) > \phi(u^*) \end{aligned}$$

Ahol használtuk az előző észrevételt és hogy valós Hilbert-téren vagyunk.

**3.2.5. Megjegyzés:** A 3.2.4. tétel szemléletes jelentése, hogy az  $Au = f$  egyenlet gyenge megoldása helyett elég a  $\phi(u)$  funkcionál minimumhelyét megkeresni. Azonban ez végtelen dimenziós Hilbert-tér esetén nehéz, a következőkben megnézzük hogyan tudjuk ezt a problémát kezelni.

Tekintsünk  $\varphi_1, \varphi_2, \dots$  lineárisan független vektorokat melyek lineáris burka sűrű  $H_A$ -ban (itt keressük a megoldást). Legyen

$$H_n := \text{span}(\varphi_1, \varphi_2, \dots, \varphi_n) \quad n \in \mathbb{N}$$



$n$  dimenziós altér  $H$ -ban (a  $\text{span}(\varphi_1, \varphi_2, \dots, \varphi_n)$  a  $\varphi_1, \varphi_2, \dots, \varphi_n$  által generált alteret jelöli). Ebben az altérben már meg tudjuk oldani a minimalizációs feladatot. A  $\phi$  funkcionál minimumhelyét ezekkel a véges dimenziós megoldásokkal fogjuk közelíteni. Ezt technikát nevezik Ritz-Galjorkin-módszernek.

**3.2.6. Megjegyzés:** Fontos, hogy ekkor  $\phi$ -re  $H_n \rightarrow \mathbb{R}$  függvényként tekintünk, hiszen a minimumhely így lesz  $H_n$ -beli. Ez a minimum szintén létezik és egyértelmű a 3.2.4. tétel miatt.

Jelölje a  $H_n$ -beli minimumot  $u_n$ , azaz

$$\phi(u_n) = \min_{H_n} \phi \quad (3.3)$$

Mivel  $\varphi_1, \varphi_2, \dots, \varphi_n$  bázis  $H_n$ -ben, ezért

$$u_n = \sum_{i=1}^n b_i \varphi_i \quad b_i \in \mathbb{R}$$

megfelelő  $b_i$  konstansokkal.  $u_n$  előállításához bevezetünk egy segédfüggvényt:

$$\tilde{\phi}: \mathbb{R}^n \rightarrow \mathbb{R}, \quad \tilde{\phi}(\mathbf{c}) := \phi\left(\sum_{i=1}^n c_i \varphi_i\right)$$

Ekkor  $\tilde{\phi}$  definíciója miatt  $\tilde{\phi}$   $\mathbb{R}^n$ -ben való minimalizása ekvivalens  $\phi$   $H_n$ -en való minimalizálásával.  $\tilde{\phi}$  minimumhelye a deriváltjának a zérushelye lesz (3.2.7. megjegyzés).

**3.2.7. Megjegyzés:** Ez a 3.2.4. tétel bizonyításából látszik: egyetlen minimumhely van, és “minél távolabb vagyunk ettől a minimumhelytől, annál nagyobb  $\phi$  értéke”. Precízebben

$$\phi(u) = \langle A(u - u^*), (u - u^*) \rangle + \phi(u^*) = \langle u - u^*, u - u^* \rangle_A + \text{konstans}$$

Ahol az első tag  $u$  és  $u^*$   $\langle \cdot, \cdot \rangle_A$ -norma szerinti távolsága, a konstans pedig  $\phi$  egyértelmű minimuma. E miatt az  $u^*$ -beli minimum lesz  $\phi$  egyetlen szélsőértéke.

A  $\tilde{\phi}$  függvény alakja, hogy lássuk mit deriválunk:

$$\begin{aligned} \tilde{\phi}(\mathbf{c}) &= \phi\left(\sum_{i=1}^n c_i \varphi_i\right) = \left\langle \sum_{i=1}^n c_i \varphi_i, \sum_{i=1}^n c_i \varphi_i \right\rangle_A - 2 \left\langle f, \sum_{i=1}^n c_i \varphi_i \right\rangle = \\ &= \sum_{i=1}^n \sum_{j=1}^n c_i c_j \langle \varphi_i, \varphi_j \rangle_A - 2 \sum_{i=1}^n c_i \langle f, \varphi_i \rangle \end{aligned}$$

Azaz a  $c_k$  szerinti parciális derivált így néz ki:

$$\partial_{c_k} \tilde{\phi}(c) = 2 \sum_{i=1}^n c_i \langle \varphi_i, \varphi_k \rangle_A - 2 \langle f, \varphi_k \rangle$$

Tehát ezek kell nullának lenni minden  $k$ -ra, ez azzal ekvivalens, hogy

$$\sum_{i=1}^n c_i \langle \varphi_i, \varphi_k \rangle_A = \langle f, \varphi_k \rangle \quad k = 1, \dots, n \quad (3.4)$$

Ezeket az egyenleteket fel tudjuk írni egy lineáris egyenletrendszerbe, aminek megoldásával megkapjuk a keresett  $u_n$   $c_i$  együtthatóit. Erre az egyenletrendszerre bevezetjük a következő jelölést:

$$Gc = b \quad (3.5)$$

ahol  $G_{ik} := \langle \varphi_i, \varphi_k \rangle_A$  és  $b_k := \langle f, \varphi_k \rangle$   $k = 1, \dots, n$ . A  $G$  mátrixot Gram-mátrixnak hívjuk. Mivel a  $\langle \cdot, \cdot \rangle_A$  skalárszorzat valós értékű, ezért a  $G$  mátrix szimmetrikus. (Nyilván minden  $n$ -hez külön  $Gc = b$  egyenletrendszer tartozik.) Ez az egyenletrendszer megoldható, mivel  $\varphi_1, \dots, \varphi_n$  lineárisan függetlenek.

**3.2.8. Állítás:** Az ezzel a módszerrel (3.5) megoldása adja  $u_n$  együtthatóit a  $\varphi_1, \dots, \varphi_n$  bázisban) keletkező  $u_n$  megoldássorozat  $\|\cdot\|_A$ -normában konvergens, és határértéke  $u^*$ .

**Bizonyítás:**  $\varphi_1, \varphi_2, \dots$  lineáris burka sűrű  $H_A$ -ban (ezt tettük fel róluk) és  $\phi$  folytonos  $H_A$ -n, ebből következően

$$\min_{H_n} \phi \rightarrow \min_{H_A} \phi \quad (n \rightarrow \infty), \quad \text{azaz} \quad \phi(u_n) \rightarrow \phi(u^*) \quad (n \rightarrow \infty)$$

$\phi$  definícióját használva

$$\begin{aligned} \phi(u_n) - \phi(u^*) &= (\langle u_n - u^*, u_n - u^* \rangle_A + \phi(u^*)) - (\langle u^* - u^*, u^* - u^* \rangle_A + \phi(u^*)) = \\ &= \langle u_n - u^*, u_n - u^* \rangle_A = \|u_n - u^*\|^2 \rightarrow 0 \end{aligned}$$

**3.2.9. Állítás:**  $u_n \in H_n$ -re az alábbiak teljesülnek:

- i.  $u_n$  a  $\phi: H_n \rightarrow \mathbb{R}$  egyértelmű minimumhelye (közelítő minimalizálás)
- ii.  $\forall v \in H_n$ -re  $\langle u_n, v \rangle_A = \langle f, v \rangle$  (vetületi egyenlet)
- iii.  $\forall v \in H_n$ -re  $\langle u_n - u^*, v \rangle_A = 0$  (a hiba ortogonalitása)

**Bizonyítás:**

- i. Ez a definíciója  $u_n$ -nek
- ii. Megmutatjuk  $H_n$  bázisvektoraira, ebből már következik tetszőleges  $v \in H_n$  -re

$$\langle u_n, \varphi_k \rangle_A = \left\langle \sum_{i=1}^n c_i \varphi_i, \varphi_k \right\rangle_A = \sum_{i=1}^n c_i \langle \varphi_i, \varphi_k \rangle_A = \langle f, \varphi_k \rangle$$

iii. Mivel  $\langle u_n, v \rangle_A = \langle f, v \rangle \quad \forall v \in H_n$ , és  $\langle u^*, v \rangle_A = \langle f, v \rangle \quad \forall v \in H_A$ , ezért

$$\langle u_n - u^*, v \rangle_A = \langle u_n, v \rangle_A - \langle u^*, v \rangle_A = \langle f, v \rangle - \langle f, v \rangle = 0$$

A Ritz-Galjorkin módszert tudjuk általánosítani azzal, ha minden  $n$ -re különböző bázisfüggvényeket választunk, azaz

$$H_n := \text{span}(\varphi_1^{(n)}, \varphi_2^{(n)}, \dots, \varphi_n^{(n)}) \quad n \in \mathbb{N}$$

és ezzel szinte az összes kapott eredmény érvényben marad, hiszen nem használtuk fel, hogy a  $(H_n)_{n \in \mathbb{N}}$  sorozat bővülő lenne. Egyedül 3.2.8. állításhoz (konvergencia) kell egy új kikötést tennünk. Erre szolgál a következő feltétel.

**3.2.10. Feltétel:** Minden  $u \in H_A$ -ra:

$$\text{dist}_A(u, H_n) := \min\{\|u - v_n\|_A : v_n \in H_n\} \rightarrow 0 \quad (n \rightarrow \infty)$$

(azért írhatunk minimumot, mert  $H_n \subset H_A$  egy nem üres, konvex, zárt halmaz és ekkor egyértelműen létezik  $v \in H_n$  amire  $\text{dist}_A(u, H_n) = \|u - v\|_A$  [5])

**3.2.11. Állítás:** Ha 3.2.10. felétel teljesül, akkor az  $u_n$  megoldássorozat  $\|\cdot\|_A$ -normában konvergens, és határértéke  $u^*$ .

**Bizonyítás:** 3.2.9. állítás iii, pontja miatt  $\forall v \in H_n$ -re  $\langle u_n - u^*, v \rangle_A = 0$ , tehát  $u_n - u^*$  merőleges  $H_n$ -re. Ez azt fejezi ki, hogy  $u^*$  vetülete  $H_n$ -re  $u_n$ , azaz a minimum  $u_n$ -en vétetik fel:

$$\min\{\|u - v_n\|_A : v_n \in H_n\} = \|u^* - u_n\|_A \rightarrow 0 \quad (n \rightarrow \infty)$$

### 3.3 A Ritz-Galjorkin módszer használata

Először bevezetjük a szükséges definíciókat. Az alfejezetben feltesszük, hogy  $\Omega \subset \mathbb{R}^n$  korlátos tartomány szakaszonként sima peremmel. Ez alatt azt értjük, hogy  $\partial\Omega$  szakaszonként egy folytonosan differenciálható függvény képe. Feltesszük továbbá, hogy  $\partial\Omega$  Lipschitz-folytonos, azaz ha  $g$   $\partial\Omega$ -nak egy paraméterezése, akkor  $g$  minden koordinátafüggvénye Lipschitz-folytonos. (Mindegy milyen normában nézve, ugyanis véges dimenziós vektortéren minden norma ekvivalens.)

**3.3.1. Definíció[2]:**  $H^k(\Omega) := \{u \in L^2(\Omega) : \partial^\alpha u \in L^2(\Omega) \ \forall |\alpha| \leq k\}$  ahol  $\partial^\alpha u$  általánosított deriváltat jelent disztribúció értelemben (ennek pontos definíciója megtalálható [3]-ban). Illetve  $H_0^1(\Omega) := \{u \in H^1(\Omega) : u|_{\partial\Omega} = 0\}$ .

A  $H_0^1(\Omega)$ -át valós értékű függvényekkel definiáljuk ebben a fejezetben.

Először megnézzük, hogy a Ritz-Galjorkin módszerben szereplő absztrakt fogalmak konkrétan minek felelnek meg egy parciális differenciálegyenlet megoldása során. Ezt a következő példán keresztül fogjuk megtenni:

$$\begin{aligned} -\operatorname{div}(p(x)\nabla u(x)) &= f(x) & \mathbf{x} \in \Omega \\ u(\mathbf{x}) &= 0 & \mathbf{x} \in \partial\Omega \end{aligned} \quad (3.6)$$

Ahol  $\operatorname{div}$  divergenciát és  $\nabla$  gradienst jelöl,  $p, u, f$  pedig  $\Omega \rightarrow \mathbb{R}$  függvények. Továbbá  $p \in L^\infty(\Omega)$ ,  $p(\mathbf{x}) \geq m > 0$  (m.m.  $\mathbf{x} \in \Omega$ ),  $f \in L^2(\Omega)$ . (Ennek egyébként speciális esete az előző fejezetben tárgyalt Poisson-egyenlet  $p \equiv 1$ -el.)

Ebben a példában a Hilbert-terünk a  $H := L^2(\Omega)$  tér lesz, az  $A$  operátorunk pedig

$$Au := -\operatorname{div}(p\nabla u)$$

Ami szimmetrikus és egyenletesen pozitív az  $u|_{\partial\Omega} = 0$  peremfeltétellel.

Az energiater a  $H_A := H_0^1(\Omega)$  tér lesz a következő skalárszorzattal:

$$\langle u, v \rangle_A = \langle Au, v \rangle = \int_\Omega Au v = \int_\Omega -\operatorname{div}(p\nabla u) v = \int_\Omega p\nabla u \cdot \nabla v \quad u, v \in H_0^1(\Omega) \quad (3.7)$$

Itt felhasználtuk Green első formuláját és hozzá hogy  $v \in H_0^1(\Omega)$ , azaz  $v|_{\partial\Omega} = 0$  és ezért hagyhattuk el a tagot amiben  $\partial\Omega$ -án való integrálás van.

A (3.6) feladat gyenge megoldása:

$$\int_\Omega -\operatorname{div}(p\nabla u) v = \int_\Omega f v \quad \forall v \in H_0^1(\Omega)$$

(3.7)-et felhasználva ez a következő alakban írható:

$$\int_\Omega p\nabla u \cdot \nabla v = \int_\Omega f v \quad \forall v \in H_0^1(\Omega)$$

**3.3.2. Tétel:** Ennek problémának egyértelműen létezik  $u^* \in H_0^1(\Omega)$  gyenge megoldása.

**Bizonyítás:** A fenti definíciókkal a feladat megfelel a 3.2.4. tételnek.

Tekintsünk egy alkalmas  $V_h \subset H_0^1(\Omega)$  véges dimenziós alteret és benne egy  $\varphi_1, \varphi_2, \dots, \varphi_n$  bázist (Hogy konkrétan ez mi legyen, arról a következő fejezetben lesz szó). Az  $u_h \in V_h$  közelítő megoldást együtthatóit a

$$Gc = b$$

lineáris egyenletrendszer megoldásával kapjuk. Itt

$$G_{ik} = G_{ki} = \int_{\Omega} p \nabla \varphi_i \cdot \nabla \varphi_k \quad \text{és} \quad b_k = \int_{\Omega} f \varphi_k$$

Ez az  $u_h$  függvény teljesíti a vetületi egyenletet, azaz

$$\int_{\Omega} p \nabla u_h \cdot \nabla v = \int_{\Omega} f v \quad \forall v \in V_h$$

Hátravan még, hogy a  $V_h$  véges dimenziós altereket hogyan választjuk. Ehhez [2]-t fogjuk követni.

**3.3.3. Definíció:** Egy  $\Omega$  tartomány triangulációjának nevezzük a

$$\mathcal{T}_h := \{T_1, \dots, T_M\}$$

halmazt, ahol

- i. Minden  $k = 1, \dots, M$  esetén  $T_k \subset \Omega$  sokszög.
- ii.  $T_1, \dots, T_M$  a tartomány nemátfedő felbontását alkotja, azaz az  $int T_k$  halmazok páronként diszjunktak és  $\cup T_k = \bar{\Omega}$
- iii. A felbontás konform, azaz  $T_k \cap T_l$  ( $k \neq l$ ) halmazok csak csúcsokból vagy teljes oldalkból (élekből vagy lapokból) állhatnak.

**3.3.4. Definíció:** A  $\mathcal{T}_h$  trianguláció finomsága alatt a

$$h := \max_{i=1, \dots, M} diam(T_k)$$

számot értjük, ahol  $diam(T_k)$  a  $T_k$  elem átmérőjét jelöli.

**3.3.5. Definíció:** A  $V_h$  altér legyen

$$V_h := \{u \in C(\bar{\Omega}) : u|_{T_k} \in P^{l_k}(T_k) \quad k = 1, 2, \dots, M\}$$

ahol  $P^{l_k}(T_k)$  jelöli a legfeljebb  $l_k$ -adfokú polinomok  $T_k$ -ra való leszűkítésének halmazát.

Általában az alábbi tulajdonságú altereket használjuk:

- i.  $T_k$ -k azonos típusúak, például a trianguláció csupa háromszögből vagy csupa négyszögből áll (illetve ezek magasabb dimenziós megfelelőiből)

- ii.  $l_k \equiv l$ , azaz azonos fokú polinomokat tekintünk az összes elemen
- iii. Alacsony fokszámú, például első vagy másodfokú polinomokkal dolgozunk, magasabb fokszám akkor indokolt, ha javítani akarunk a megoldás simaságán vagy konvergenciarendjén

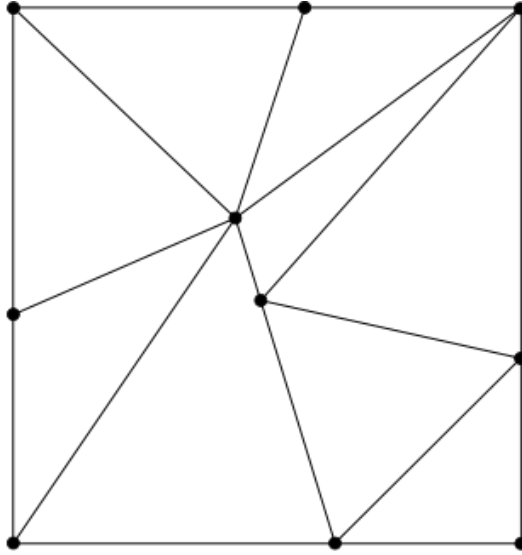
**3.3.6. Megjegyzés:** Az  $l$ -edfokú polinomokat a  $T_k$ -ben kijelölt csomóponti értékek határozzák meg. Ezek lehetnek függvényértékek és deriváltértékek is. Ha a csomóponti értékek csak függvényértékek, akkor a bázis olyan polinomokból áll, amelyek egy adott csomópontban 1-et, a többin 0-át vesznek fel. Tehát ha  $x_1, \dots, x_r$  csomópontok, akkor

$$\varphi_i(x_j) := \delta_{ij} = \begin{cases} 1 & \text{ha } i = j \\ 0 & \text{egyébként} \end{cases}$$

Példaként nézzünk meg egy egyszerű esetet: 2 dimenzióban vagyunk, az elemeink háromszögek és  $u|_{T_k}$  legyen elsőfokú polinom. Ekkor  $T_k$  elemből három csomópontra lesz szükségünk, hiszen az

$$(x, y) \mapsto a_k + b_k x + c_k y$$

függvényben három szabad paraméter van. Válasszuk csomópontoknak a háromszög csúcsait. Ez azért lesz jó, mert így az egész  $\bar{\Omega}$ -on értelmes folytonos függvényt kapunk, mivel egy-egy oldalra való leszűkítésen a két csúcsbeli függvényérték egyértelműen meghatározza az oldalon vett lineáris függvényt. A folytonosságot tehát a csomópontok megfelelő választásával garantáljuk.



2. ábra. Trianguláció és a hozzá tartozó csomópontok

### 3.4 Stabilitás és konvergencia

A stabilitást és konvergenciát a (3.6) feladatra felírt végelem módszerre fogjuk megnézni.

**3.4.1. Állítás:** Egy  $\{V_h\}_{h>0}$  család esetén, ha a felosztás finomsága 0-hoz tart, akkor az elemekre és polinomokra tett nem túl szigorú feltételekkel[5] elérhető, hogy bármely  $u \in H_0^1(\Omega)$  esetén

$$\text{dist}_{H_0^1(\Omega)}(u, V_h) = \min \{ \|u - v_h\|_{H_0^1(\Omega)} : v_h \in V_h \} \rightarrow 0 \quad (h \rightarrow 0)$$

**3.4.2. Következmény:** A 3.4.1. állításban szereplő feltételekkel a felírt végelem módszer konvergens.

**3.4.3. Állítás:** A felírt végelem módszerrel az  $u_h \in V_h$  megoldásra

$$\|u_h\| \leq \frac{1}{m} \|f\|$$

**Bizonyítás:** 3.2.9. állítás ii, pontja miatt

$$m\|u_h\|^2 \leq \langle Au_h, u_h \rangle = \langle u_h, u_h \rangle_A = \langle f, u_h \rangle \leq \|f\| \|u_h\|$$

Ezt átrendezve kapjuk az állítást.

**3.4.4. Következmény:** Ha  $f_1$  és  $f_2$  jobboldalakhoz tartozó megoldásvektorok  $u_1^h$  és  $u_2^h$  akkor

$$\|u_1^h - u_2^h\| \leq \frac{1}{m} \|f_1 - f_2\|$$

# 4 Véges térfogat módszer

## 4.1 Megoldási ötlet

A módszert azzal kezdjük, hogy felosztjuk a tartományunkat résztartományokra. A résztartományok csúcsai lesznek a rácspontok, amin a keresett függvény értékeit becsüljük. Egy adott rácspontbeli érték becsüléséhez, az adott rácspont környezetén vett integrált (térfogatot) fogjuk használni. Ezt nyilván nem ismerjük, de a pont környezetén vett integrálást a Gauss-Osztrogradszkij tétel segítségével át tudjuk alakítani a környezet peremén vett integrálásra. A peremen vett integrálra adunk valami közelítést (pl: kvadratúraformulával) a rácspontbeli függvényértékek felhasználásával. Ezt felírva azt összes rácspontra egy lineáris egyenletrendszerhez jutunk, mely megoldásával kapjuk a keresett függvény rácsponti értékeire való becsülésünket.

Ahhoz, hogy a Gauss-Osztrogradszkij tétel használni tudjuk annyi kell, hogy a differenciálegyenlet valamilyen fluxus divergenciáját tartalmazza és adott forrásokat. Pontosan ilyen formában írhatók fel a fizika megmaradási tételei. Ezért mérnöki számításoknál gyakran ezt a módszert használják, például a konvekció-diffúzió és az elsőrendű hiperbolikus egyenletek könnyen kezelhetők vele.

## 4.2 A módszer bemutatása

Legyen  $\Omega \subset \mathbb{R}^2$ , véges tartomány,  $\partial\Omega := \Gamma = \Gamma_1 \cup \Gamma_{23}$  olyan zárt töröttvonal (0 belső és külső szögek nélkül), melynek legfeljebb véges sok töréspontja van. A módszert az alábbi feladaton fogjuk megnézni:

$$\begin{aligned} \operatorname{div}(k(\mathbf{x})\nabla u(\mathbf{x})) + f(\mathbf{x}) &= 0 & \mathbf{x} \in \Omega \\ u(\mathbf{x}) &= g_1(\mathbf{x}) & \mathbf{x} \in \Gamma_1, \quad k(\mathbf{x})\frac{\partial u}{\partial n}(\mathbf{x}) + \sigma(\mathbf{x})u(\mathbf{x}) &= g_{23}(\mathbf{x}) & \mathbf{x} \in \Gamma_{23} \end{aligned} \tag{4.1}$$

ahol  $n$  a perem külső normálvektora. Az együtthatókról feltesszük, hogy

$$\begin{aligned} \sigma(\mathbf{x}) &\geq \sigma_0 > 0 & \mathbf{x} \in \Gamma_3, \quad \sigma(\mathbf{x}) &\equiv 0 & \mathbf{x} \in \Gamma_2, \quad \Gamma_2 \cup \Gamma_3 = \Gamma_{23} \\ 0 < k_0 &\leq k(\mathbf{x}) \leq k_1 & \mathbf{x} \in \bar{\Omega} \end{aligned}$$

Továbbá megköveteljük az alábbi simaságra vonatkozó feltételek teljesülését  $j = 0,1$ -re:



$$k \in C^{1+j}(\bar{\Omega}), f \in H^{1+j}(\Omega), \sigma, g_{23} \in H^{1+j}(\Gamma_{23}), u \in H^{2+j}(\Omega), g_1 \in C(\Gamma_1)$$

Most megnézzük, hogy ezen módszer esetén hogyan történik a diszkretizáció. Első lépésként felosztjuk a tartományunkat egy téglalap alakú nem ekvidisztáns rács segítségével, melyek egyenesei a kordinátatengelyekkel párhuzamosak. Az  $x_1$  és  $x_2$  tengelyeken lévő osztópontokat  $x_{1i}$  illetve  $x_{2j}$  -vel indexeljük. A rács egy tipikus pontjára az  $\mathbf{x}^0 = (x_{1i}, x_{2j})$  jelölést használjuk. A rácsot olyannak választjuk, hogy a perem töréspontjai,  $\Gamma_1$  és  $\Gamma_{23}$  közötti átmenetek (itt változnak a peremfeltételek), és  $\Gamma$  metszéspontjai a rács egyenesével rácsponatok legyenek. Ez elérhető, ha elég finomnak választjuk a rácsot. Ez a  $\Omega$  -át belül “elemi” téglalapokra osztja fel, a peremnél téglalapok helyett keletkezhetnek derékszögű háromszögek is. A rács egyenesei közötti távolságokra bevezetjük a következő jelölést:

$$h_{1,i\pm 1/2} = h_{1i}^\pm := |x_{1,i\pm 1} - x_{1i}|, \quad h_{2,j\pm 1/2} = h_{2j}^\pm := |x_{2,j\pm 1} - x_{2j}|$$

illetve ezek maximumát jelöljük  $h$ -val. A korábbi rácsra vonatkozó feltételen felül elvárjuk, hogy az egyenesek távolsága tegyen eleget a

$$0 < \frac{h}{\kappa} \leq h_{1i}^\pm, h_{2j}^\pm \leq \kappa h \quad (\kappa \geq 1)$$

feltételnek, azt garantálja hogy ne legyen túl irreguláris a rácsunk.

A rácsponokat az második fejezethez hasonlóan  $w_h$ , illetve  $\bar{w}_h$  -val jelöljük. A  $\Gamma_{23}$  -beli pontokra bevezetjük a

$$\gamma_{23,h} := \Gamma_{23} \cap \bar{w}_h$$

jelölést. A  $\Gamma_1$  -beli rácsponatokkal külön nem foglalkozunk, ezeket a pontokat tudjuk eliminálni a 2.3.12. megjegyzés szerint.

A differenciálegyenlet megoldásához minden  $\mathbf{x}^0 \in w_h \cup \gamma_{23,h}$  rácsponat körül kijelölünk egy  $H(\mathbf{x}^0)$  “elemi tartományt” (véges térfogatot) a téglalapok oldalfelező pontjainak használatával.  $H(\mathbf{x}^0)$  peremét  $S(\mathbf{x}^0)$  -al jelöljük. Az  $S(\mathbf{x}^0)$  peremek  $\Omega$  -nak egy másodlagos felosztását adják.

Most nézzük meg mi a differenciálegyenlet diszkretizációjának alapja. Egy adott  $\mathbf{x}^0 \in w_h \cup \gamma_{23,h}$  ponthoz a (4.1) egyenletet  $H(\mathbf{x}^0)$  felett integráljuk és az első tagot a Gauss-Osztrogadszkij tétel használatával átalakítjuk a következőképpen:

$$\int_{H(\mathbf{x}^0)} \operatorname{div}(k\nabla u) d\mathbf{x} = \int_{S(\mathbf{x}^0)} k\nabla u \cdot \vec{n} ds = \int_{S(\mathbf{x}^0)} k \frac{\partial u}{\partial n} ds$$

ahol  $S(\mathbf{x}^0)$  pozitívan irányított és  $\vec{n}$  ennek külső normálvektora. Ezt adódik az úgynevezett mérlegegyenlet:

$$\int_{S(\mathbf{x}^0)} k \frac{\partial u}{\partial n} ds + \int_{H(\mathbf{x}^0)} f d\mathbf{x} = 0 \quad (4.2)$$

A (4.2) lokális tulajdonság globálisan is igaz, hiszen két szomszédos elemi tartományhoz tartozó egyenlet összeadásakor az első tagokban integrandusok egymás  $(-1)$ -szeresei a közös határon, így ezen a részen kiesnek, a második tagok összeadása pedig megfelel a tartományok unióján való integrálásnak.

Most nézzük meg, hogy (4.2) alapján a diszkretizálás hogyan történik. Erre több mód is van, mi egy lehetséges példát adunk rá, mégpedig az integrálok kvadratura formulával való helyettesítését.  $\mathbf{x}^0 \in w_h \cup \gamma_{23,h}$  pontban az  $f$ -hez tartozó integrált egyponos képlettel közelítjük:

$$\int_{H(\mathbf{x}^0)} f d\mathbf{x} \approx |H(\mathbf{x}^0)| f(\mathbf{x}^0)$$

A  $k \frac{\partial u}{\partial n}$ -hoz tartozó integrálhoz felhasználjuk, hogy  $S(\mathbf{x}^0)$  oldalai a koordinátatengelyekkel párhuzamosak, és hosszúságuk

$$h_{1i} := \frac{1}{2}(h_{1i}^+ + h_{1i}^-), \text{ illetve } h_{2j} := \frac{1}{2}(h_{2j}^+ + h_{2j}^-)$$

$S(\mathbf{x}^0)$  oldalait jelöljük  $s_1^\pm$ , illetve  $s_2^\pm$ -val. Az  $\mathbf{x}^0$  és  $(x_{1i}, x_{2,j\pm 1})$  rácspontok közötti szakaszt  $s_1^\pm$ , az  $\mathbf{x}^0$  és  $(x_{1,i\pm 1}, x_{2j})$  rácspontok közötti szakaszt  $s_2^\pm$  felezi. Emiatt a következő közelítéseket használjuk:

$$\int_{s_1^\pm} k \frac{\partial u}{\partial n} ds \approx \pm h_{1i} (k v_{x_2})_{i,j\pm 1/2}, \quad \int_{s_2^\pm} k \frac{\partial u}{\partial n} ds \approx \pm h_{2j} (k v_{x_1})_{i\pm 1/2,j}$$

ahol

$$k_{i\pm 1/2,j} := k\left(x_{1i} \pm \frac{h_{1i}^\pm}{2}, x_{2j}\right), \quad k_{i,j\pm 1/2} := k\left(x_{1i}, x_{2j} \pm \frac{h_{2j}^\pm}{2}\right)$$

$v$  pedig az  $u(\mathbf{x})$  approximációja a  $\mathbf{x} \in w_h$  rácspontokban, és a  $v$  diszkrét iránymenti deriváltjai

$$\pm v_{x_1, i \pm 1/2, j} := \frac{v(x_1 \pm h_1^\pm, x_2) - v(\mathbf{x})}{h_1^\pm} \quad (4.3)$$

$$\pm v_{x_2, i, j \pm 1/2} := \frac{v(x_1, x_2 \pm h_2^\pm) - v(\mathbf{x})}{h_2^\pm} \quad (4.4)$$

Azaz

$$(kv_{x_2})_{i, j \pm 1/2} = \pm k \left( x_{i1}, x_{2j} \pm \frac{h_{1i}^\pm}{2} \right) \frac{v(x_1, x_2 \pm h_2^\pm) - v(\mathbf{x})}{h_2^\pm}$$

és

$$(kv_{x_1})_{i \pm 1/2, j} = \pm k \left( x_{1i} \pm \frac{h_{1i}^\pm}{2}, x_{2j} \right) \frac{v(x_1 \pm h_1^\pm, x_2) - v(\mathbf{x})}{h_1^\pm}$$

Ezekre bevezetünk rövidítéseket, hogy ne bonyolódjanak el a jelölések:  $v_{x_1, \pm}$ ,  $v_{x_2, \pm}$ ,  $k_{i \pm, j}$ ,  $k_{i, j \pm}$ .

Tehát az ezen módon kapott approximációk a következőképpen néznek ki:

i, Ha  $\mathbf{x}^0 \in w_h$ , azaz a tartomány belsejében van, és a hozzá tartozó  $H(\mathbf{x}^0)$  tartomány egy téglalap:

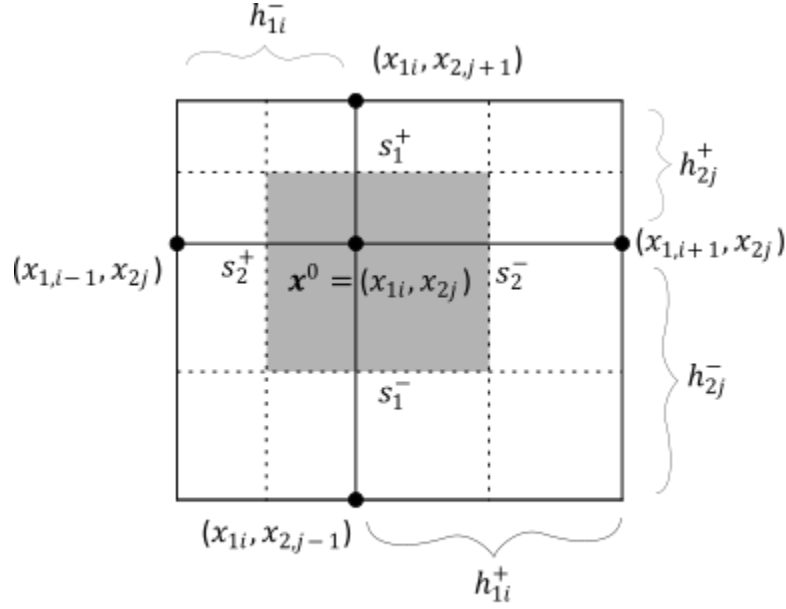
$$h_{1i} [(kv_{x_2})_{i, j+} - (kv_{x_2})_{i, j-}] + h_{2j} [(kv_{x_1})_{i+, j} - (kv_{x_1})_{i-, j}] + |H(\mathbf{x}^0)| f(\mathbf{x}^0) = 0 \quad (4.5)$$

Mivel  $H(\mathbf{x}^0)$  egy téglalap, ezért  $|H(\mathbf{x}^0)| = h_{1i} h_{2j}$ .  $|H(\mathbf{x}^0)|$ -vel való osztás és átrendezés után:

$$-\left( \frac{(kv_{x_2})_{i, j+} - (kv_{x_2})_{i, j-}}{h_{2j}} + \frac{(kv_{x_1})_{i+, j} - (kv_{x_1})_{i-, j}}{h_{1i}} \right) = f(\mathbf{x}^0)$$

Behelyettesítve azt kapjuk hogy

$$-\left( \frac{k_{i, j+} \frac{v_{i, j+1} - v_{i, j}}{h_2^+} - k_{i, j-} \frac{v_{i, j} - v_{i, j-1}}{h_2^-}}{h_{2j}} + \frac{k_{i+, j} \frac{v_{i+1, j} - v_{i, j}}{h_1^+} - k_{i-, j} \frac{v_{i, j} - v_{i-1, j}}{h_1^-}}{h_{1i}} \right) = f(\mathbf{x}^0)$$



3. ábra. Belül lévő pont környezete

ii,  $\mathbf{x}^0 \in \gamma_{23,h}$  és az  $\mathbf{x}^+, \mathbf{x}^0, \mathbf{x}^-$  pontokat összekötő töröttvonalnak része  $\Gamma_{23}$ , ahol  $\mathbf{x}^\pm := \left( x_{1i} \pm \frac{h_{1i}^\pm}{2}, x_{2j} \pm \frac{h_{2j}^\pm}{2} \right)$ , a másik kettő felhasznált rácspontok pedig a tartományon belüli:  $(x_{1,i+1}, x_{2j}), (x_{1i}, x_{2,j-1}) \in w^h$ . Ekkor a

$$k(\mathbf{x}) \frac{\partial u}{\partial n}(\mathbf{x}) + \sigma(\mathbf{x})u(\mathbf{x}) = g_{23}(\mathbf{x}) \quad \mathbf{x} \in \Gamma_{23}$$

peremfeltételt felhasználva az  $\mathbf{x}^0$ -hoz tartozó közelítés:

$$-h_{1i}(kv_{x_2})_{i,j-} + h_{2j}(kv_{x_1})_{i+,j} + (|s_\gamma^+| + |s_\gamma^-|)(g_{23} - \sigma v)(\mathbf{x}^0) + |H(\mathbf{x}^0)|f(\mathbf{x}^0) = \quad (4.6)$$

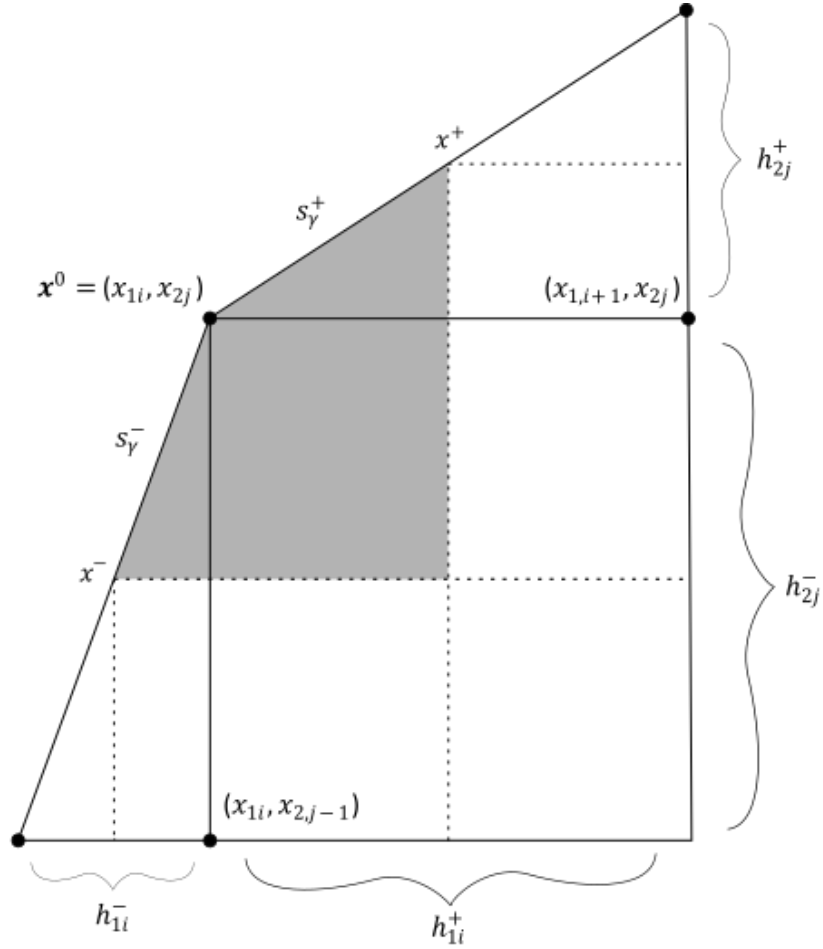
Ahol

$$|s_\gamma^\pm| := \frac{1}{2} \left( (h_{1i}^\pm)^2 + (h_{2j}^\pm)^2 \right)^{\frac{1}{2}}$$

illetve

$$|H(\mathbf{x}^0)| := \frac{h_{1i}^+ h_{2j}^+}{2} + \frac{h_{1i}^- h_{2j}^-}{2} + \frac{h_{1i}^+ h_{2j}^-}{2} = \frac{h_{1i}^+}{2} \left( \frac{h_{2j}^-}{2} + \frac{h_{2j}^+}{4} \right) + \frac{h_{1i}^- h_{2j}^-}{8}$$

Itt  $|s_\gamma^\pm|$ -át Pithagorasz tétellel számoltuk,  $|H(\mathbf{x}^0)|$  pedig két derékszögű háromszög és egy téglalap területeinek összege.



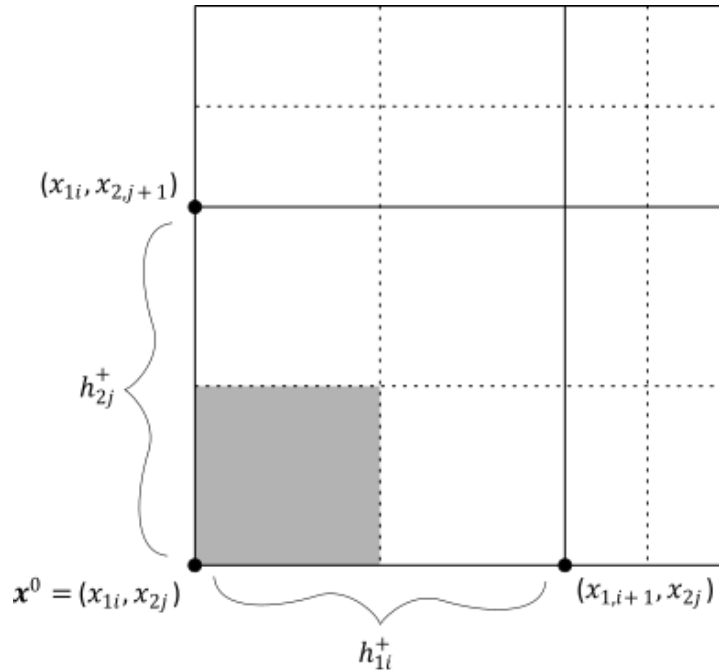
4. ábra. Peremen lévő pont környezete

iii,  $\mathbf{x}^0 \in \gamma_{23,h}$  és az  $\mathbf{x}^0$  lévő belső szög  $\frac{\pi}{2}$

$$\frac{h_{2j}^+}{2} (kv_{x_1})_{i+,j} + \frac{h_{1i}^+}{2} (kv_{x_2})_{x_{i,j+}} + \left( \frac{h_{2j}^+}{2} + \frac{h_{1i}^+}{2} \right) (g_{23} - \sigma v)(\mathbf{x}^0) + |H(\mathbf{x}^0)| f(\mathbf{x}^0) = 0 \quad (4.7)$$

Ahol

$$|H(\mathbf{x}^0)| := \frac{h_{2j}^+}{2} \frac{h_{1i}^+}{2}$$



5. ábra. Sarkon lévő pont környezete

**4.2.1. Megjegyzés:** Ezzel a felosztással és integrálközelítéssel a kapott (4.5) egyenletek  $|H(\mathbf{x}^0)|$ -val való osztás után azonosak az ugyanerre a rácsra felírt véges differencia módszerrel.

**4.2.2. Állítás:** A felírt egyenletekből keletkező mátrix M-mátrix, így a lineáris egyenletrendszernek létezik egyértelmű megoldása.

**4.2.3. Megjegyzés:** A boxmódszerhez természetesen nem muszáj a bemutatott téglalaprácsot használni a tartomány felosztásához. Más, például csupa háromszög trianguláció esetében is elérhető, hogy a kapott mátrix M-mátrix legyen.

## 4.2 Konvergencia

**4.3.1. Tétel:** Ha  $y_h$  a véges térfogat módszerrel kapott megoldásvektor, akkor az  $u$  pontos megoldás rácspontbeli értékeinek  $u_h$  vektorára

$$\|y_h - u_h\|_{C(w_h)} \leq M h^{1+j/2} \left( \ln \frac{1}{h} \right)^{1/2}$$

ahol  $j = 0, 1$  és  $M$  alkalmas konstansok továbbá a norma

$$\|u\|_{C(w_h)} := \max_{\mathbf{x} \in w_h} |u(\mathbf{x})|$$

# 5 Összefoglalás

Szakdolgozatomban parciális differenciálegyenletek numerikus megoldására néztünk meg három alapvető módszert, melyeken egy-egy példán keresztül részletesen végigmentünk. Ezek elég speciális alakúak voltak, természetesen más típusú egyenletekre nem pontosan ezek sémák, de az alapgondolatok és alaptechnikák megértéséhez ennyi is elegendő volt. Mindhárom bemutatott módszer végeredményben csak egy eljárást adott, amivel differenciálegyenletet egy lineáris egyenletrendszerre vezettük vissza. Azonban láthattuk, hogy ez a visszavezetés milyen változatos módokon történhet, a legegyszerűbb véges differencia módszertől, a véges elem módszerig, mely absztrakt terekre vonatkozó eredményekre épül. Ez is remekül mutatja, hogy egy adott problémát milyen sokféleképpen lehet megközelíteni.

## Irodalomjegyzék

- [1] STOYAN GISBERT, TAKÓ GALINA, *Numerikus Módszerek III*, Typotex Kiadó, 2008, 978-963-7546-77-8
- [2] KARÁTSON JÁNOS, HORVÁTH RÓBERT, IZSÁK FERENC, *Parciális differenciálegyenletek numerikus módszerei számítógépes alkalmazásokkal*, 2013, egyetemi jegyzet
- [3] BESENYEI ÁDÁM, Komornik Vilmos, Simon László, *Parciális differenciálegyenletek*, Typotex Kiadó, 2013, 978 963 279 259 0
- [4] STOYAN GISBERT, TAKÓ GALINA, *Numerikus Módszerek II*, Typotex Kiadó, 2012, 978-963-7546-53-2
- [5] KARÁTSON JÁNOS, *Numerikus funkcionálanalízis*, Typotex Kiadó, 2012, 978-963-279-239-2