

Eötvös Loránd Tudományegyetem
Meteorológiai Tanszék

Egydimenziós véges elemes módszerek alkalmazása advekciós feladatokra



Készítette: Krüzselyi Ilona

Témavezető: Dr. Havasi Ágnes
(ELTE Meteorológiai Tanszék)

Budapest, 2009

Tartalomjegyzék

1. Bevezetés	3
2. Az advekción feladat	5
2.1. <i>A szennyezőanyag-terjedési egyenlet</i>	5
2.2. <i>Operátorszeletelés</i>	6
2.3. <i>Az advekción</i>	8
2.4. <i>Numerikus módszerek</i>	9
3. A véges elemes módszerről	10
4. A Galjorkin-módszer	12
4.1. <i>A hagyományos véges elemes módszer</i>	14
4.2. <i>A Petrov–Galjorkin-módszer</i>	18
4.3. <i>A legkisebb négyzetek módszere</i>	21
4.4. <i>Periodikus peremfeltétel</i>	24
5. A tesztelés eredményei	27
6. Fourier-analízis	35
6.1. <i>A hagyományos véges elemes módszer</i>	37
6.2. <i>A Petrov–Galjorkin-módszer</i>	39
6.3. <i>A legkisebb négyzetek módszere</i>	40
6.4. <i>Eredmények</i>	41
7. A kétdimenziós advekción feladat	47
8. Összefoglalás	51
Függelék	53
Irodalomjegyzék	57

1. Bevezetés

Napjainkban igen meghatározó probléma a légszennyezés, illetve annak következményei, mint a savas esők, a klímaváltozás, vagy a bioszférában okozott kár. Adott szennyezőanyagok légköri tartózkodási időjüktől függően akár több ezer kilométerre is elsodródhatnak a kibocsátás helyétől, így regionális, sőt globális skálán is problémákat okoznak.

Ilyen regionális probléma, hogy a nitrogén-oxidok, a kén-dioxid és a szulfátok száraz, illetve nedves ülepedésével a talaj és a felszíni vizek savasodnak, ami a növények és vízi élőlények károsodásához, illetve kipusztulásához vezet. Egy másik regionális példa az ózonszennyezés, ami igen negatív hatással van az emberi egészségre és a vegetációra. A troposzférikus ózon másodlagos szennyező, azaz különböző elsődleges szennyezők, nitrogén-oxidok, szén-monoxid, metán és illékony szerves anyagok (VOC – Volatile Organic Compounds) fotokémiai reakciói során keletkezik.

Az egy hétnél hosszabb tartózkodási idővel rendelkező anyagok globális skálán is éreztetik hatásukat. Az ötvenezer éves átlagos tartózkodási idővel rendelkező CFC-gázok a fő okozói a sztratoszférikus ózommennyiség csökkenésének, és így annak is, hogy a káros UV-B sugarak elérik a földfelszínt. A megnövekedett koncentrációjú üvegházhatású gázok pedig (vízgőz, CO_2 , CH_4 , N_2O , O_3 , CFC-gázok) hozzájárulnak a globális klímaváltozáshoz. Mindezekről a problémákról bővebben olvashatunk Stanners és Bourdeau munkájában (1995).

Adott légszennyező anyag legmagasabb koncentrációját vagy ülepedését, aminek még nincs káros hatása a bioszférára egy elégségesen hosszú időszak alatt, az adott anyag elfogadható vagy kritikus koncentrációjának vagy ülepedésének nevezzük. Elkerülendő a kritikus szint túllépését, sok anyag esetében csökkenteni kell a kibocsátást, de nem jobban, mint szükséges, hiszen az emisszió-csökkentés drága, gazdasági következményei is vannak. Mindemellett a különböző szennyezőanyagok közti kölcsönhatásokat is figyelembe kell venni.

Ahhoz, hogy meg tudjuk határozni, hogy egy lehetséges emissziós forrás mennyire terhelheti meg a környezetet, illetve azt, hogy a kibocsátás visszaszorítása mennyire befolyásolná annak szennyezettségét, matematikai modellek kidolgozására és azok elemzésére van szükség. Erre a célra az ún. szennyezőanyag-terjedési modelleket használják. Ezek a modellek a tömegmegmaradás elvén alapulnak, amit egy parciális

differenciálegyenlet-rendszerrel fejezhetünk ki, az ún. szennyezőanyag-terjedési egyenletekkel. A nagytávolságú légszennyezés öt különböző részfolyamat együttes eredménye, úgymint

1. az advekció,
2. a diffúzió,
3. az ülepedés,
4. az emisszió
5. és a kémiai reakciók.

A terjedési egyenletek ezeket a jelenségeket írják le. Analitikusan ezeket az egyenleteket csak irreális egyszerűsítések mellett lehet megoldani, ezért numerikus módszerek alkalmazására van szükség. Az első lépés lehet, hogy az eredeti rendszert felbontjuk a fenti fizikai folyamatok szerint egyszerűbben kezelhető alrendszerekre az operátorszeletelés segítségével. A szakdolgozatban egy ilyen részfeladat, az advekciós feladat numerikus megoldásával foglalkozunk.

Az advekciós feladat numerikus megoldásánál a térbeli diszkretizációra véges elemes módszereket használunk. A hagyományos véges elemes (Galjorkin-) módszer hiperbolikus feladatok esetén általában hamis oszcillációkat ad az eredményben (Hunsdorfer and Verwer, 2003), ezért más eljárásokat is tanulmányozunk, hogy megvizsgáljuk, csökkenthető-e az oszcilláció mértéke, ill. kaphatunk-e pontosabb megoldást. A dolgozatban a hagyományos véges elemes módszer, egy Petrov–Galjorkin-féle módszer és a legkisebb négyzetek módszerének használhatóságát, ill. a Fourier-analízis segítségével azok tulajdonságait vizsgáljuk.

A 2. fejezet vázolja, hogyan lehet a terjedési egyenletektől eljutni az advekciós feladat megoldásáig. A 3. fejezetben a véges elemes módszerek használatáról, használhatóságáról lesz szó. A 4. fejezet tartalmazza a Galjorkin-módszerek ismertetését, míg az 5. fejezet az ezekkel végzett kísérletek eredményeit. A módszerek Fourier-analízise a 6. fejezetben található. A 7. fejezetben a kétdimenziós advekciós feladat megoldásával foglalkozunk, majd a főbb eredmények összefoglalása zárja a dolgozatot.

2. Az advekciós feladat

2.1. A szennyezőanyag-terjedési egyenlet

Tegyük fel, hogy m különböző kémiai anyag koncentrációváltozását vizsgáljuk. Jelölje $\underline{c} \in R^m$ vektor az m anyag koncentrációját, ami általában térben és időben változik, azaz $\underline{c} = \underline{c}(\underline{x}, t)$. Ha $\underline{u} = \underline{u}(\underline{x}, t)$ jelöli a háromdimenziós sebességvektort, és c egyetlen anyag koncentrációját az m közül, akkor c változását a

$$\frac{\partial c}{\partial t} + \nabla(\underline{u}c) = \nabla(\underline{K}\nabla c) + R(\underline{x}, \underline{c}) + E + \sigma c$$

parciális differenciálegyenlettel és a megfelelő kezdeti és peremfeltételekkel írhatjuk le (Zlatev, 1995). A bal oldal második tagja írja le az advekciót. A jobb oldal első tagja fejezi ki a turbulens diffúziót, ahol $\underline{K} = \underline{K}(\underline{x}, t)$ a diffúziós együttható-mátrix. $R(\underline{x}, \underline{c})$ reprezentálja az adott anyaggal végmenő kémiai reakciókat, ez a tag köti össze az m darab egyenletet. E írja le az emissziót, a természetes és antropogén szennyezőanyag-kibocsátást. A σc tag az adott anyag nedves és száraz ülepedését fejezi ki. A fenti rendszert nem lehet megoldani analitikusan, a kivitelezhető és megbízható numerikus megoldásra pedig a legjobb, és sokszor egyetlen módszere az operátorszeletelés elsődleges alkalmazása.

2.2. Operátorszeletelés

Az operátorszeletelés lényege, hogy az eredeti modellt több, egyszerűbb alrendszerrel helyettesítjük, amelyek megoldására akár különböző numerikus eljárásokat is használhatunk. A következőkben az operátorszeletelést Csomós et al. (2005) és Havasi et al. (2001) munkája alapján ismertetjük.

Legyen S egy normált tér, és tekintsük a

$$\begin{aligned} \frac{dw(t)}{dt} &= Aw(t) \quad t \in (0, T] \\ w(0) &= w_0 \end{aligned}$$

kezdetiérték-feladatot, ahol $w: [0, T] \rightarrow S$ ismeretlen függvény, és $A: S \rightarrow S$ lineáris operátor. Tegyük fel, hogy az A operátor felbontható két egyszerű operátor összegére, nevezetesen az A_1 és A_2 operátoróra. A legegyszerűbb operátorszeletelési módszer az ún. szekvenciális szeletelés, amelynek lényege a következő. A feladat időintervallumát n darab részintervallumra bontjuk. (Az egyszerűség kedvéért tegyük fel, hogy ezek egyenlő hosszúságúak.) A részintervallumok $\tau = T/n$ hosszúságát szeletelési időlépcsőnek nevezzük. Első lépésben az első részintervallumon megoldjuk az A_1 operátorú részfeladatot a megadott kezdeti feltételből kiindulva. Ezután a kiszámított megoldást kezdeti feltételként alkalmazva megoldjuk az A_2 operátorú részfeladatot. Ezt a folyamatot ismételjük az összes szeletelési időlépcsőben úgy, hogy a kezdeti feltétel mindig az előbb megoldott részfeladat megoldása lesz. Képlettel tehát a megoldandó feladatsorozat a következőképpen adható meg:

$$\frac{dw_k^{(1)}(t)}{dt} = A_1 w_k^{(1)}(t) \quad t \in ((k-1)\tau, k\tau]$$

$$w_k^{(1)}((k-1)\tau) = w_{k-1}^{(2)}((k-1)\tau)$$

és

$$(1) \quad \frac{dw_k^{(2)}(t)}{dt} = A_2 w_k^{(2)}(t) \quad t \in ((k-1)\tau, k\tau]$$

$$(2) \quad w_k^{(2)}((k-1)\tau) = w_k^{(1)}(k\tau),$$

ahol $k = 1, 2, \dots, n$ és $w_0^{(2)}(0) = w_0$. Ez az eljárás természetesen több részoperátorra is alkalmazható.

Az operátorszeletelés alkalmazása rendszerint hibát okoz a megoldásban. A megoldást vizsgáljuk csak az első szeletelési lépésben, azaz $k=1$ -re, azzal a feltételezéssel, hogy az eredeti feladat és az összes részfeladat is megoldható pontosan! Ha az (1)(2) feladat megoldását $\hat{w}(\tau)$ -val jelöljük, akkor az

$$Err_{sp}(\tau) := \hat{w}(\tau) - w(\tau)$$

kifejezést lokális szeletelési hibának nevezzük. Jelölje

$$[A_1, A_2] := A_1 \circ A_2 - A_2 \circ A_1$$

A_1 és A_2 kommutátorát. Belátható, hogy a szekvenciális szeletelés lokális szeletelési hibája pontosan akkor tűnik el minden kezdeti feltételre, ha az

$$[A_1, A_2] = 0$$

feltétel teljesül.

A szekvenciális operátorszeletelésen kívül egyéb operátorszeletelési eljárásokat is alkalmaznak a légszennyezési modellekben, például a szimmetrikus, a súlyozott szekvenciális és a súlyozott szimmetrikus operátorszeletelési módszereket (Strang, 1963, 1968; Csomós et al., 2005). Az operátorszeletelés segítségével tehát a terjedési egyenletet részfeladatokra bonthatjuk. Ezek egyike az advekción feladat, amellyel foglalkozni kívánunk.

2.3. Az advekción

Az advekción a szélmezővel való szállítódás folyamatát jelenti. Legyen $c(x,y,z,t)$ egy adott anyag koncentrációja az $(x,y,z) \in D \subset R^3$ pontban és a $t \in [0,T]$ időpillanatban. Jelölje a szélesség három komponensét $u(x,y,z,t)$, $v(x,y,z,t)$ és $w(x,y,z,t)$. Tegyük fel, hogy a szélességet ismerjük az egész D tartományon, illetve az egész $[0,T]$ időintervallumon. Ekkor az adott anyag szélmezővel való terjedését a következő parciális differenciálegyenlettel írhatjuk le:

$$\frac{\partial c}{\partial t} = -\frac{\partial(uc)}{\partial x} - \frac{\partial(vc)}{\partial y} - \frac{\partial(wc)}{\partial z} \quad \forall (x,y,z) \in D \subset R^3, \forall t \in [0,T]$$

Egyszerűsíthetjük az egyenletet, ha feltételezzük, hogy a sebességmező divergenciamentes, azaz, ha teljesül, hogy

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad \forall (x,y,z) \in D \subset R^3, \forall t \in [0,T].$$

Így a

$$\frac{\partial c}{\partial t} = -u \frac{\partial c}{\partial x} - v \frac{\partial c}{\partial y} - w \frac{\partial c}{\partial z} \quad \forall (x,y,z) \in D \subset R^3, \forall t \in [0,T]$$

advekcións egyenletet kapjuk. Tovább egyszerűsödik az egyenlet, ha az áramlási mező konstans. Ekkor az advekcións egyenlet alakja az alábbi:

$$(3) \quad \frac{\partial c}{\partial t} = -u_0 \frac{\partial c}{\partial x} - v_0 \frac{\partial c}{\partial y} - w_0 \frac{\partial c}{\partial z} \quad \forall (x,y,z) \in D \subset R^3, \forall t \in [0,T]$$

A fenti egyenletekkel analóg módon fel lehet írni az egy- és kétdimenziós advekcións egyenleteket is. Természetesen az egyenletekhez tartoznak megfelelő kezdeti és peremfeltételek is. A (3) egyszerűsített egyenlet gyakran előfordul a szennyezőanyag-terjedési modellek advekcións részeként. A megoldására használt numerikus módszerek alkalmazhatóságának ellenőrzésére, vagy elméleti vizsgálatokban alkalmazzák (Kreiss and Olinger, 1972).

2.4. Numerikus módszerek

Az advekcios egyenletek általában nem oldhatók meg analitikusan, de numerikus megoldásukra több eljárás is kínálkozik, és legtöbbjük besorolható az alábbi négy módszer közé:

- karakterisztikák módszere
- véges különbséges módszer
- véges elemes módszer
- spektrális módszer

A legelső módszer, amelyet az advekcios feladat megoldására használtak, a karakterisztikák módszere. Ez a legkényelmesebb módszer, ha u , v és w kielégít bizonyos feltételeket. Azonban a terjedési modellekben a sebesség általános függvénye (x,y,z,t) -nek, ezért használata kevésbé népszerű, ennek ellenére sok helyen alkalmazzák. Bővebben lásd Richtmyer and Morton (1967) és Smith (1978).

A véges különbséges módszereket a XX. század elejétől kezdve használják a differenciálegyenletek megoldására, a '60-as, '70-es évektől pedig egyre szélesebb körben alkalmazzák terjedési modellekben is. Általában a legalább harmad- és negyedrendű sémák az elterjedtek. Bővebben lásd Chock (1985, 1991) és Long and Pepper (1981).

A spektrális és véges elemes módszer az ún. Galjorkin-módszer két változata. A spektrális módszereknél ortogonális függvényrendszer segítségével írjuk fel az ismeretlen változókat. Ennek is két változata van: a spektrális módszernél Fourier-együtthatókat, az ún. pszeudospektrális módszernél Csebisev-polinomokat használnak. Bővebben lásd Kreiss and Olinger (1972) és Orszag (1971, 1972).

3. A véges elemes módszerről

A véges elemes módszer (VEM) egy olyan diszkretizációs módszer, amely során az ismeretlen változókat csak lokálisan nem nullaértékű polinomok lineáris kombinációjaként közelítjük. A véges elemes séma számos előnyös tulajdonsággal rendelkezik. A meteorológiai egyenletekre való alkalmazása nagyon pontos fázisterjedést ad, és igen jól kezeli a nemlinearitásokat (Haltiner and Williams, 1980). A módszer használható változó felbontásnál (Côté et al., 1998) vagy szabálytalan határoknál (Sklarz et al., 1979). A VEM legnagyobb hátránya az, hogy minden időlépcsőben, minden változóra egy nagy mátrixot kell invertálni (Haltiner and Williams, 1980). A légszennyezési modellekben főként egyszerűbb, például elsőrendű lineáris egydimenziós véges elemes sémákat használnak (Chock 1985, 1991; Pepper et al., 1979), de próbálkoznak fejlettebb módszerekkel is, amelyeket lokálisan nagyobb felbontásnál használhatnak (Neta and Williams, 1986; Neta, 1992). A diffúziós-advekción-reakción egyenletre Codina számos véges elemes módszert összegyűjtött és összehasonlított cikkében (Codina, 1998).

A véges elemes módszer használata egyre nagyobb tért hódít a meteorológiai modellezés más területein is. Côté és Staniforth (1990) két globális szemi-lagrange-i sekélyvízmodellt hasonlított össze, amelyek között csak a térbeli diszkretizációban volt különbség. Az egyik modellben spektrális, a másikban véges elemes módszert alkalmaztak. Arra jutottak, hogy a két modell majdnem azonos minőségű előrejelzést ad rögzített időlépcsőre és összehasonlítható térbeli felbontásra, de a véges elemes módszert használó modell kicsit jobbnak bizonyult. A VEM másik előnye, hogy nagy felbontásnál költséghatékonyabb, mint a spektrális módszer.

A Kanadai Meteorológiai Központban a rövidtávú regionális előrejelzésekre évekig a Regionális Véges Elem (RFE – Regional Finite Element) (Mailhot et al. 1997) modellt használták, mígnem 1997 februárjában felváltották a GEM (Global Environmental Multiscale) modellel (Côté et al. 1998), ami egy globális, változó rácsfelbontású, multiskálájú modell, és akárcsak az RFE modell, háromdimenziós véges elemes módszert alkalmaz a térbeli diszkretizációra. A GEM modell három különböző skálájú konfigurációval rendelkezik. Az ekvidisztáns felbontású, globális skálájú konfigurációt az időjárás-előrejelzésen kívül a klímaváltozás és a nagytávolságú szennyezőanyag-terjedés modellezésére is használják. A változó felbontású szinoptikus skálájú konfigurációval rövidtávú előrejelzéseket készítenek, és a regionális klímát és légszennyezettséget

modellezik. Végül a változó felbontású mezo- β , ill. mezo- γ skálájú konfigurációt még pontosabb előrejelzésekre és szimulációkra alkalmazzák.

Európában is van példa a VEM operatív használatára. Untch és Hortal (2003) megvizsgálta, hogy mennyire hatékony az egydimenziós véges elemes módszer az operatíván működő véges különbséggel szemben a szemi-lagrange-i sémát alkalmazó ECMWF (European Centre for Medium Range Weather Forecasts) előrejelző modellben a vertikális diszkretizációnál. A VEM lineáris bázisfüggvényekkel negyedrendben, harmadfokúakkal nyolcadrendben pontos, míg a véges különbséges módszer csak másodrendű. A véges elemes séma csökkentette a vertikális hőmérsékleti profil számítási módusainak amplitúdóját és a modellben jelentkező globális ózonvesztést a 60 napos előrejelzésekben. Négydimenziós variációs adatasszimilációs (4D-Var) kísérleteket is végeztek, és eredményül azt kapták, hogy a VEM csökkentette a vertikális zajokat és a negatív hőmérsékleti anomáliákat a sztratoszférában. A pozitív eredmények következményeként az ECMWF 2002 januárjától a determinisztikus és ensemble előrejelzésekben, ill. a variációs adatasszimilációban operatíván használja a vertikális diszkretizációra a véges elemes módszert harmadfokú bázisfüggvényekkel, sőt még a 40 éves re-analízisnél (ERA40) is alkalmazták.

4. A Galjorkin-módszer

A Galjorkin-módszert és a három véges elemes sémát, amelyekkel a dolgozatban foglalkozom, a konkrét advekción feladat segítségével ismertetem. A módszerek elméleti vizsgálatát és tesztelését is ezen a feladaton végeztem.

A feladat egy térben egydimenziós, homogén, hiperbolikus parciális differenciálegyenlet inhomogén kezdeti és peremfeltétellel:

$$(4) \quad \frac{\partial c}{\partial t} + u_0 \frac{\partial c}{\partial x} = 0, \quad u_0 > 0, \text{ konstans}, \quad t \in [0, T], \quad x \in \Omega = [0; \pi]$$

$$(5) \quad c(0, t) = g(t), \quad t \in [0, T]$$

$$(6) \quad c(x, 0) = c_0(x), \quad x \in \Omega = [0; \pi]$$

Az u_0 konstans sebességet advekción sebességnek nevezzük. Az egydimenziós modellek általában nem reálisak a szennyezőanyagok terjedésének vizsgálatánál, bár vannak speciális esetek, ahol alkalmazhatók, például a rövidskálájú szállítódásnál, ahol a szélirány közelítőleg állandónak tekinthető. Egyszerűségük azonban alkalmassá teszik őket új numerikus módszerek tesztelésére.

Jelölje Ω_T az $\Omega \times [0, T]$ halmazt. A (4)-(6) feladat klasszikus megoldása az x és t szerint egyszer folytonosan differenciálható függvények körébe tartozik, azaz $c \in C^{1,1}(\Omega_T)$, ahol c kielégíti a kezdeti és peremfeltételeket. Ha egy feladat nem oldható meg analitikusan, akkor közelítő megoldásra van szükség. A $C^{1,1}(\Omega_T)$ tér függvényeit nehéz közelíteni, ezért $C^{1,1}(\Omega_T)$ -nél bővebb függvényosztályban keresünk egy ún. gyenge megoldást.

A gyenge megoldás értelmezéséhez először vegyük észre, hogy ha c klasszikus megoldása a (4)-(6) feladatnak, akkor c kielégíti az

$$(7) \quad \int_0^\pi \left(\eta(x) \frac{\partial c}{\partial t}(x, t) + u_0 \eta(x) \frac{\partial c}{\partial x}(x, t) \right) dx = 0$$

egyenletet minden $t \in [0, T]$ és minden $\eta \in L^2(\Omega)$ esetén. Itt $L^2(\Omega)$ az Ω halmazon négyzetesen integrálható függvények Hilbert-terét jelöli a $\langle c, \eta \rangle = \int_\Omega c \eta$ skalárszorozattal.

Vegyük észre, hogy (7) már nem csak $C^{1,1}(\Omega_T)$ -beli függvényekre értelmes, hanem olyan c függvényre is, amely

1. t szerint differenciálható, és amelyre minden rögzített $t \in [0, T]$ esetén az $x \rightarrow \frac{\partial c}{\partial t}(x, t)$ függvény $L^2(\Omega)$ -beli;
2. amelyre minden rögzített $t \in [0, T]$ esetén az $x \rightarrow c(x, t)$ függvény $H^1(\Omega)$ -beli, azaz önmaga és első deriváltja is $L^2(\Omega)$ -beli.

Írjuk fel a (4) egyenletet a

$$(8) \quad D_t c + A c = 0$$

alakban, ahol $D_t = \frac{\partial}{\partial t}$ és $A = u_0 \frac{\partial}{\partial x}$, és vezessük be a $V := \{ c(x, t) : \exists \frac{\partial c}{\partial t}, \forall t : \frac{\partial c}{\partial t} \in L^2(\Omega);$

$\forall t : x \rightarrow c(x, t) \in H^1(\Omega) \}$ jelölést. Ekkor azt mondjuk, hogy $c \in V$ gyenge megoldása a (8)-(5)-(6) feladatnak, ha

$$(9) \quad \langle D_t c, \eta \rangle + \langle A c, \eta \rangle = 0 \quad \forall \eta \in L^2(\Omega)$$

teljesül minden rögzített $t \in [0, T]$ esetén, és c kielégíti az (5) peremfeltételt és a (6) kezdeti feltételt. A (9) egyenletet (8) gyenge alakjának, az $\eta(x)$ függvényt pedig próbafüggvénynek nevezzük.

A gyenge megoldás közelítését egy $V_h \subset V$ véges dimenziós altérben keressük, továbbá a (9) egyenlőség teljesülését nem minden $\eta \in L^2(\Omega)$ függvényre követeljük meg, hanem csak egy $W_h \subset L^2(\Omega)$ altér elemeire. Azaz keresünk egy olyan c_h közelítést a V_h térből, amelyre

$$(10) \quad \langle D_t c_h, \eta_h \rangle + \langle A c_h, \eta_h \rangle = 0 \quad \forall \eta_h \in W_h \subset L^2(\Omega)$$

teljesül minden $t \in [0, T]$ esetén, és amely c_h kielégíti az (5)-(6) feltételeket (Kan et al., 2006).

A V_h és W_h terek bázisfüggvény-rendszereinek megválasztásától függően többféle véges elemes Galjorkin-módszer létezik.

4.1. A hagyományos véges elemes módszer

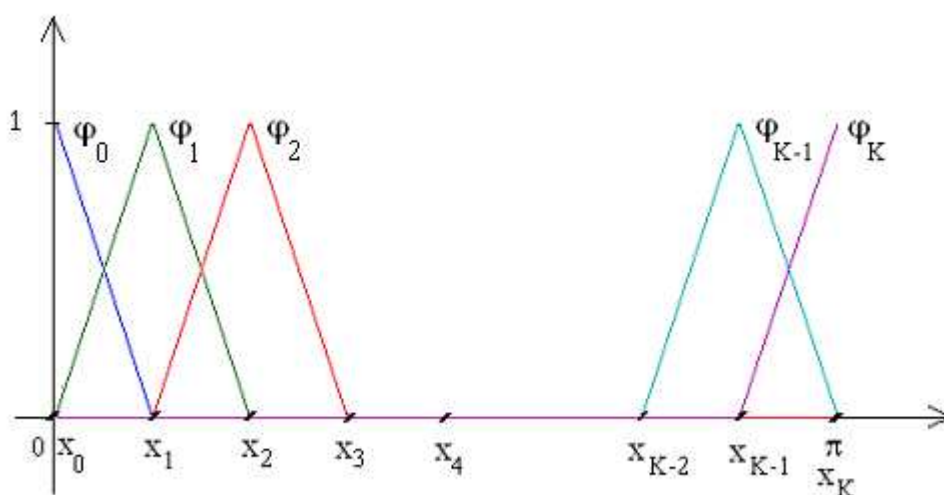
A hagyományos VEM esetében a c_h közelítő megoldás és az η_h próbafüggvény ugyanabból a térből való ($\forall c_h, \eta_h \in V_h$), és megegyeznek bázisfüggvényeik is.

A véges elemes módszerek lényege, hogy az Ω tartományt résztartományokra bontjuk, és a megoldást minden időpontban olyan folytonos függvénnyel közelítjük, amely minden ilyen résztartományon általában azonos fokszámú, de különböző alakú alacsonyrendű polinomok összegeként áll elő. Azaz V_h olyan folytonos függvényekből áll, amelyeket az egyes résztartományokon alacsonyrendű polinomok lineáris kombinációjával adunk meg. Ezen tér bázisfüggvényei egy kisszámú résztartománytól eltekintve mindenhol nulla értéket vesznek fel, véges számú pont kivételével deriválhatók, és maguk és deriváltjaik is Riemann-integrálhatók.

A mi esetünkben a bázisfüggvények az 1. ábrán is látható szakaszonként lineáris kalapfüggvények:

$$\varphi_i(x_j) = \delta_{ij},$$

ahol δ_{ij} a Kronecker-delta, $x_j = \frac{\pi \cdot j}{K}$ ($j=0, \dots, K$), K a résztartományok száma (Kan et al., 2006).



1. ábra – A hagyományos véges elemes módszer bázisfüggvényei

Tekintsük a (10) egyenletet kifejtve:

$$(11) \quad \int_0^\pi \left(\eta_h \frac{\partial c_h}{\partial t} + u_0 \eta_h \frac{\partial c_h}{\partial x} \right) dx = 0 \quad \forall \eta_h \in V_h, \forall t \in [0, T]$$

ahol c_h -t a következő alakban keressük:

$$(12) \quad c_h(t, x) = \sum_{i=0}^K \alpha_i(t) \varphi_i(x),$$

és az $\alpha_i(t)$ egyelőre ismeretlen függvényei t -nek. Mivel minden η_h felírható a φ_j bázisfüggvények lineáris kombinációjaként, (11) egyenlőséget elég a η_h -k helyett minden bázisfüggvényre igazolni. Ezt figyelembe véve, és (12)-t behelyettesítve (11)-be, a

$$(13) \quad \sum_{i=0}^K \left(\alpha_i'(t) \int_0^\pi \varphi_i \varphi_j dx + u_0 \alpha_i(t) \int_0^\pi \varphi_i' \varphi_j dx \right) = 0 \quad j=0, \dots, K$$

egyenletrendszeret kapjuk minden t időpontra.

Az $\alpha(t)$ függvényt a θ -módszerrel (Hunsdorfer and Verwer, 2003) közelítjük. Ez az egyik legelterjedtebb időintegrálási módszer, amelynek működését a

$$w'(t) = f(t, w)$$

általános időfüggő közönséges differenciálegyenleten mutatjuk be. Jelölje N az időlépések számát, és Δt az időlépcsők hosszát. Ekkor a

$$w_{n+1} = w_n + (1 - \theta) \Delta t \cdot f(t_n, w_n) + \theta \Delta t \cdot f(t_{n+1}, w_{n+1}) \quad n = 1, \dots, N; \quad N = \frac{T}{\Delta t}$$

sémát alkalmazzuk w közelítő kiszámítására, ahol $\theta \in [0, 1]$. Ha $\theta > 0$, akkor a séma implicit. A θ értékétől függően a θ -módszernek létezik néhány speciális esete:

- $\theta = 0$ explicit Euler-séma
- $\theta = 0,5$ trapézformula
- $\theta = 1$ implicit Euler-séma

A (13) összefüggésben az időbeli diszkretizációra a fenti θ -módszert alkalmazva a

$$\sum_{i=0}^K \left((\alpha_i^n - \alpha_i^{n-1}) \int_0^\pi \varphi_i \varphi_j dx + u_0 \Delta t \left[(1 - \theta) \alpha_i^n + \theta \alpha_i^{n+1} \right] \int_0^\pi \varphi_i' \varphi_j dx \right) = 0$$

$j=0, \dots, K; \quad n=1, \dots, N$

egyenletrendszert kapjuk, ami átrendezve a

$$(14) \quad \sum_{i=0}^K \left(\int_0^{\pi} \varphi_i \varphi_j dx + u_0 \Delta t \theta \int_0^{\pi} \varphi_i' \varphi_j dx \right) \alpha_i^n = \sum_{i=0}^K \left(\int_0^{\pi} \varphi_i \varphi_j dx - u_0 \Delta t (1-\theta) \int_0^{\pi} \varphi_i' \varphi_j dx \right) \alpha_i^{n-1}$$

$j=0, \dots, K; \quad n=1, \dots, N$

alakba írható. Felhasználva a peremfeltételt, azaz, hogy $\alpha_0^n = g(t_n)$ ($n = 1, \dots, N$), jutunk el (15)-höz.

$$(15) \quad \sum_{i=1}^K \left(\int_0^{\pi} \varphi_i \varphi_j dx + u_0 \Delta t \theta \int_0^{\pi} \varphi_i' \varphi_j dx \right) \alpha_i^n = \sum_{i=1}^K \left(\int_0^{\pi} \varphi_i \varphi_j dx - u_0 \Delta t (1-\theta) \int_0^{\pi} \varphi_i' \varphi_j dx \right) \alpha_i^{n-1}$$

$$- g(t_n) \left(\int_0^{\pi} \varphi_0 \varphi_j dx + u_0 \Delta t \theta \int_0^{\pi} \varphi_0' \varphi_j dx \right) + g(t_{n-1}) \left(\int_0^{\pi} \varphi_0 \varphi_j dx - u_0 \Delta t (1-\theta) \int_0^{\pi} \varphi_0' \varphi_j dx \right)$$

$j=1, \dots, K; \quad n=1, \dots, N$

Így végül minden n időrétegre egy K egyenletről álló lineáris egyenletrendszerhez jutunk, amelyben az ismeretlenek az α_i^n ($i=1, \dots, K$) együtthatók. Az α_i^{n-1} értékek ismertek az előző időrétegen kapott megoldásból, illetve $n=1$ esetén a kezdeti feltételből.

Az integrálokat kiszámolva kapjuk meg a végső alakját a lineáris egyenletrendszerünknek:

$$(16) \quad (\underline{S}_1 + u_0 \Delta t \theta \underline{S}_2) \underline{\alpha}^n = (\underline{S}_1 - u_0 \Delta t (1-\theta) \underline{S}_2) \underline{\alpha}^{n-1} - g(t_n) \underline{f}_1 + g(t_{n-1}) \underline{f}_2 \quad n=1, \dots, N,$$

ahol $\underline{\alpha}^n$ vektor tartalmazza az α_i^n együtthatókat, valamint

$$\underline{S}_1 = \frac{h}{6} \begin{bmatrix} 4 & 1 & & & & \\ 1 & 4 & 1 & & & 0 \\ & 1 & \ddots & & & \\ & & \ddots & \ddots & 1 & \\ & & & 1 & 4 & 1 \\ & 0 & & & 1 & 2 \end{bmatrix}, \quad \underline{S}_2 = \frac{1}{2} \begin{bmatrix} 0 & 1 & & & & \\ -1 & 0 & 1 & & & 0 \\ & -1 & \ddots & & & \\ & & \ddots & \ddots & -1 & \\ & & & -1 & 0 & 1 \\ & & & & -1 & 1 \end{bmatrix},$$

$$\underline{f}_1 = \begin{bmatrix} \frac{h}{6} - \frac{u_0 \Delta t \theta}{2} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{és} \quad \underline{f}_2 = \begin{bmatrix} \frac{h}{6} + \frac{u_0 \Delta t (1 - \theta)}{2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

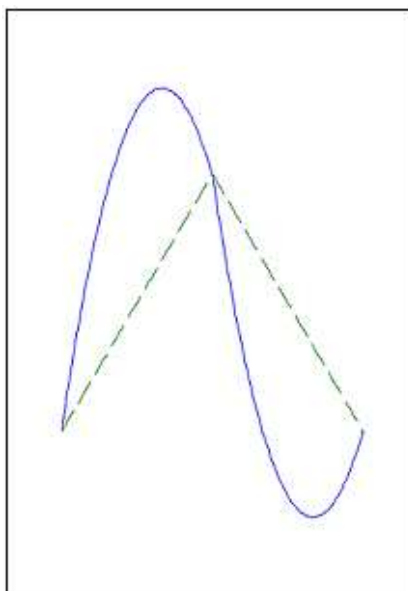
A képletekben $h = \frac{\pi}{K}$ a lépéshosszt jelöli. Látható, hogy a mátrixok csak a tridiagonálisokban tartalmaznak nemnulla elemet, mivel $\varphi_i \varphi_j$, ill. $\varphi_i' \varphi_j'$ értéke nulla, ha $|i-j| > 2$.

4.2. A Petrov–Galjorkin-módszer

A Petrov–Galjorkin-módszernél a próbafüggvény és a keresett közelítő megoldás tere nem egyezik meg, és bázisfüggvényeik különbözőek. Megválasztásuktól függően többféle Petrov–Galjorkin-módszer használatos. Az általunk választott módszert Hundsdorfer and Verwer (2003) alapján mutatjuk be. A W_h tér bázisfüggvényei, a tényleges próbafüggvények (2. ábra) az alábbi alakban írhatók fel:

$$\psi_j(x) = \varphi_j(x) + \kappa\sigma_j(x),$$

$$\text{ahol } \sigma_j(x) = \begin{cases} \frac{3}{h^2}(x-x_{j-1})(x_j-x), & x_{j-1} \leq x \leq x_j \\ -\frac{3}{h^2}(x-x_j)(x_{j+1}-x), & x_j \leq x \leq x_{j+1} \end{cases}$$



2. ábra – A próbafüggvény (folytonos kék vonal) és a bázisfüggvénye (szaggatott zöld vonal)

és $\kappa = \frac{e^\mu + 1}{e^\mu - 1} - \frac{2}{\mu}$, illetve $\mu = \frac{ah}{d}$ az ún. Péclet-szám. A módszert advekción-diffúziós differenciálegyenleteknél szokták használni, ahol a d a diffúziós együttható, ami a mi esetünkben nulla. Így μ végtelen, és κ határértéke 1.

Ekkor (13) helyett a

$$\sum_{i=0}^K \alpha_i'(t) \left(\int_0^\pi \varphi_i \varphi_j dx + \int_0^\pi \varphi_i \sigma_j dx \right) + u_0 \alpha_i(t) \left(\int_0^\pi \varphi_i' \varphi_j dx + \int_0^\pi \varphi_i' \sigma_j dx \right) = 0 \quad j=0, \dots, K$$

$$\underline{f_3} = \begin{bmatrix} \frac{h}{6} + \frac{h}{4} - \frac{u_0 \Delta t \theta}{2} - \frac{u_0 \Delta t \theta}{2} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{és} \quad \underline{f_4} = \begin{bmatrix} \frac{h}{6} + \frac{h}{4} + \frac{u_0 \Delta t (1-\theta)}{2} + \frac{u_0 \Delta t (1-\theta)}{2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

4.3. A legkisebb négyzetek módszere

A legkisebb négyzetek módszerének ismertetéséhez az előzőektől eltérően nem időfüggő feladatból indulunk ki. Tekintsük az $Lw = f$ alakú egyenletet, ahol L egy operátor, w ismeretlen, f pedig ismert $L^2(\Omega)$ -beli függvény. Ennek az egyenletnek az

$$\langle Lw, L\eta \rangle = \langle f, L\eta \rangle \quad \forall \eta \in L^2(\Omega)$$

gyenge alakjában a próbafüggvény η helyett $L\eta$ alakú. A diszkrét feladat

$$(19) \quad \langle Lw_h, L\eta_h \rangle = \langle f, L\eta_h \rangle \quad \forall \eta_h \in V_h, \quad w_h \in V_h.$$

A (19) egyenletnek létezik egyetlen w_h megoldása, és érvényes rá a

$$\|w - w_h\| \leq \frac{\|L\|}{\alpha} \inf_{\eta_h \in V_h} \|w - \eta_h\| \quad (\forall w \in L_2(\Omega): \|Lw\| \geq \alpha \|w\|)$$

hibabecslés (Ern et al., 2004).

Mivel a legkisebb négyzetek módszerét csak a térbeli diszkrétizációra használjuk, ezért első lépésként az (4) egyenletben szereplő időderiváltra alkalmazzuk a θ -módszert!

$$(20) \quad u_0 \Delta t \theta \cdot c^n \prime(x) + c^n(x) = u_0 \Delta t (\theta - 1) c^{n-1} \prime + c^{n-1}(x) \quad n=1, \dots, N$$

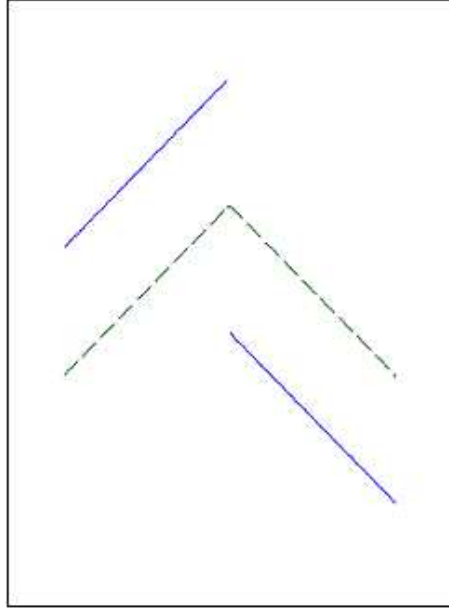
Ekkor az L operátorra $Lc^n = u_0 \Delta t \theta \frac{d}{dx} c^n + 1 \cdot c^n$, (20) diszkrét gyenge alakja pedig

$$\begin{aligned} \int_0^\pi \left(c_h^n + u_0 \Delta t \theta \frac{d}{dx} c_h^n \right) \left(\eta_h + u_0 \Delta t \theta \frac{d}{dx} \eta_h \right) dx = \\ \int_0^\pi \left(c_h^{n-1} + u_0 \Delta t (\theta - 1) \frac{d}{dx} c_h^{n-1} \right) \left(\eta_h + u_0 \Delta t \theta \frac{d}{dx} \eta_h \right) dx \end{aligned} \quad \forall \eta_h \in V_h; \quad n=1, \dots, N,$$

ahol c_h^n közelítő megoldást a következő alakban keressük: $c_h^n(x) = \sum_{i=0}^K \alpha_i^n \varphi_i(x)$. Az $L\eta_h$

próbafüggvény helyett elegendő $\varphi_j + u_0 \Delta t \theta \varphi_j \prime$ függvényeket ($j=0, \dots, K$) (3. ábra) vizsgálni.

Így



3. ábra – A próbafüggvény (folytonos vonal) és a keresett függvény bázisfüggvénye (szaggatott vonal)

$$\sum_{i=0}^K \left(\int_0^{\pi} \varphi_i \varphi_j + u_0 \Delta t \theta \int_0^{\pi} \varphi_i \varphi_j' + u_0 \Delta t \theta \int_0^{\pi} \varphi_i' \varphi_j + (u_0 \Delta t \theta)^2 \int_0^{\pi} \varphi_i' \varphi_j' \right) \alpha_i^n =$$

$$\sum_{i=0}^K \left(\int_0^{\pi} \varphi_i \varphi_j + u_0 \Delta t \theta \int_0^{\pi} \varphi_i \varphi_j' + u_0 \Delta t (\theta - 1) \int_0^{\pi} \varphi_i' \varphi_j + (u_0 \Delta t)^2 \theta (\theta - 1) \int_0^{\pi} \varphi_i' \varphi_j' \right) \alpha_i^{n-1}$$

$j = 0, \dots, K; \quad n = 1, \dots, N.$

Felhasználva a peremfeltételt a

$$\sum_{i=0}^K \left(\int_0^{\pi} \varphi_i \varphi_j + u_0 \Delta t \theta \int_0^{\pi} \varphi_i \varphi_j' + u_0 \Delta t \theta \int_0^{\pi} \varphi_i' \varphi_j + (u_0 \Delta t \theta)^2 \int_0^{\pi} \varphi_i' \varphi_j' \right) \alpha_i^n =$$

$$\sum_{i=0}^K \left(\int_0^{\pi} \varphi_i \varphi_j + u_0 \Delta t \theta \int_0^{\pi} \varphi_i \varphi_j' + u_0 \Delta t (\theta - 1) \int_0^{\pi} \varphi_i' \varphi_j + (u_0 \Delta t)^2 \theta (\theta - 1) \int_0^{\pi} \varphi_i' \varphi_j' \right) \alpha_i^{n-1}$$

(21)

$$- g(t_n) \left(\int_0^{\pi} \varphi_0 \varphi_j + u_0 \Delta t \theta \int_0^{\pi} \varphi_0 \varphi_j' + u_0 \Delta t \theta \int_0^{\pi} \varphi_0' \varphi_j + (u_0 \Delta t)^2 \theta^2 \int_0^{\pi} \varphi_0' \varphi_j' \right)$$

$$+ g(t_{n-1}) \left(\int_0^{\pi} \varphi_0 \varphi_j + u_0 \Delta t \theta \int_0^{\pi} \varphi_0 \varphi_j' + u_0 \Delta t (\theta - 1) \int_0^{\pi} \varphi_0' \varphi_j + (u_0 \Delta t)^2 \theta (\theta - 1) \int_0^{\pi} \varphi_0' \varphi_j' \right)$$

$j = 0, \dots, K; \quad n = 1, \dots, N.$

egyenletrendszeret kapjuk. Az integrálokat kiszámolva jutunk az

$$(22) \quad \left(\underline{S}_1 + u_0 \Delta t \theta (\underline{S}_5 + \underline{S}_2) + (u_0 \Delta t \theta)^2 \underline{S}_6 \right) \underline{\alpha}^n =$$

$$\left(\underline{S}_1 + u_0 \Delta t \theta \underline{S}_5 + u_0 \Delta t (\theta - 1) \underline{S}_2 + (u_0 \Delta t)^2 \theta (\theta - 1) \underline{S}_6 \right) \underline{\alpha}^{n-1} - g(t_n) \underline{f}_5 + g(t_{n-1}) \underline{f}_6$$

$n = 1, \dots, N$

egyenletrendszerhez, ahol

$$\underline{\underline{S}}_5 = \frac{1}{2} \begin{bmatrix} 0 & -1 & & & & \\ 1 & 0 & -1 & & 0 & \\ & 1 & \ddots & & & \\ & & & \ddots & -1 & \\ 0 & & & 1 & 0 & -1 \\ & & & & 1 & 1 \end{bmatrix}, \quad \underline{\underline{S}}_6 = \frac{1}{h} \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & 0 & \\ & -1 & \ddots & & & \\ & & & \ddots & -1 & \\ 0 & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{bmatrix},$$

$$\underline{f}_5 = \begin{bmatrix} \frac{h}{6} - \frac{u_0 \Delta t \theta}{2} + \frac{u_0 \Delta t \theta}{2} - \frac{u_0^2 (\Delta t)^2 \theta^2}{h} \\ \vdots \\ 0 \end{bmatrix}$$

$$\text{és } \underline{f}_6 = \begin{bmatrix} \frac{h}{6} - \frac{u_0 \Delta t (\theta - 1)}{2} + \frac{u_0 \Delta t}{2} - \frac{u_0^2 (\Delta t)^2 \theta (\theta - 1)}{h} \\ \vdots \\ 0 \end{bmatrix}.$$

4.4. Periodikus peremfeltétel

Az oldalsó helyett periodikus peremfeltétellel is megvizsgáltuk az advekción feladatot. Ekkor tehát a

$$c(0, t) = c(2\pi, t)$$

peremfeltétel használtuk, amitől megváltozott a végső lineáris egyenletrendszerek alakja.

Tekintsük először a hagyományos véges elemes módszert! A (14) egyenletbe az $\alpha_0^n = \alpha_K^n$ ($n = 1, \dots, N$) összefüggést helyettesítjük, és a

$$\begin{aligned} \sum_{i=1}^K \left(\int_0^\pi \varphi_i \varphi_j dx + u_0 \Delta t \theta \int_0^\pi \varphi_i' \varphi_j dx \right) \alpha_i^n + \left(\int_0^\pi \varphi_0 \varphi_j dx + u_0 \Delta t \theta \int_0^\pi \varphi_0' \varphi_j dx \right) \alpha_K^n = \\ \sum_{i=1}^K \left(\int_0^\pi \varphi_i \varphi_j dx - u_0 \Delta t (1 - \theta) \int_0^\pi \varphi_i' \varphi_j dx \right) \alpha_i^{n-1} + \left(\int_0^\pi \varphi_0 \varphi_j dx - u_0 \Delta t (1 - \theta) \int_0^\pi \varphi_0' \varphi_j dx \right) \alpha_K^{n-1} \end{aligned}$$

$j=1, \dots, K; \quad n=1, \dots, N$

egyenletrendszer kapjuk, amely az integrálok kiszámolása után az

$$\left(\underline{\underline{S_1'}} + u_0 \Delta t \theta \underline{\underline{S_2'}} \right) \underline{\underline{\alpha}}^n = \left(\underline{\underline{S_1'}} - u_0 \Delta t (1 - \theta) \underline{\underline{S_2'}} \right) \underline{\underline{\alpha}}^{n-1} \quad n=1, \dots, N$$

alakot ölti, ahol

$$\underline{\underline{S_1'}} = \frac{h}{6} \begin{bmatrix} 4 & 1 & & & 1 \\ 1 & 4 & 1 & & 0 \\ & 1 & \ddots & & \\ & & & \ddots & 1 \\ 0 & & & 1 & 4 & 1 \\ & & & & 1 & 2 \end{bmatrix}, \quad \underline{\underline{S_2'}} = \frac{1}{2} \begin{bmatrix} 0 & 1 & & & -1 \\ -1 & 0 & 1 & & 0 \\ & -1 & \ddots & & \\ & & & \ddots & -1 \\ 0 & & & -1 & 0 & 1 \\ & & & & -1 & 1 \end{bmatrix}$$

A Petrov–Galjorkin-módszer esetén a peremfeltétel felhasználása után (17) helyett

a

$$\begin{aligned}
& \sum_{i=1}^K \left(\int_0^\pi \varphi_i \varphi_j dx + \int_0^\pi \varphi_i \sigma_j dx + u_0 \Delta t \theta \int_0^\pi \varphi_i' \varphi_j dx + u_0 \Delta t \theta \int_0^\pi \varphi_i' \sigma_j dx \right) \alpha_i^n \\
& \left(\int_0^\pi \varphi_0 \varphi_j dx + \int_0^\pi \varphi_0 \sigma_j dx + u_0 \Delta t \theta \int_0^\pi \varphi_0' \varphi_j dx + u_0 \Delta t \theta \int_0^\pi \varphi_0' \sigma_j dx \right) \alpha_K^n = \\
& \sum_{i=1}^K \left(\int_0^\pi \varphi_i \varphi_j dx + \int_0^\pi \varphi_i \sigma_j dx - u_0 \Delta t (1-\theta) \int_0^\pi \varphi_i' \varphi_j dx - u_0 \Delta t (1-\theta) \int_0^\pi \varphi_i' \sigma_j dx \right) \alpha_i^{n-1} \\
& + \left(\int_0^\pi \varphi_0 \varphi_j dx + \int_0^\pi \varphi_0 \sigma_j dx - u_0 \Delta t (1-\theta) \int_0^\pi \varphi_0' \varphi_j dx - u_0 \Delta t (1-\theta) \int_0^\pi \varphi_0' \sigma_j dx \right) \alpha_K^{n-1} \\
& \qquad \qquad \qquad j=1, \dots, K; \quad n=1, \dots, N.
\end{aligned}$$

egyenletrendszerhez jutunk, ami az integrálok kiszámítása után a következőképpen alakul:

$$\left(\underline{S_1'} + \underline{S_3'} + u_0 \Delta t \theta (\underline{S_2'} + \underline{S_4'}) \right) \underline{\alpha}^n = \left(\underline{S_1'} + \underline{S_3'} + u_0 \Delta t (1-\theta) (\underline{S_2'} + \underline{S_4'}) \right) \underline{\alpha}^{n-1}$$

$n=1, \dots, N,$

ahol

$$\underline{S_3'} = \frac{h}{4} \begin{bmatrix} 0 & -1 & & & 1 \\ 1 & 0 & -1 & & 0 \\ & 1 & \ddots & & \\ & & & \ddots & -1 \\ 0 & & & 1 & 0 & -1 \\ & & & & 1 & 1 \end{bmatrix} \text{ és } \underline{S_4'} = \frac{1}{2} \begin{bmatrix} 2 & -1 & & & & -1 \\ -1 & 2 & -1 & & & 0 \\ & -1 & \ddots & & & \\ & & & \ddots & -1 & \\ 0 & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{bmatrix}.$$

A legkisebb négyzetek módszerének alkalmazásánál a periodikus peremfeltétel mellett az alábbi K egyenletet kapjuk minden n -re

$$\begin{aligned}
& \sum_{i=0}^K \left(\int_0^\pi \varphi_i \varphi_j + u_0 \Delta t \theta \int_0^\pi \varphi_i \varphi_j' + u_0 \Delta t \theta \int_0^\pi \varphi_i' \varphi_j + (u_0 \Delta t \theta)^2 \int_0^\pi \varphi_i' \varphi_j' \right) \alpha_i^n + \\
& \left(\int_0^\pi \varphi_0 \varphi_j + u_0 \Delta t \theta \int_0^\pi \varphi_0 \varphi_j' + u_0 \Delta t \theta \int_0^\pi \varphi_0' \varphi_j + (u_0 \Delta t)^2 \theta^2 \int_0^\pi \varphi_0' \varphi_j' \right) \alpha_K^n = \\
& \sum_{i=0}^K \left(\int_0^\pi \varphi_i \varphi_j + u_0 \Delta t \theta \int_0^\pi \varphi_i \varphi_j' + u_0 \Delta t (\theta-1) \int_0^\pi \varphi_i' \varphi_j + (u_0 \Delta t)^2 \theta (\theta-1) \int_0^\pi \varphi_i' \varphi_j' \right) \alpha_i^{n-1} \\
& + \left(\int_0^\pi \varphi_0 \varphi_j + u_0 \Delta t \theta \int_0^\pi \varphi_0 \varphi_j' + u_0 \Delta t (\theta-1) \int_0^\pi \varphi_0' \varphi_j + (u_0 \Delta t)^2 \theta (\theta-1) \int_0^\pi \varphi_0' \varphi_j' \right) \alpha_K^{n-1} \\
& \qquad \qquad \qquad j=1, \dots, K; \quad n=1, \dots, N,
\end{aligned}$$

ami az átalakítások után az

$$\begin{aligned} & \left(\underline{\underline{S_1'}} + u_0 \Delta t \theta \left(\underline{\underline{S_5'}} + \underline{\underline{S_2'}} \right) + (u_0 \Delta t \theta)^2 \underline{\underline{S_6'}} \right) \underline{\underline{\alpha}}^n = \\ & \left(\underline{\underline{S_1'}} + u_0 \Delta t \theta \underline{\underline{S_5'}} + u_0 \Delta t (\theta - 1) \underline{\underline{S_2'}} + (u_0 \Delta t)^2 \theta (\theta - 1) \underline{\underline{S_6'}} \right) \underline{\underline{\alpha}}^{n-1} \end{aligned} \quad n=1, \dots, N$$

lineáris algebrai egyenletrendszert eredményezi, ahol

$$\underline{\underline{S_5'}} = \frac{1}{2} \begin{bmatrix} 0 & -1 & & & 1 \\ 1 & 0 & -1 & & 0 \\ & 1 & \ddots & & \\ & & & \ddots & -1 \\ & 0 & & 1 & 0 & -1 \\ & & & & 1 & 1 \end{bmatrix}$$

és

$$\underline{\underline{S_6'}} = \frac{1}{h} \begin{bmatrix} 2 & -1 & & & -1 \\ -1 & 2 & -1 & & 0 \\ & -1 & \ddots & & \\ & & & \ddots & -1 \\ & 0 & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{bmatrix}.$$

5. A tesztelés eredményei

A megoldást Matlab program segítségével számítottuk. A tesztelésnél a

$$c_0(x) = \sin(x)$$

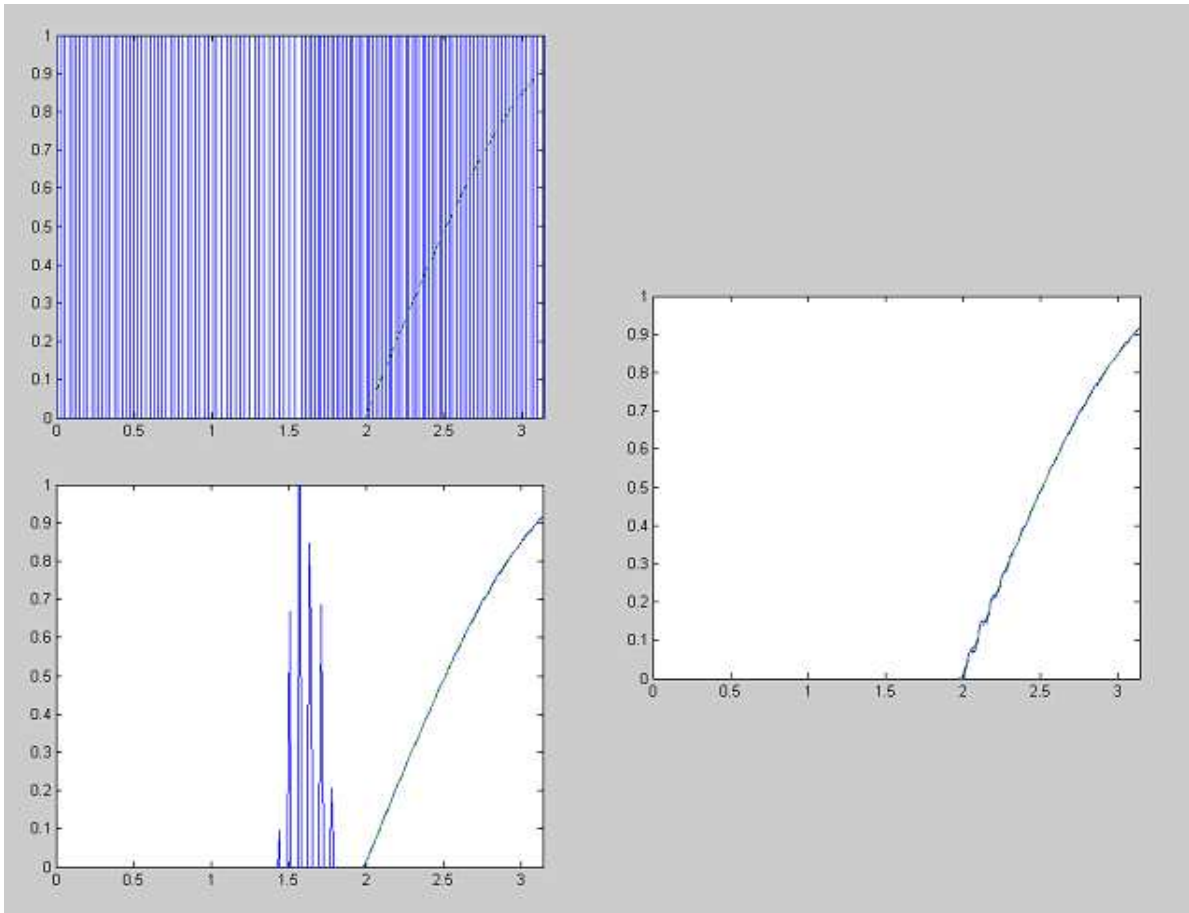
kezdeti feltételt, és az oldalsó peremfeltételes esetben a

$$g(t) = \sin(-u_0 t)$$

peremfeltételt, továbbá az időintervallum hosszára a $T = 1$ és az advekción sebességre az $u_0 = 2$ értékeket választottuk meg. Ekkor az analitikus megoldás $c(x, t) = \sin(x - 2t)$.

A megoldó programot több különböző θ -ra lefuttattuk $N=K=200$ mellett, és ábráztuk az eredményeket. Azt figyelhettük meg, hogy $\theta = 0,5$ mellett mindhárom véges elemes módszer számított megoldása egybesimul a pontos megoldással. A $\theta < 0,5$ esetekre az instabilitás, míg a $\theta > 0,5$ esetekre a számított megoldás csillapodása a jellemző. Minél távolabb esik θ a 0,5-től, annál valószínűbb ezek előfordulása.

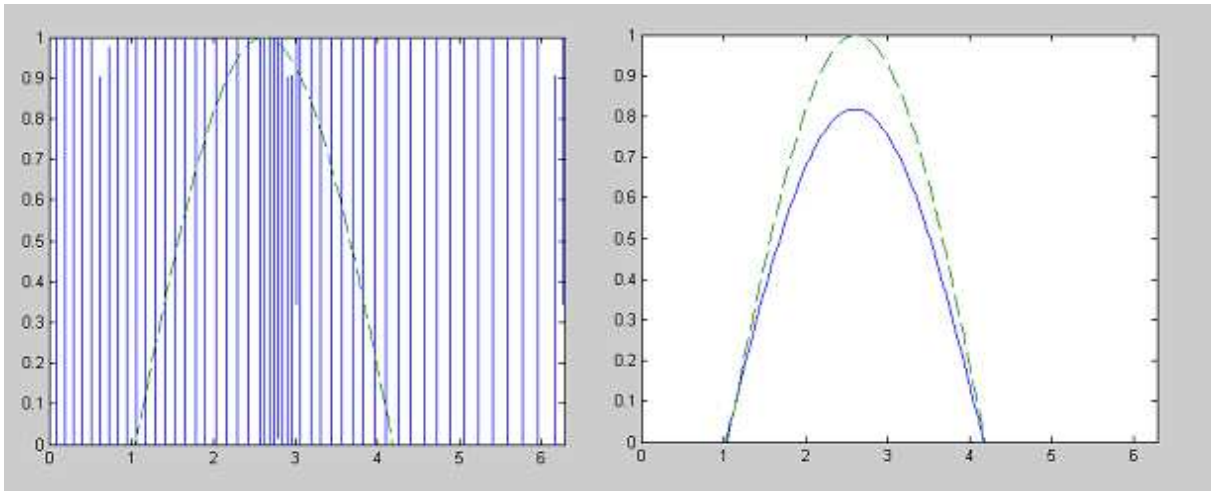
Oldalsó peremfeltétel mellett a három módszer közül a hagyományos esetben már $\theta = 0,4$ -nél megjelenik az instabilitás, míg a másik két módszer esetén csak 0,3-nél (a legkisebb négyzetek módszerének esetében a programkódot lásd a Függelékben). A 4. ábrán a $\theta = 0,3$ melletti eredmények láthatók mindhárom módszer esetén. A Petrov–Galjorkin-módszernél csak kis oszcilláció figyelhető meg, míg a hagyományos VEM megoldása instabil. A csillapodás csak $\theta = 0,9$ mellett figyelhető meg nagyon kis mértékben.



4. ábra – A hagyományos véges elemes (bal felső), a legkisebb négyzetek (bal alsó) és a Petrov–Galjorkin- (jobb oldali) módszerek (kék vonal) összehasonlítása a pontos megoldással (zöld vonal) a $T=1$ időpillanatban $\theta=0,3$ mellett

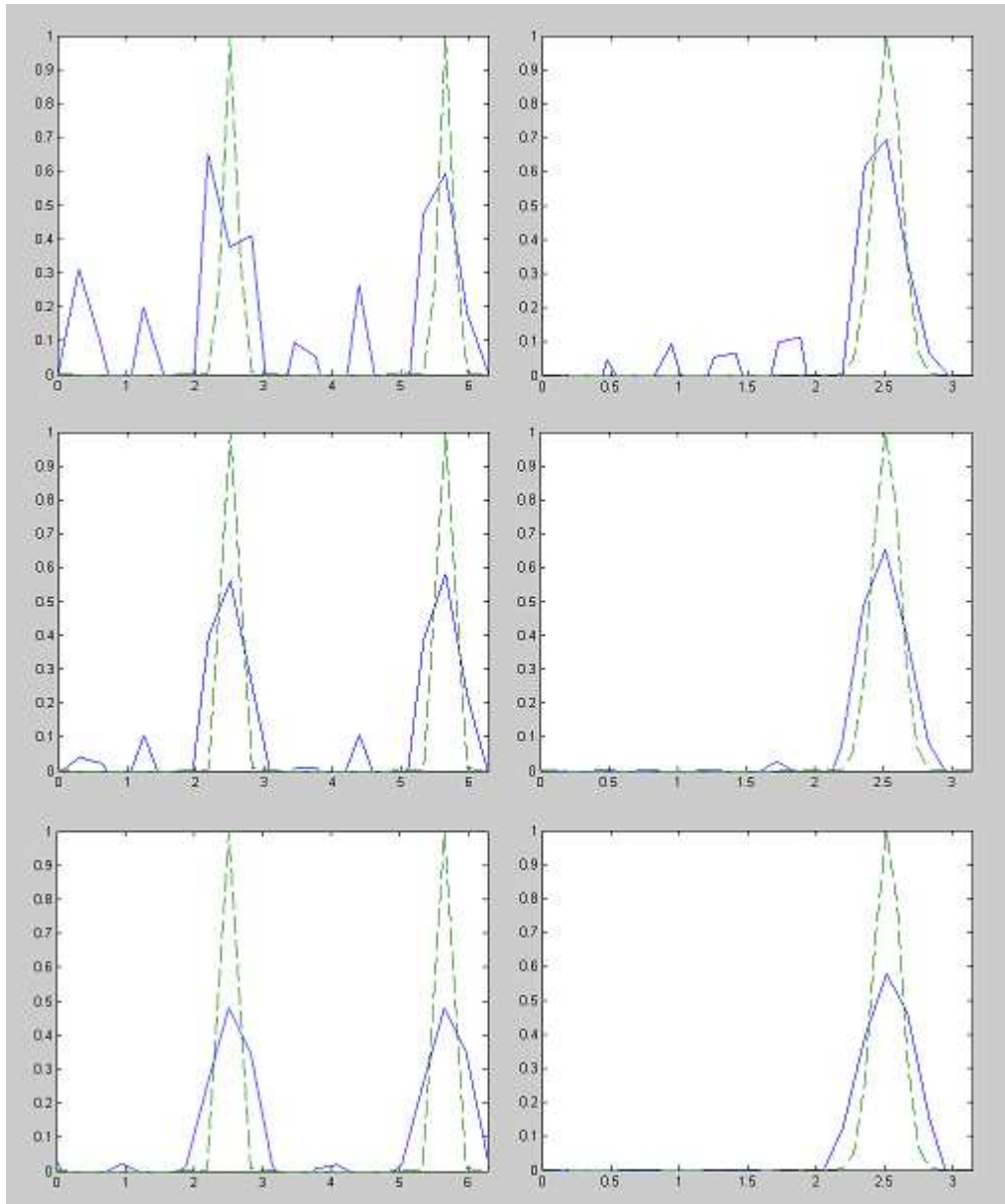
Periodikus peremfeltételnél már kisebb θ -nál (0,7-0,8) jelentkezik a csillapodás, viszont kis oszcilláció a legrosszabb esetben is csak 0,2-nél. Ez a legrosszabb eset a hagyományos véges elemes módszernél volt tapasztalható, miközben a Petrov–Galjorkin-módszernél (a programkódot lásd a Függelékben) csak $\theta = 0,05$ mellett volt oszcilláció.

Ha $T = 1$ helyett $T = 10$ -zel számolunk, akkor minden módszer esetén már $\theta = 0,4$ -nél instabilitás, $\theta = 0,6$ -nél csillapodás figyelhető meg. Az 5. ábrán a Petrov–Galjorkin-módszer eredménye látható ebben a két esetben.



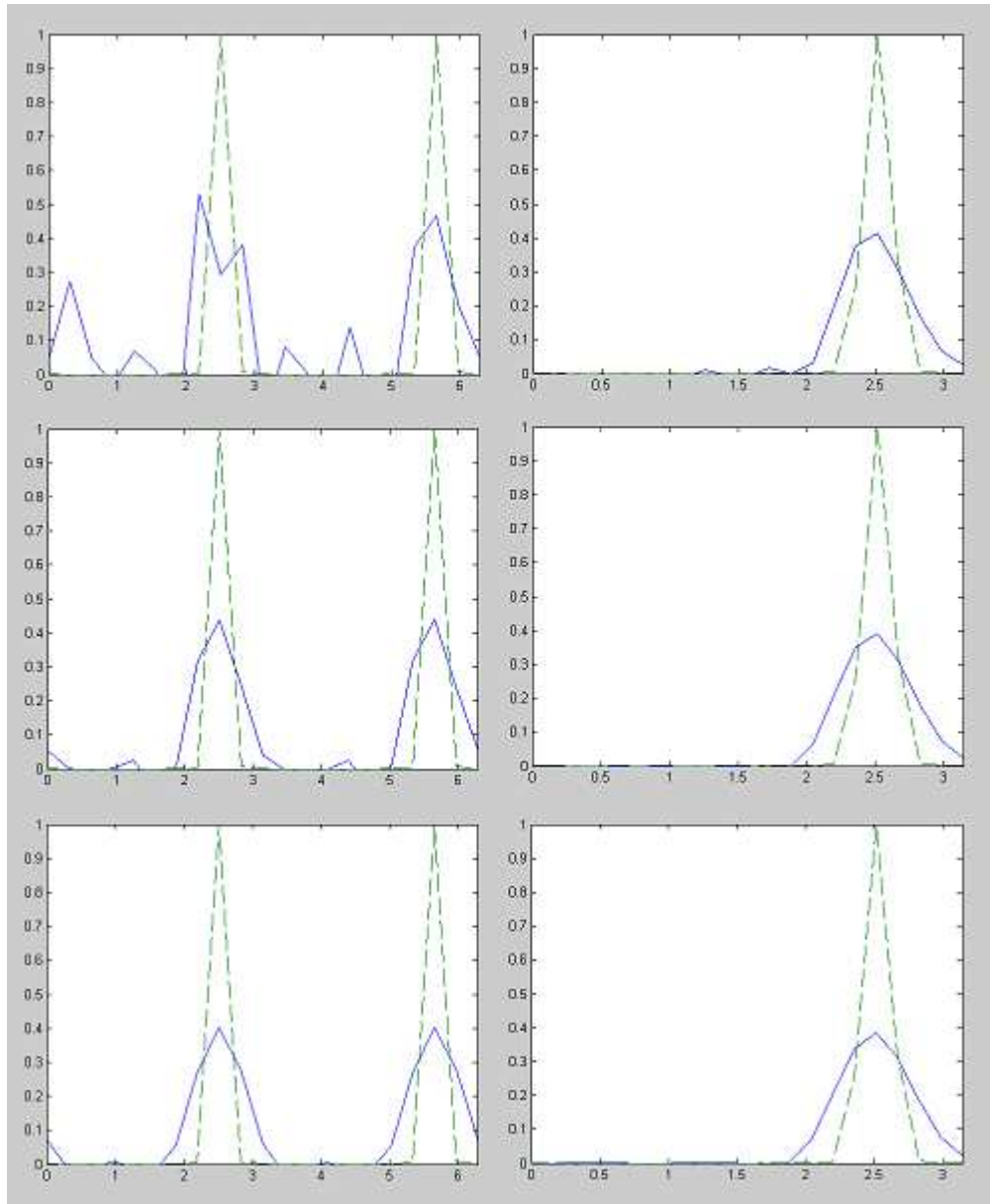
5. ábra – A Petrov–Galjorkin-módszerrel számított numerikus (kék folytonos vonal) és a pontos (zöld szaggatott) megoldás $T=10$ mellett $\theta=0,4$ (bal oldalon) és $\theta=0,6$ (jobb oldalon) esetén

A dolgozatban a θ -módszer két speciális esetében vizsgáljuk a véges elemes módszereket, nevezetesen $\theta = 0,5$ és $\theta = 1$ esetében. Megfelelő felbontás és kezdeti és peremfeltételek mellett ezen θ -k mellett is találkozhatunk oszcillációval (bár a megoldás stabil, lásd 6.fejezet). Legyenek ezek a feltételek a következők: $c_0(x) = \sin^{100}(x)$ kezdeti feltétel, $g(t) = \sin^{100}(-t)$ periodikus peremfeltétel, $u_0 = 1$ advekción sebesség, $N=K=20$ rácsfelbontás. A 6. és 7. ábra a három véges elemes séma eredményeit mutatja be $\theta = 0,5$ (6. ábra) és $\theta = 1$ (7. ábra) mellett.



6. ábra – A pontos és számított megoldás $T=1$ és $\theta=0,5$ mellett felülről rendre a hagyományos véges elemes, a legkisebb négyzetek és a Petrov–Galjorkin-módszerek esetében, a bal oldali ábrákon $[0;2\pi]$ -n periodikus peremfeltétellel, a jobb oldali ábrákon $[0;\pi]$ -n oldalsó peremfeltétellel

Mindkét ábra alapján azt mondhatjuk, hogy a hagyományos véges elemes módszerrel kapott megoldás oszcillál leginkább, míg a Petrov–Galjorkin melletti a legkevésbé, és az oldalsó peremfeltételnél kisebb az oszcilláció mértéke. Az időbeli diszkretizációs sémákat összehasonlítva pedig azt látjuk, hogy az implicit Euler-módszer ($\theta = 1$) adja a kisebb oszcillációt.



7. ábra – A pontos és számított megoldás $T=1$ és $\theta=1$ mellett felülre rendre a hagyományos véges elemes, a legkisebb négyzetek és a Petrov-Galjorkin módszerek esetében, a baloldali ábrákon $[0; 2\pi]$ -n periodikus peremfeltétellel, a jobb oldali ábrákon $[0; \pi]$ -n oldalsó peremfeltétellel

A megoldás ábrázolása mellett kiszámoltuk az oldalsó peremfeltétellel rendelkező advekción feladat hibáját is, amit a pontos és a számított megoldás különbségének diszkrét l^2 -beli normájával fejeztünk ki. Jelölje e_i^n a pontos és a közelítő megoldás különbségét az n -edik időpillanatban, az x_i helyen, e pedig az összes tér- és időbeli osztópontban értelmezett diszkrét hibafüggvényt. Ekkor az e hiba diszkrét l_2 -beli normáját az

$$\|e\|_{l^2(\Omega_T)} := \sqrt{\sum_{i=0}^K \sum_{n=0}^N (e_i^n)^2 h \Delta t}$$

képlettel értelmezzük. A hibaszámolásnál az advekción sebesség $u_0 = 3$ volt, így az analitikusan is megkapható megoldás $c(x,t) = \sin(x - 3t)$.

A hibákat $\theta = 0,5$ mellett az 1-3. táblázatok tartalmazzák. K a térbeli, N az időbeli osztásközök száma. Adott oszlopban rögzített időlépcsőhöz tartozó hibák található feleződő lépésköz mellett. Látható, hogy nagy K mellett feleakkora időlépcsőnél nagyjából negyedakkora a hiba, tehát a módszer időben másodrendűen viselkedik, amit vártunk is, hiszen a trapézformula, amit használtunk, másodrendű időintegrálási séma.

K	N=100	N=200	N=300	N=600	N=1000
4	0,072057	0,07205	0,072048	0,072048	0,072048
8	0,017304	0,017303	0,017303	0,017303	0,017303
16	0,004273	0,004272	0,004272	0,004272	0,004272
32	0,001071	0,001065	0,001064	0,001065	0,001065
64	0,000296	0,000268	0,000266	0,000266	0,000266
128	0,000149	0,000074	0,000068	0,000067	0,000066
256	0,000136	0,000037	0,000022	0,000017	0,000017
512	0,000135	0,000034	0,000015	0,000006	0,000004
1024	0,000135	0,000034	0,000015	0,000004	0,000002

1. Táblázat – A hagyományos véges elemes módszer hibája $\theta=0,5$ mellett

A hagyományos véges elemes módszer (1. táblázat) $K \sim 100$ alatt másodrendűen viselkedik, kisebb lépésközökre már csak elsőrendűen, majd egyáltalán nem csökken a hiba, mivel az idődiszkretizáció hibája gátat szab a további javulásnak. Látható, hogy az időlépések számát növelve nagyobb mértékben csökken a hiba a K növekedésével. $N \sim 200$ és $K \sim 100$ esetén a hiba már csak 10^{-5} nagyságrendű.

K	N=100	N=200	N=300	N=600	N=1000
4	0,0416	0,041642	0,04165	0,041655	0,041656
8	0,006172	0,006195	0,006199	0,006202	0,006202
16	0,000923	0,000928	0,000929	0,000931	0,000931
32	0,000192	0,000149	0,000147	0,000147	0,000148
64	0,000136	0,000041	0,000028	0,000025	0,000024
128	0,000135	0,000034	0,000015	0,000006	0,000004
256	0,000135	0,000034	0,000015	0,000004	0,000002
512	0,000135	0,000034	0,000015	0,000004	0,000001
1024	0,000135	0,000034	0,000015	0,000004	0,000001

2. Táblázat – A Petrov–Galjorkin-módszer hibája $\theta=0,5$ mellett

A Petrov–Galjorkin-módszer hibája minden adott rácsfelbontásnál kisebb, mint a hagyományos véges elemé. Maximum 10^{-5} nagyságrendű hibát már $N \sim 200$ és $K \sim 60$ körüli

értékeknél kaphatunk. Adott K értékek mellett a módszer akár jobban is viselkedik, mint egy másodrendű séma.

K	N=100	N=200	N=300	N=600	N=1000
4	0,061923	0,066621	0,068344	0,070151	0,070899
8	0,010272	0,012859	0,014075	0,015538	0,016204
16	0,001347	0,001861	0,00224	0,002919	0,003346
32	0,000219	0,000238	0,000287	0,000402	0,000513
64	0,000136	0,000045	0,000039	0,000051	0,000065
128	0,000135	0,000034	0,000016	0,000007	0,000008
256	0,000135	0,000034	0,000015	0,000004	0,000002
512	0,000135	0,000034	0,000015	0,000004	0,000001
1024	0,000135	0,000034	0,000015	0,000004	0,000001

3. Táblázat – A legkisebb négyzetek módszerének hibája $\theta=0,5$ mellett

A legkisebb négyzetek módszerének hibáját a 3. táblázat szemlélteti. Adott rácsfelbontásnál a hiba kisebb, mint a hagyományos véges elemes módszernél, de nagyobb, mint a Petrov–Galjorkin-séma esetében.

Az implicit Euler-módszert használva az időbeli diszkretizációra, mindhárom véges elemes módszer természetesen elsőrendűen viselkedik, térben kis K értékeknél másodrendűen, nagyobbaknál elsőrendűen viselkedik, míg már nincs további javulás. A hibák nagysága a három sémánál nagyjából azonos. A hagyományos véges elemes módszer hibáját a 4. táblázat tartalmazza. Az $N \sim K \sim 1000$ felbontásnál kapott hiba is 10^{-3} nagyságrendű, miközben ehhez $\theta = 0,5$ mellett már $N \sim 100$ és $K \sim 15$ rácsfelbontás is elegendő.

K	N=100	N=200	N=300	N=600	N=1000
4	0,075109	0,072901	0,072457	0,072171	0,072102
8	0,02686	0,020473	0,01893	0,017839	0,017553
16	0,019978	0,010919	0,008068	0,005592	0,004852
32	0,01922	0,009809	0,006644	0,003505	0,002305
64	0,019106	0,009675	0,006486	0,003276	0,00199
128	0,019083	0,009652	0,006461	0,003247	0,001954
256	0,019077	0,009646	0,006456	0,003241	0,001948
512	0,019076	0,009645	0,006455	0,00324	0,001947
1024	0,019076	0,009645	0,006454	0,00324	0,001947

4. Táblázat – A hagyományos véges elemes módszer hibája $\theta=1$ mellett

Összefoglalva, mind a hiba nagyságát, mind az oszcilláció mértékét vizsgálva elmondhatjuk, hogy az advekción feladat esetében a legcélszerűbb a Petrov–Galjorkin-módszert használni, mert ez a séma generálja a legkisebb zajokat és adja a legpontosabb

eredményeket. A következőkben a három vizsgált véges elemes módszer elméleti vizsgálatát végezzük el.

6. Fourier-analízis

Ha a (4) egydimenziós advekción egyenlethez π szerint periodikus kezdeti és peremfeltételt adunk meg, akkor a feladat pontos megoldása előáll diszkrét hullámmódusok összegeként, és mindegyik módus az u_0 advekción sebességgel terjed változatlan amplitúdóval. A megoldás tulajdonságai így végső soron a hullámösszetevők tulajdonságaitól függenek. Ezért egy numerikus sémát jól jellemez, hogy mennyire pontosan írja le egyetlen hullámmódus, azaz egy

$$(23) \quad c(x,0) = c_0(x) = C_k e^{ikx}$$

alakú kezdeti függvény továbbterjedését a k hullámszám függvényében. A Fourier-analízis során a hullámmegoldásokat a stabilitás, a fiktív csillapodás, a relatív fázissebesség és a számítási diszperzió szempontjából vizsgáljuk (Riddaway, 1986).

A (23) kezdeti feltételhez tartozó folytonos megoldás a $c(x,t) = C_k e^{ik(x-u_0t)}$ függvény, ahol C_k a $k = \frac{2\pi}{L}$ hullámszámtól függő konstans. $L = lh$ a hullámhossz, és l az adott hullámhosszhoz tartozó osztásközök száma (egynél nagyobb egész szám).

A $c_i^0 = C_k \exp(ikx_i)$ kezdeti feltételhez tartozó diszkrét feladat megoldása

$$(24) \quad c_i^n = (\lambda_k)^n C_k \exp(ikx_i)$$

alakú, ahol λ_k egy k hullámszámtól függő komplex mennyiség, amelyet $\lambda_k = |\lambda_k| e^{i\phi}$ alakban felírva (24) helyett a

$$c_i^n = C_k |\lambda_k|^n \exp\left(ik\left(x_i + \frac{n\phi}{k}\right)\right)$$

képletet kapjuk.

Egy sémát *stabilnak* nevezünk, ha a bemenő adatoktól folytonosan függ a megoldás. Más szóval, ha létezik olyan univerzális állandó, amellyel megszorozva a bemenő adatok normáját, a megoldás normáját felülről becsülhetjük. Ha a kezdeti függvény diszkrét hullámmódusok összege, akkor a numerikus megoldás akkor lesz stabil, ha ezen hullámmódusok egyikének sem válik végtelenné az amplitúdója az időben, azaz ha

a (24) numerikus megoldásban $|\lambda_k| \leq 1$ minden k hullámszámra, hiszen különben nem létezne a megkövetelt felső becslés.

Sok esetben a stabilitás elvár valamilyen összefüggést az idő- és térbeli lépésköz között. A gyakorlatban ez azt jelenti, hogy adott térbeli lépésköz mellett csak korlátozott időlépcsővel számolhatunk. Ha egy numerikus sémánál nincs ilyen megkötés, akkor azt a sémát abszolút stabilnak nevezzük.

Az advekción feladat pontos megoldásában a hullámok változatlan amplitúdóval haladnak, ezért a stabilitás mellett azt is szeretnénk, ha az amplitúdók nem csillapodnának az idővel, azaz ha a $D_k := |\lambda_k|$ *fiktív csillapodásra* $D_k = 1$ teljesülne. Abban az esetben, ha ez teljesül, a numerikus sémát neutrálisnak nevezzük.

Míg a pontos megoldás hullámösszetevői az u_0 advekción sebességgel terjednek, a numerikus megoldásban a k hullámszámú hullám terjedési sebességét (más szóval fázissebességét) a

$$c_k = -\frac{\phi}{k\Delta t}$$

képlet adja meg. Annak jellemzésére, hogy egy numerikus módszer mennyire jól írja le az adott hullám terjedési sebességét, az

$$r_k = \frac{c_k}{u_0}$$

relatív fázissebességét használjuk. Ha $r_k = 1$, akkor a k hullámszámú hullám fázissebességét pontosan írja le a módszer. Ha $r_k < 1$, akkor a numerikus módszer alulbecsüli, ha pedig $r_k > 1$, akkor felülbecsüli a pontos terjedési sebességet. Ha $r_k < 0$, akkor a terjedés irányát is hibásan írja le a módszer, vagyis ilyenkor az adott hullám a pozitív helyett a negatív x irányban terjed.

A numerikus séma *diszperziót* okozhat a megoldásban még akkor is, ha az eredeti egyenlet nem diszperzív. Ez azt jelenti, hogy a zavar nem őrzi meg az alakját elmozdulás

közben, azaz a $c_{g,k} = \frac{\partial}{\partial k}(kc_k)$ csoportsebesség nem egyezik meg a c_k fázissebességgel. A

csoportsebességhez tartozó relatív csoportsebesség képlete $r_{g,k} = \frac{c_{g,k}}{u_0}$.

6.1. A hagyományos véges elemes módszer

A (16) egyenletrendszer alapján a véges elemes séma a

$$\begin{aligned} & \left(-\frac{a\Delta t\theta}{2} + \frac{h}{6}\right)\alpha_{i-1}^n + \frac{4}{6}h\alpha_i^n + \left(\frac{a\Delta t\theta}{2} + \frac{h}{6}\right)\alpha_{i+1}^n = \\ & \left(\frac{h}{6} + \frac{a\Delta t(1-\theta)}{2}\right)\alpha_{i-1}^{n-1} + \frac{4}{6}\alpha_i^{n-1} + \left(\frac{h}{6} + \frac{a\Delta t(1-\theta)}{2}\right)\alpha_{i+1}^{n-1} \end{aligned}$$

egyenlettel írható le. Ezt $\frac{6}{h}$ -val megszorozva és bevezetve az $\varepsilon = \frac{u_0\Delta t}{h}$ jelölést az

$$(1-3\varepsilon\theta)\alpha_{i-1}^n + 4\alpha_i^n + (3\varepsilon\theta+1)\alpha_{i+1}^n = (1+3\varepsilon(1-\theta))\alpha_{i-1}^{n-1} + 4\alpha_i^{n-1} + (1-3\varepsilon(1-\theta))\alpha_{i+1}^{n-1}$$

összefüggést kapjuk. Az ε számot Courant-számnak nevezzük.

Mivel adott rácspontban a bázisfüggvények közül csak egy vesz fel nemnulla értéket, nevezetesen az egyet, így adott pontban $c_i^n = \alpha_i^n$. Tehát a (24) kifejezést a fenti egyenletbe helyettesítve az

$$\begin{aligned} & (1-3\varepsilon\theta)(\lambda_k)^n C_k \exp(ik(x_i-h)) + 4(\lambda_k)^n C_k \exp(ikx_i) + \\ & (3\varepsilon\theta+1)(\lambda_k)^n C_k \exp(ik(x_i+h)) = (\lambda_k)^{n-1} C_k [(1+3\varepsilon(1-\theta))\exp(ik(x_i-h)) + \\ & 4\exp(ikx_i) + (1-3\varepsilon(1-\theta))\exp(ik(x_i+h))] \end{aligned}$$

egyenletet kapjuk. Egyszerűsítsünk a közös $C_k(\lambda_k)^{n-1}\exp(ikx_i)$ tényezővel!

$$(1-3\varepsilon\theta)\lambda_k e^{-ikh} + 4\lambda_k + (3\varepsilon\theta+1)\lambda_k e^{ikh} = (1+3\varepsilon(1-\theta))e^{-ikh} + 4 + (1-3\varepsilon(1-\theta))e^{ikh}$$

Ebből kifejezhető λ_k a következőképpen:

$$\begin{aligned} \lambda_k &= |\lambda_k|(\cos\theta + i\sin\theta) = \frac{(1+3\varepsilon(1-\theta))e^{-ikh} + 4 + (1-3\varepsilon(1-\theta))e^{ikh}}{(1-3\varepsilon\theta)e^{-ikh} + 4 + (3\varepsilon\theta+1)e^{ikh}} = \\ & \frac{(2+\cos kh)^2 + 9\varepsilon^2\theta(\theta-1)\sin^2 kh}{(2+\cos kh)^2 + 9\varepsilon^2\theta^2\sin^2 kh} - i \frac{3\varepsilon\sin kh(2+\cos kh)}{(2+\cos kh)^2 + 9\varepsilon^2\theta^2\sin^2 kh} \end{aligned}$$

Mivel $|\lambda_k|^2 = |\lambda_k|^2 \cos^2\phi + |\lambda_k|^2 \sin^2\phi$, így a

$$|\lambda_k|^2 = \frac{(2+\cos kh)^4 + 9\varepsilon^2(1-2\theta+2\theta^2)\sin^2 kh(2+\cos kh)^2}{[(2+\cos kh)^2 + 9\varepsilon^2\sin^2 kh]^2}$$

kifejezést kapjuk, melynek értéke a stabilitás érdekében nem lehet nagyobb egynél. Ezt a feltételt felhasználva az

$$(25) \quad \frac{(1-\theta)^2}{\theta^4} (2 + \cos kh)^2 - 9\varepsilon^2 \sin^2 kh \leq 0$$

egyenlőtlenséghez jutunk. Az implicit Euler-sémával ($\theta = 1$) a hagyományos véges elemes módszer esetén (25) mindig fennáll, azaz a módszer abszolút stabil. Más θ mellett csak a feltételes stabilitás állhat fenn, ha teljesül, hogy

$$(26) \quad \Delta t \geq \frac{1-\theta}{\theta^2} \frac{h}{3a} \frac{2 + \cos kh}{\sin kh}.$$

Nem tapasztalunk csillapodást az implicit Euler-séma esetén, ha $D_k = 1$, azaz ha $9\varepsilon^2 \sin^2 kh = 0$. Ez akkor teljesül, ha $kh = \frac{2\pi}{l} = p\pi$ ($p \in Z$), vagyis ha a $2/l$ felírható egy egész szám feleként. Ilyen l érték csak egy létezik, nevezetesen a 2. A trapézformula esetén a neutralitáshoz az egyenlőségnek kell teljesülnie a (26) feltételben.

Mivel $tg\phi = \frac{\text{Im } \lambda_k}{\text{Re } \lambda_k}$, a fázissebesség és a relatív fázissebesség a következő alakot

ölti:

$$c_k = -\frac{1}{k\Delta t} \arctg \left(-\frac{3\varepsilon \sin kh(2 + \cos kh)}{(2 + \cos kh)^2 + 9\varepsilon^2 \theta(\theta - 1) \sin^2 kh} \right),$$

$$r_k = -\frac{1}{\varepsilon h k} \arctg \left(-\frac{3\varepsilon \sin kh(2 + \cos kh)}{(2 + \cos kh)^2 + 9\varepsilon^2 \theta(\theta - 1) \sin^2 kh} \right).$$

A csoportsebesség és a relatív csoportsebesség $\theta = 1$ mellett:

$$c_{g,k} = \frac{1}{\Delta t} \frac{3\varepsilon h(1 + 2 \cos kh)}{(2 + \cos kh)^2 + 9\varepsilon^2 \sin^2 kh},$$

$$r_{g,k} = \frac{3(1 + 2 \cos kh)}{(2 + \cos kh)^2 + 9\varepsilon^2 \sin^2 kh}.$$

6.2. A Petrov–Galjorkin-módszer

A (18) egyenletrendszerből a séma

$$\begin{aligned} & \left(\frac{h}{6} + \frac{h}{4} - a\Delta t\theta \right) \alpha_{i-1}^n + \left(\frac{4}{6}h + a\Delta t\theta \right) \alpha_i^n + \left(\frac{h}{6} - \frac{h}{4} \right) \alpha_{i+1}^n = \\ & \left(\frac{h}{6} + \frac{h}{4} + a\Delta t(1-\theta) \right) \alpha_{i-1}^{n-1} + \left(\frac{4}{6}h - a\Delta t(1-\theta) \right) \alpha_i^{n-1} + \left(\frac{h}{6} - \frac{h}{4} \right) \alpha_{i+1}^{n-1} \end{aligned}$$

alakú. Ezt átalakítva az előzőekhez hasonlóan, a

$$|\lambda_k|^2 = \frac{(2(1+3\varepsilon(1-\theta))\cos kh + 2(2-3\varepsilon(1-\theta)))^2 + 9(1+2\varepsilon(1-\theta))^2 \sin^2 kh}{[2(1-3\varepsilon\theta)\cos kh + 2(2+3\varepsilon\theta)]^2 + 9(1-2\varepsilon\theta)^2 \sin^2 kh}.$$

kifejezést kapjuk. A stabilitás $|\lambda_k|^2 \leq 1$ feltétele teljesül, ha

$$(1 - \cos kh)^2 + 6\varepsilon(2\theta - 1)(1 - \cos kh) \geq 0.$$

Ez mindig igaz, ha $\theta \geq 0,5$. Kisebbs θ értékekre csak feltételesen stabil a módszer.

A $D_k = 1$ feltétel teljesül, azaz a módszer neutrális, ha $\cos kh = 1$. Ez akkor következik be, ha $kh = \frac{2\pi}{l} = 2p\pi$ ($p \in \mathbb{Z}$), azaz ha l reciproka egész szám. Ilyen l érték nincs, ezért mindig van csillapodás.

A fázissebességet és a relatív fázissebességet a következő képletek adják meg:

$$\begin{aligned} c_k &= -\frac{1}{k\Delta t} \operatorname{arctg} \left(\frac{6\varepsilon \sin kh (\cos kh - 7)}{(-1 - 6\varepsilon)\cos^2 kh + (6 + 3\varepsilon - 15\varepsilon\theta - 6\varepsilon^2\theta)\cos kh + 25 - 6\varepsilon - 36\varepsilon\theta} \right), \\ r_k &= -\frac{1}{kh\varepsilon} \operatorname{arctg} \left(\frac{6\varepsilon \sin kh (\cos kh - 7)}{(-1 - 6\varepsilon)\cos^2 kh + (6 + 3\varepsilon - 15\varepsilon\theta - 6\varepsilon^2\theta)\cos kh + 25 - 6\varepsilon - 36\varepsilon\theta} \right). \end{aligned}$$

A csoportsebesség és relatív csoportsebesség képlete nagyon bonyolult, ezért ezeket itt nem tüntetjük fel, de az utóbbi l -től való függését a 12-13. ábra szemlélteti.

6.3. A legkisebb négyzetek módszere

A séma (22) alapján:

$$\begin{aligned} & \left(\frac{h}{6} - \frac{a^2(\Delta t)^2 \theta^2}{h} \right) \alpha_{i-1}^n + \left(\frac{4}{6} h + \frac{2a^2(\Delta t)^2 \theta^2}{h} \right) \alpha_i^n + \\ & \left(\frac{h}{6} - \frac{a^2(\Delta t)^2 \theta^2}{h} \right) \alpha_{i+1}^n = \left(\frac{h}{6} + \frac{a\Delta t}{2} - \frac{a^2(\Delta t)^2(\theta-1)\theta}{h} \right) \alpha_{i-1}^{n-1} + \\ & \left(\frac{4}{6} h + \frac{2a^2(\Delta t)^2(\theta-1)\theta}{h} \right) \alpha_i^{n-1} + \left(\frac{h}{6} - \frac{a\Delta t}{2} - \frac{a^2(\Delta t)^2(\theta-1)\theta}{h} \right) \alpha_{i+1}^{n-1} \end{aligned}$$

Innen az előzőekhez hasonlóan levezethető, hogy

$$|\lambda_k|^2 = \frac{\left[(1 - 6\varepsilon^2 \theta(\theta - 1)) \cos kh + 2(2 + 6\varepsilon^2 \theta(\theta - 1)) \right]^2 + 9\varepsilon^2 \sin^2 kh}{\left[(1 - 6\varepsilon^2 \theta^2) \cos kh + (2 + 6\varepsilon^2 \theta^2) \right]^2}.$$

A $|\lambda_k|^2 \leq 1$ egyenlőtlenségbe behelyettesítve, átalakítások után a

$$\left[(4\theta - 3) \cos^2 kh + 4\theta \cos kh - 8\theta + 3 \right] + 12\varepsilon^2 \theta(1 - 2\theta)(\cos kh - 1)^2 \leq 0,$$

egyenlőtlenséget kapjuk, amelyről belátható, hogy minden ε -ra teljesül $\theta \geq 0,5$ esetén, azaz ekkor a módszer abszolút stabil.

Az egyenlőség akkor igaz $\theta \geq 0,5$ mellett, ha

$$\left[(4\theta - 3) \cos^2 kh + 4\theta \cos kh - 8\theta + 3 \right] = 0$$

és $12\varepsilon^2 \theta(1 - 2\theta)(\cos kh - 1)^2 = 0.$

Ez csak akkor következik be, ha $\cos kh = 1$, tehát a módszer neutrális, ha $kh = \frac{2\pi}{l} = 2p\pi$

($p \in Z$). Ilyen l érték nincs, ezért mindig van csillapodás.

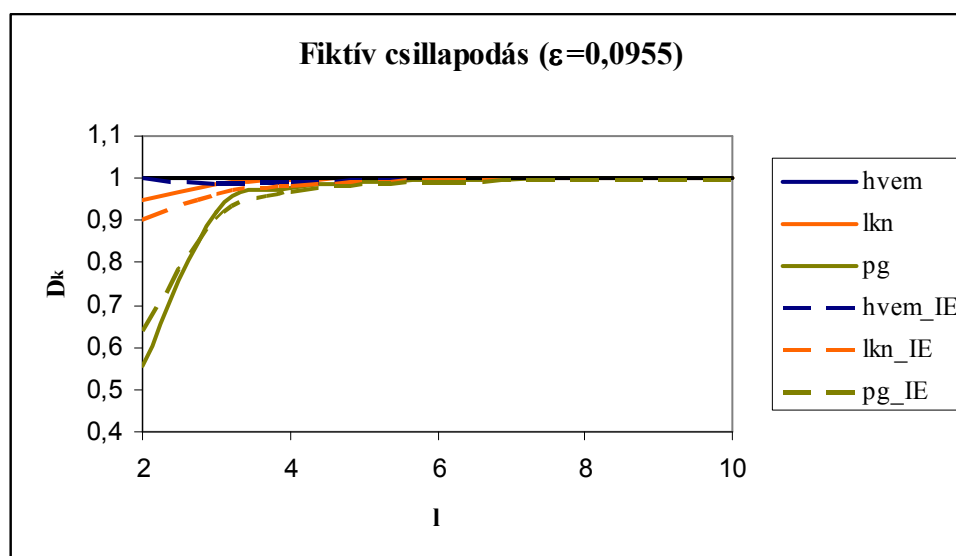
Az implicit Euler idődiszkrétizációs sémát használva a legkisebb négyzetek és a hagyományos véges elemes módszernél azonos fázis- és csoportsebességeket kapunk. A fázissebesség és relatív fázissebesség:

$$c_k = -\frac{1}{k\Delta t} \operatorname{arctg} \left(-\frac{3\varepsilon \sin kh}{2 + \cos kh} \right),$$

$$r_k = -\frac{1}{\varepsilon hk} \operatorname{arctg} \left(-\frac{3\varepsilon \sin kh}{2 + \cos kh} \right).$$

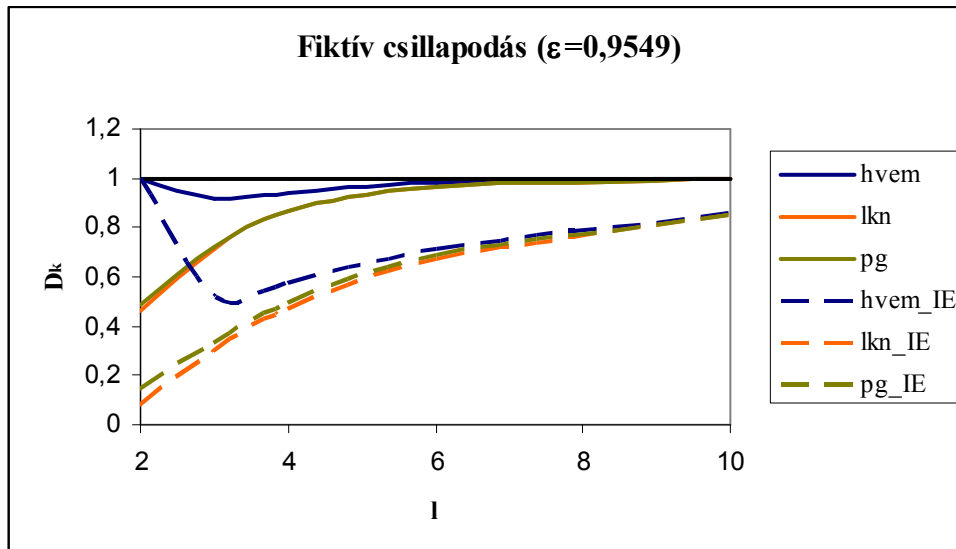
6.5. Eredmények

A 8-13. ábrák szemléltetik a fiktív csillapodást, a relatív fázissebességet, illetve a relatív csoportsebességet az l osztásközszám, így közvetve az L hullámhossz függvényében. Mindhárom értéknél azt szeretnénk, ha egyenlők lennének eggyel, hiszen ez azt jelentené, hogy a séma neutrális, továbbá a fázis,- ill. csoportsebesség megegyezik az advekción sebességével. Minden ábrán hat görbét láthatunk, a három véges elemes módszer eredményeit a két speciális idődiszkrétizáció esetén, nevezetesen a trapézformula és az implicit Euler-séma (IE) esetén. A 9., 10. és 12. ábrán látható eredmények $\varepsilon = 0,9549$ ($N = K = 200$) mellett, míg a 8., 11. és 13. ábra eredményei $\varepsilon = 0,0955$ ($N = 1000$ és $K = 100$) mellett érvényesek. Mindegyik sémánál a nagyobb l értékek, azaz nagyobb hullámhosszak és kisebb Courant-számok esetében lesznek az értékek közelebb egyhez.



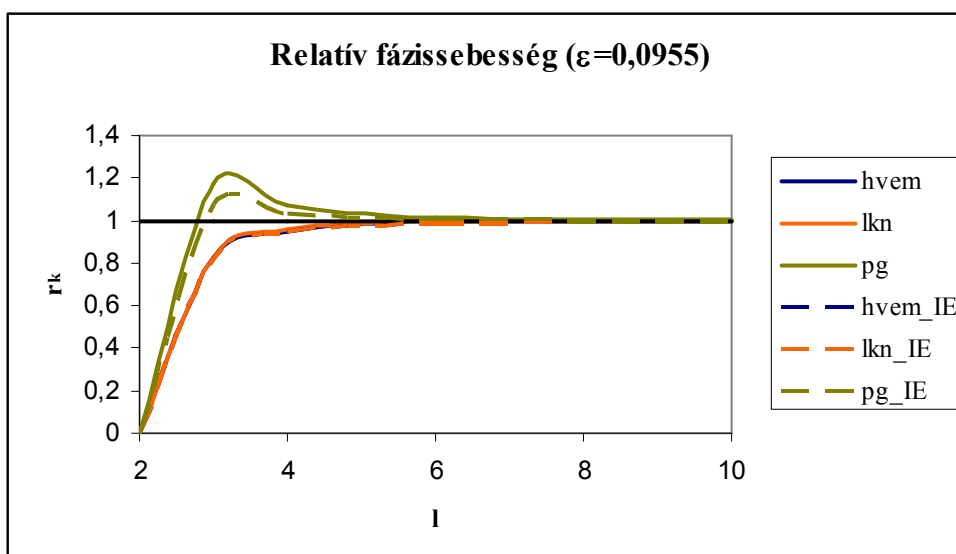
8. ábra – A hagyományos véges elem, a Petrov–Galjorkin-módszer és a legkisebb négyzetek módszerének trapézformulát, ill. implicit Euler-sémát alkalmazó változatának fiktív csillapodása l függvényében $\varepsilon=0,0955$ mellett

A fiktív csillapodás (8-9. ábra) értékei a hullámhossz növekedésével az egyhez tartanak. Az eredmények a hagyományos véges elemes módszer esetén a legjobbak mindkét idődiszkrétizációs sémánál, ill. Courant-számnál. Adott véges elemes módszernél a $\theta = 0,5$ esetben kapunk kedvezőbb D_k mennyiségeket.

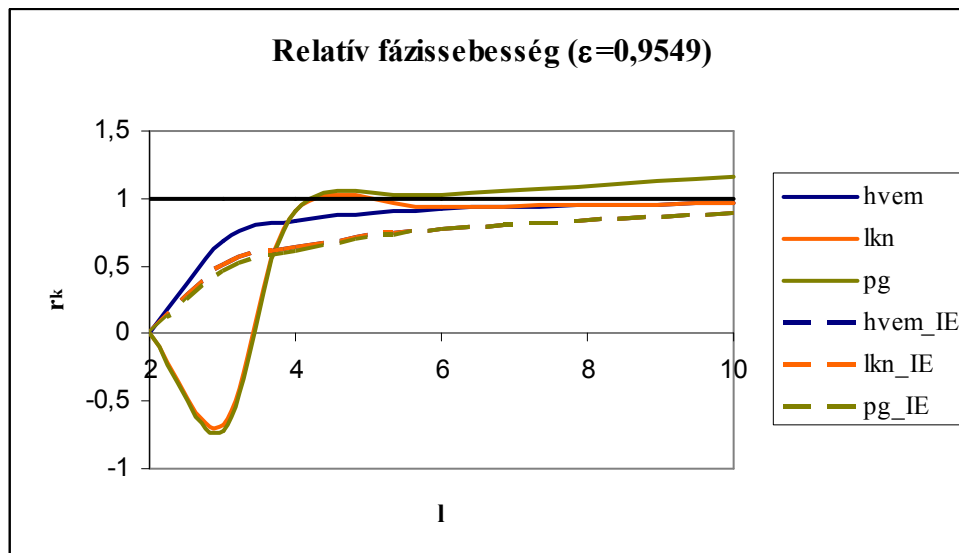


9. ábra – A hagyományos véges elem, a Petrov–Galjorkin-módszer és a legkisebb négyzetek módszerének trapézformulát, ill. implicit Euler-sémát alkalmazó változatának fiktív csillapodása l függvényében $\varepsilon=0,9549$ mellett

A relatív fázissebesség (10-11. ábra) értékei (a relatív csoportsebességeké is) a hagyományos véges elem és legkisebb négyzetek módszerénél megegyeznek $\theta = 1$ esetén, és nagyon hasonlóak $\theta = 0,5$ és $\varepsilon = 0,0955$ mellett. Ezek a sémák adják a legjobb eredményeket mindkét θ -nál. A Petrov–Galjorkin-módszernél egynél nagyobb értékeket is kapunk, ami azt jelenti, hogy ebben az esetben a fázissebesség nagyobb, mint az advekciós sebesség. $\varepsilon = 0,9549$ mellett kis l értékekre pedig negatív eredményt ad a módszer a legkisebb négyzetek módszerével együtt, azaz a hullám terjedésének iránya is rossz.

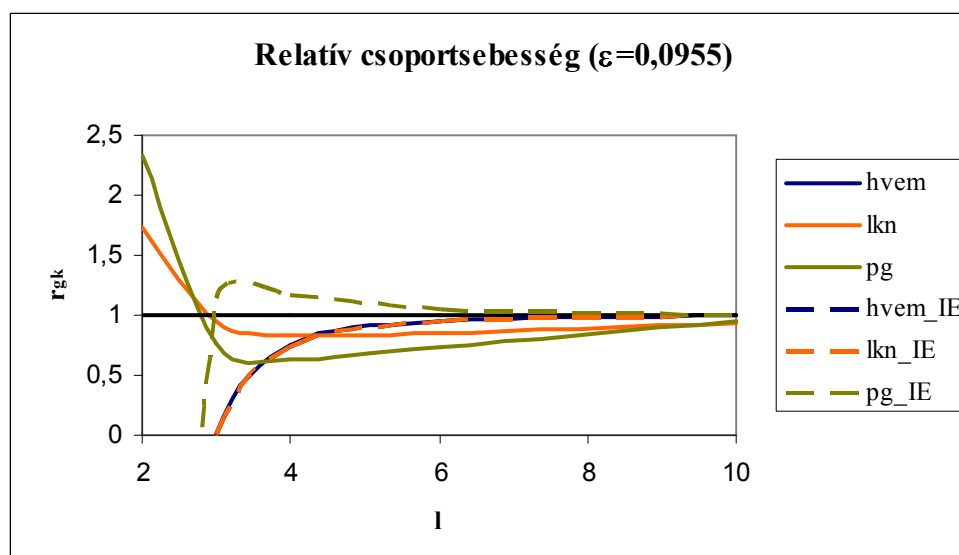


10. ábra – A hagyományos véges elem, a Petrov–Galjorkin-módszer és a legkisebb négyzetek módszerének trapézformulát, ill. implicit Euler-sémát alkalmazó változatának relatív fázissebessége l függvényében $\varepsilon=0,0955$ mellett

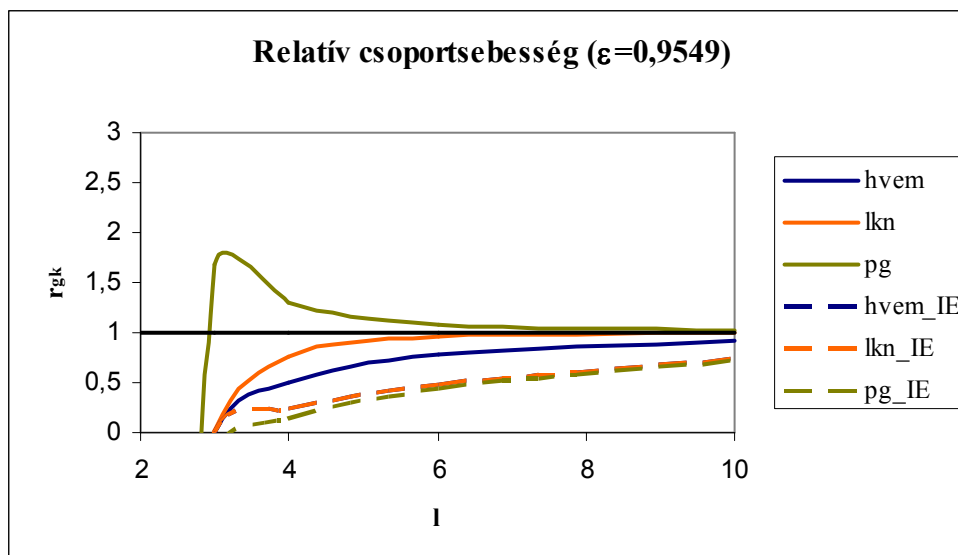


11. ábra – A hagyományos véges elem, a Petrov–Galjorkin-módszer és a legkisebb négyzetek módszerének trapézformulát, ill. implicit Euler-sémát alkalmazó változatának relatív fázissebessége l függvényében $\varepsilon=0,9549$ mellett

A relatív csoportsebességeknél nagy negatív értékeket is kaptunk kis l -ekre, ezért az ábrán az r_{gk} értékei csak 0-tól vannak feltüntetve, hogy a nagyobb l értékeknél jobban látszódjon az eredmény. Mindkét Courant-szám esetén a legkisebb négyzetek módszere - implicit Euler-módszernél a hagyományos véges elemes módszerrel együtt - adta a legjobb eredményeket. Nagyobb ε -ra a trapézformulás, míg kisebbre az implicit Euler-sémát alkalmazó módszerek a jobbak.



12. ábra – A hagyományos véges elem, a Petrov–Galjorkin-módszer és a legkisebb négyzetek módszerének trapézformulát, ill. implicit Euler-sémát alkalmazó változatának relatív csoportsebessége l függvényében $\varepsilon=0,0955$ mellett



13. ábra – A hagyományos véges elem, a Petrov–Galjorkin-módszer és a legkisebb négyzetek módszerének trapézformulát, ill. implicit Euler-sémát alkalmazó változatának relatív csoportsebessége l függvényében $\varepsilon=0,9549$ mellett

Ha a fázis- és csoportsebesség nem egyezik meg, diszperziót kapunk a megoldásban. Az 5. és 6. táblázat a relatív fázis- és csoportsebességeket tartalmazza l függvényében a három VEM esetén $\theta = 0,5$ mellett. Minél nagyobb a különbség a két sebesség között, annál erősebb a diszperzió. Az 5. táblázatbeli értékek $\varepsilon = 0,9549$,

hvem	l	2	3	4	6	10
	r_k	0	0,6910	0,8287	0,9212	0,9707
	r_{gk}	-3	0	0,4957	0,7703	0,9133
lkn	l	2	3	4	6	10
	r_k	0	-0,6760	0,9024	0,9389	0,9734
	r_{gk}	-3,0844	0,0274	0,7551	0,9591	0,9946
pg	l	2	3	4	6	10
	r_k	0	-0,7217	0,9066	1,0284	1,1619
	r_{gk}	-12,7871	1,6897	1,3027	1,0718	1,0113

5. Táblázat – A hagyományos véges elem, a Petrov–Galjorkin-módszer és a legkisebb négyzetek módszerének relatív fázis- és csoportsebessége $\varepsilon=0,9549$ és $\theta=0,5$ esetén

míg a 6. táblázatbeliek $\varepsilon = 0,0955$ mellett érvényesek. Látható, hogy kis l értékek mellett erős a diszperzió, de a hullámhosszak növekedésével csökken mindhárom módszernél. A legkisebb diszperziókat a legkisebb négyzetek módszerénél (lkn), míg a legerősebbet a Petrov–Galjorkin-módszernél (pg) tapasztalhatjuk. $\varepsilon = 0,0955$ esetén a legkisebb négyzetek módszere mellett a hagyományos VEM (hvem) is kis diszperziót ad.

hvem	1	2	3	4	6	10
	r_k	0	0,8251	0,9533	0,9916	0,9988
	r_{gk}	-3	0	0,7462	0,9576	0,9945
lkn	1	2	3	4	6	10
	r_k	0	0,8307	0,9549	0,9918	0,9988
	r_{gk}	1,7284	0,9529	0,8254	0,8517	0,9261
pg	1	2	3	4	6	10
	r_k	0	1,1749	1,0674	1,0163	1,0039
	r_{gk}	2,3277	0,7612	0,6350	0,7251	0,9545

6. Táblázat – A hagyományos véges elem, a Petrov–Galjorkin-módszer és a legkisebb négyzetek módszerének relatív fázis- és csoportsebessége $\varepsilon=0,0955$ és $\theta=0,5$ esetén

A 7. táblázat az implicit Euler-sémához tartozó relatív fázis- és csoportsebességeket tartalmazza. Itt a nagyobb Courant-szám mellett sokkal erősebb diszperziókat kapunk, mint a trapézformulás esetben. Egy eddig még nem tapasztalt eredményt is leolvashatunk a táblázatból, mégpedig azt, hogy $\varepsilon = 0,0955$ mellett a Petrov–Galjorkin-módszer adja a legkisebb diszperziót.

$\varepsilon=0,9549$	hvem lkn	1	2	3	4	6	10	
		r_k	0	0,5135	0,6409	0,7816	0,9000	
		r_{gk}	-3	0	0,2458	0,4837	0,7322	
	pg	1	2	3	4	6	10	
		r_k	0	0,4574	0,6113	0,7712	0,8979	
		r_{gk}	-1,7832	-0,1755	0,1469	0,4439	0,7230	
	$\varepsilon=0,0955$	hvem lkn	1	2	3	4	6	10
			r_k	0	0,8196	0,9485	0,9892	0,9979
			r_{gk}	-3	0	0,7349	0,9506	0,9918
pg		1	2	3	4	6	10	
		r_k	0	1,0752	1,0371	1,0078	1,0004	
		r_{gk}	-7,6289	1,1310	1,1601	1,0417	1,0044	

7. Táblázat – A hagyományos véges elem, a Petrov–Galjorkin-módszer és a legkisebb négyzetek módszerének relatív fázis- és csoportsebessége $\varepsilon=0,9549$ és $\varepsilon=0,0955$ esetén $\theta=1$ mellett

A Fourier-analízis eredményeit összefoglalva elmondható, hogy adott séma a fiktív csillapodását, relatív sebességeit vagy diszperzióját nézve a nagyobb hullámhosszakra és kis Courant-számokra a jobb. A legkisebb csillapodást a hagyományos véges elemes módszer adja, míg a legpontosabb fázis- és csoportsebességeket a legkisebb négyzetek módszere, természetesen adott esetben a hagyományos véges elemes módszerrel együtt. A

diszperzió a hagyományos véges elemes módszernél, ill. a legkisebb négyzetek módszerénél a legkisebb, kivéve implicit Euler-formula és kis Courant-szám mellett, ahol a Petrov–Galjorkin-séma a legjobb. Az idődiszkretizációs séma megválasztásánál ajánlott figyelembe venni, hogy a trapézformulás sémák kisebb csillapodást okoznak, illetve azt is, hogy ha ezekkel számolunk, akkor nagyobb Courant-számoknál pontosabb fázis- és csoportsebességet kapunk.

7. A kétdimenziós advekciós feladat

Az egydimenziós véges elemes módszerekkel akár kétdimenziós feladatokat is megoldhatunk operátorszeletelés segítségével. Ilyenfajta dimenziós szeletelést sikerrel alkalmazott például Hunsdorfer és Spee (1995) a kétdimenziós advekciós feladatra, bár ők a térbeli diszkretizációt véges térfogat sémával oldották meg. Cullen (1973) megmutatta, hogy a lineáris hagyományos véges elemes módszer, ha leap-frog sémát használunk az időbeli diszkretizációra, a kétdimenziós advekciós feladatra megköveteli az

$$\frac{|u_0|\Delta t}{d} \leq \frac{1}{\sqrt{6}}$$

stabilitási feltételt, ahol d a csomópontok közötti távolság. Egy dimenzióban ez a feltétel (Haltiner and Williams, 1980) gyengébb, az

$$\frac{|u_0|\Delta t}{h} \leq \frac{1}{\sqrt{3}}$$

egyenlőtlenséggel írhatjuk le. Bár mi θ -módszert használtunk az időbeli diszkretizációra, ami $\theta \geq 0,5$ esetben abszolút stabil megoldást ad a kipróbált véges elemes módszerekkel, de megvizsgálni ebben az esetben is érdemes a problémát. Másrészt több dimenzióban a 4. fejezet egyenletrendszereiben az integrálok kiszámolása sokkal bonyolultabb és számításigényesebb.

Tekintsük az alábbi kétdimenziós advekciós feladatot!

$$\frac{\partial c}{\partial t} + u_0 \frac{\partial c}{\partial x} + v_0 \frac{\partial c}{\partial y} = 0$$

$$c(x, y, 0) = \sin x + \sin y$$

$$c(0, y, t) = c(2\pi, y, t)$$

$$c(x, 0, t) = c(x, 2\pi, t),$$

ahol u_0 és v_0 a konstans advekciós sebességek. A pontos megoldás:
 $c(x, y, t) = \sin(x - u_0 t) + \sin(y - v_0 t)$.

A numerikusan megoldáshoz először tehát operátorszeletelést használunk, így a 2.2. alfejezet alapján két egyszerűbben megoldható egydimenziós részfeladathoz jutunk:

$$\frac{\partial c_{(1)}^n}{\partial t} + u_0 \frac{\partial c_{(1)}^n}{\partial x} = 0$$

$$c_{(1)}^n(x, y, (n-1)\Delta t) = c_{(2)}^{n-1}(x, y, (n-1)\Delta t) \quad n=1, \dots, N$$

és

$$\frac{\partial c_{(2)}^n}{\partial t} + v_0 \frac{\partial c_{(2)}^n}{\partial y} = 0$$

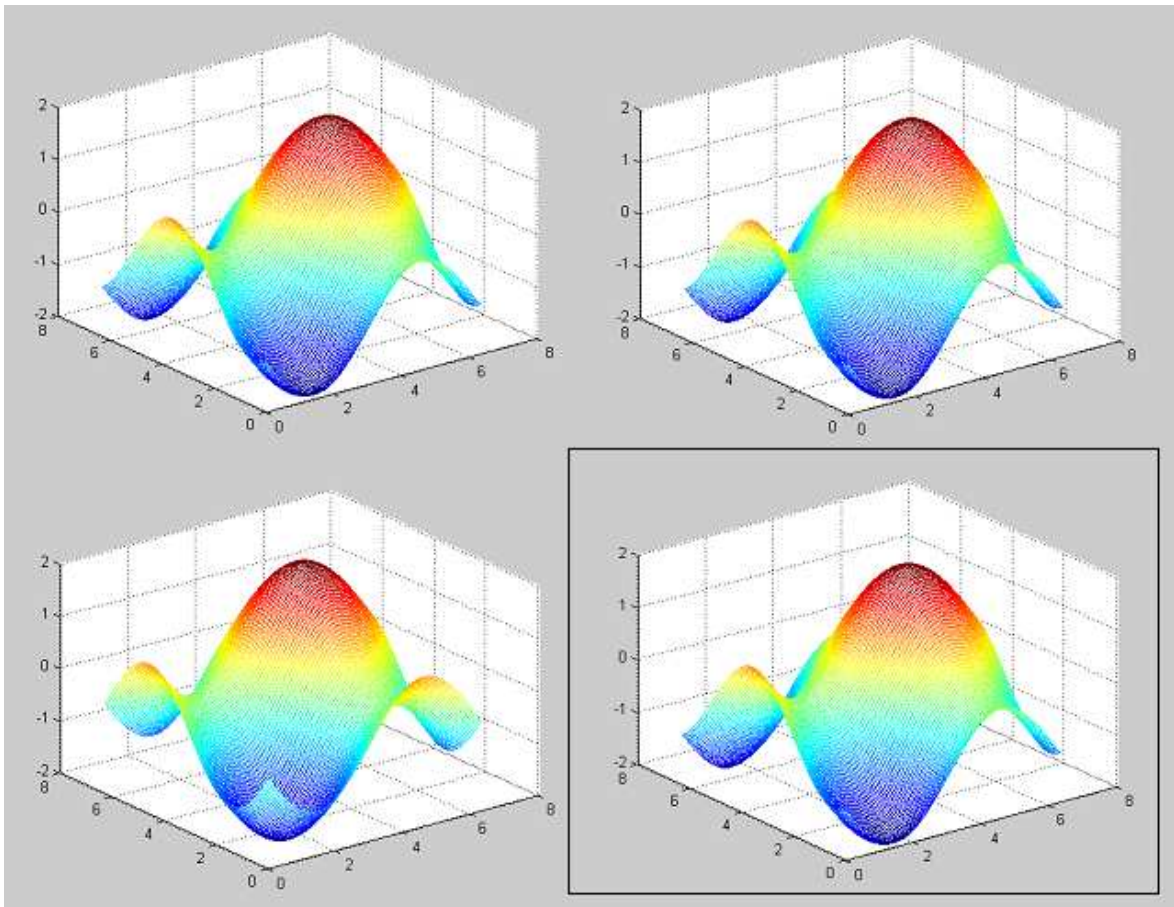
$$c_{(2)}^n(x, y, (n-1)\Delta t) = c_{(1)}^{n-1}(x, y, n\Delta t) \quad n=1, \dots, N,$$

ahol $c_{(2)}^0(x, y, 0) = \sin x + \sin y$.

Legyen $A_1 := u_0 \frac{\partial}{\partial x}$ és $A_2 := v_0 \frac{\partial}{\partial y}$. Ennél a feladatnál periodikus peremfeltételnél

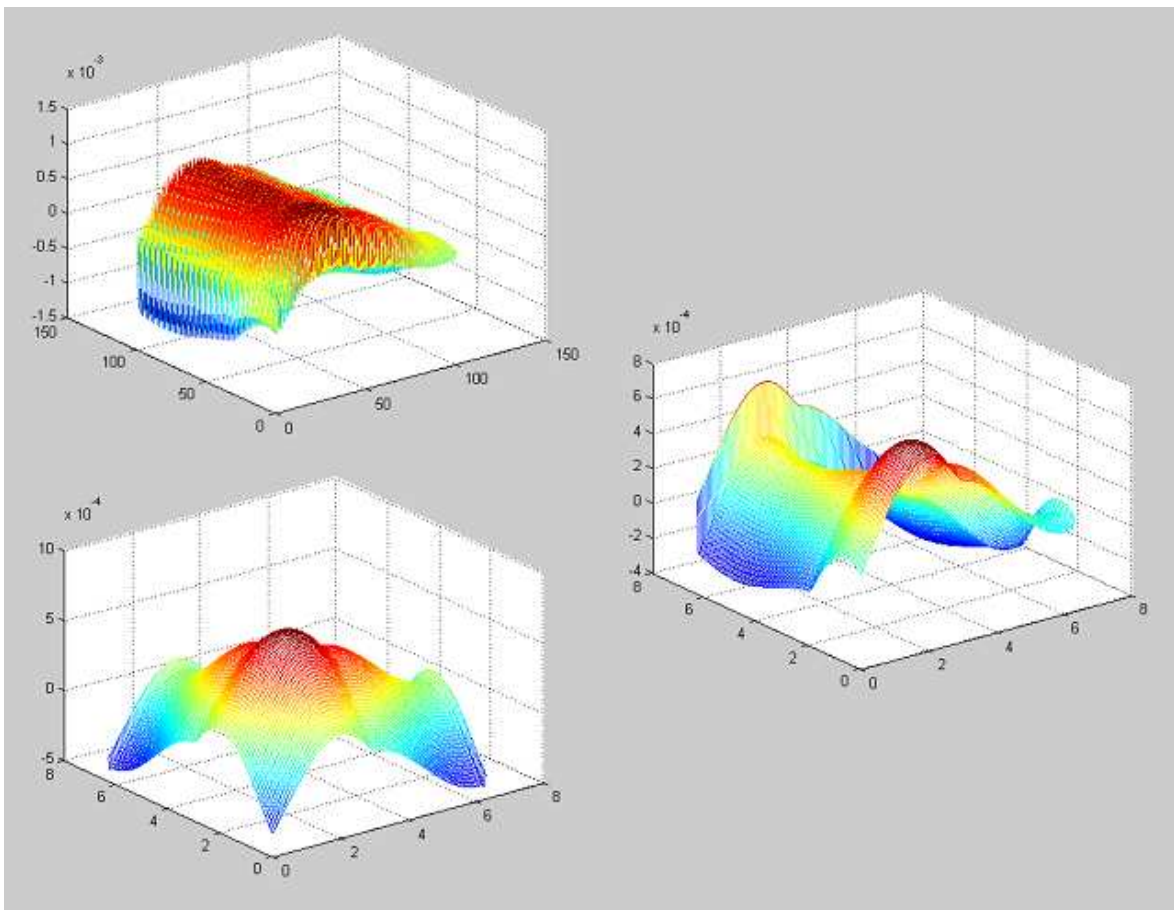
az operátorszorzat nem okoz hibát a megoldásban, mert A_1 és A_2 kommutál, hiszen

$$[A_1, A_2](c) = u_0 \frac{\partial c}{\partial x} \left(v_0 \frac{\partial c}{\partial y} \right) - v_0 \frac{\partial c}{\partial y} \left(u_0 \frac{\partial c}{\partial x} \right) = u_0 v_0 \frac{\partial^2 c}{\partial x \partial y} - v_0 u_0 \frac{\partial^2 c}{\partial y \partial x} = 0.$$



14. ábra – A hagyományos véges elemes (fent balra), a Petrov-Galjorkin (fent jobbra) módszer és legkisebb négyzetek módszerének (lent balra) számított megoldása és a pontos megoldás (lent jobbra) $T=1$ -ben

Adott időrétegen megoldjuk egymás után a két advekción feladatot, úgy, ahogy azt egy dimenzióban tettük, kezdeti feltételként mindig az utoljára kapott eredményt felhasználva. Idődiszkretizációs sémaként a kísérleteknél a trapézformulát használtuk, térben pedig kipróbáltuk mindhárom véges elemes módszert (a Petrov–Galjorkin-módszer programkódját lásd a Függelékben). Az időintervallum hossza $T=1$, a két advekción sebesség pedig $u_0=2$ és $v_0=3$ volt, így a pontos megoldás: $c(x,y,t)=\sin(x-2t)+\sin(y-3t)$. Jelölje a térbeli osztásközök számát x irányban K , y irányban J , az időbeli osztásközök számát pedig N . A 14. ábrán $N=K=J=100$ mellett $T=1$ -ben látható a három módszerrel számított numerikus és a pontos megoldás. Ránézésre csak a legkisebb négyzetek módszerének megoldása tér el a pontos megoldástól.



15. ábra – A hagyományos véges elemes (bal felső), a legkisebb négyzetek (bal alsó) és a Petrov–Galjorkin-módszer (jobb oldali) megoldásának eltérése a pontos megoldástól $T=1$ -ben

A 15. ábra a három módszer megoldásának pontos megoldásától vett eltérését szemlélteti $N=K=J=100$ esetén $T=1$ -ben. A hagyományos véges elemes módszer hibája egy nagyságrenddel nagyobb a másik két módszerénél, és oszcillál. A Petrov–

Galjorkin-módszer hibája hasonló alakú, de oszcilláció nélkül. A peremek mellett nagy gradiensek figyelhetők meg. A hibafüggvények ugyan nagy változékonyságot mutatnak, de a megoldásban ez – a hibák kis nagyságrendje miatt – nem látszódik meg, csak a legkisebb négyzetek módszerének esetén.

A módszerek hibáját adott időpillanatban a pontos és a közelítő megoldás különbségének diszkrét l_2 -beli normájával fejezzük ki az 5. fejezetben található hibával hasonló módon, de most (x,t) sík helyett (x,y) síkon számolunk. Tehát e diszkrét hibafüggvény diszkrét l_2 -beli normáját

$$\|e\|_{l^2(\Omega_T)} := \sqrt{\sum_{i=0}^K \sum_{j=0}^J (e_i^n)^2} h^2$$

képlettel számoltuk. A 8. táblázatban $K = J$ mellett a három véges elemes módszer hibája látható. Az időbeli osztópontok száma 200, és az értékek $T = 1$ -ben érvényesek. A módszerek térben másodrendűen viselkednek, és a legkisebb hibákat most is a Petrov–Galjorkin-módszer adja. A hibák $K = 128$ mellett 10^{-3} nagyságrendűek, két nagyságrenddel nagyobbak, mint ennél a felbontásnál az egydimenziós feladatnál kapott hibák, amikor trapézformulát használtunk, de kisebbek, mint mikor az implicit Euler-sémával oldottuk meg a feladatot.

K	hvem	lkn	pg
4	2,108638	2,146028	2,563549
8	0,445411	0,424743	0,463315
16	0,10849	0,083077	0,087396
32	0,026854	0,017605	0,018392
64	0,006712	0,004316	0,004209
128	0,001711	0,001129	0,001041

8. Táblázat – A három módszer hibája

8. Összefoglalás

A szakdolgozatban konstans áramlási sebességű advekción feladatokat vizsgáltunk. Három véges elemes módszert alkalmaztunk a feladatok megoldására: a hagyományos véges elemes és egy Petrov–Galjorkin-módszert, ill. a legkisebb négyzetek módszerét.

Az időbeli diszkretizációra a θ -módszert használtuk, melynek főként két speciális esetével, a trapézformulával ($\theta = 0,5$) és az implicit Euler-sémával ($\theta = 1$) foglalkoztunk. $0,5$ -nél kisebb θ -val nem érdemes számolni, mert ezekben az esetekben többnyire instabil megoldást kapunk. Megállapítottuk, hogy megfelelően megválasztott kezdeti és peremfeltételek, ill. rácsfelbontás mellett $\theta \geq 0,5$ esetén is megjelenhet oszcilláció. Ezeket a zajokat leginkább a Petrov–Galjorkin-módszer képes elnyomni.

Egydimenziós esetben a közelítő megoldás hibája a trapézformulával számolva már $N \sim 200$ és $K \sim 60$ rácsfelbontás mellett maximum 10^{-5} nagyságrendű. Ezt a pontosságot az implicit Euler-módszerrel még $N \sim K \sim 1000$ mellett sem érhetjük el. Az implicit Euler-séma átlagosan két-három nagyságrenddel rosszabb eredményeket ad, mint a trapézformula. A véges elemes módszerek időben a választott idődiszkretizációs sémának megfelelően viselkednek, térben pedig kis lépésközökre másod-, nagyobbakra elsőrendűen viselkednek, míg az idődiszkretizáció gátat nem szab a további javulásnak. A legkisebb hibákat a Petrov–Galjorkin-módszernél kapjuk.

A 6. fejezetben elvégeztük a véges elemes módszerek Fourier-analízisét. Mindhárom séma abszolút stabilnak bizonyult $\theta \geq 0,5$ mellett. Csillapodás minden hullámhosszra mindhárom módszernél jellemző, kivéve a hagyományos véges elemes módszert $l = 2$ -nél. A fázis- és csoportsebesség sem egyezik meg az advekción sebességgel egyik séma esetében sem, bármilyen hullámhosszt is veszünk. Ugyanakkor megállapítottuk, hogy nagyobb hullámhosszakra és kis Courant-számra a csillapodás jelentősen csökken, a fázis- és csoportsebesség pontosabb, és a diszperzió is kisebb. A legkisebb csillapodást a hagyományos véges elemes módszer esetén kaptuk, míg a legpontosabb fázis- és csoportsebességeket, ill. a legkisebb diszperziót a legkisebb négyzetek módszere esetén, természetesen adott esetben a hagyományos véges elemes módszerrel együtt. Az idődiszkretizációs séma megválasztásánál ajánlatos figyelembe venni, hogy a trapézformula kisebb csillapodást okoz, valamint nagyobb Courant-számoknál pontosabb fázis- és csoportsebességet eredményez. Habár a legkisebb hibákat

és zajokat a Petrov–Galjorkin-módszer adja, a hullámmódust leíró tulajdonságok ennél a sémánál a legrosszabbak.

Az egydimenziós véges elemes módszerekkel a kétdimenziós advekciós feladat is megoldható bizonyult operátorszeletelés segítségével. A legkisebb hibákat az egydimenziós feladathoz hasonlóan a Petrov–Galjorkin-módszer adta.

Függelék

1. Az egydimenziós, oldalsó peremfeltételes advekciós feladat legkisebb négyzetek módszerével való megoldására használt Matlab programkód:

```
function lknir_theta(k,N)
%Legkisebb négyzetek módszere

%du/dt+a*du/dx=0      a=konst, a>0
%u(0,t)=g(t)         t:[0,T]
%u(x,0)=u0(x)        x:[0,pi]

%g(t)=-sin(a*t)
%u0=sin(x)

a=1;
T=1;
dt=T/N;
h=pi/k;
theta=0.5;

s1=diag(4*(ones(1,k)))+diag(ones(1,k-1),1)+diag(ones(1,k-1),-1);
s1(k,k)=2;
s2=diag(zeros(1,k))+diag(0.5*ones(1,k-1),1)+diag(-0.5*ones(1,k-1),-1);
s2(k,k)=0.5;
s3=diag(zeros(1,k))+diag(-0.5*ones(1,k-1),1)+diag(0.5*ones(1,k-1),-1);
s3(k,k)=0.5;
s4=diag(2*(ones(1,k)))+diag(-1*ones(1,k-1),1)+diag(-1*ones(1,k-1),-1);
s4(k,k)=1;
S=h/6*s1+a*dt*theta*s2+a*dt*theta*s3+a*a*dt*dt*theta*theta/h*s4;

sorx=1:k;
x=sorx*h;
sort=0:N;
t=sort*dt;
x_teljes=[0 x];
sorx2=1:(2*k);
x2=sorx2*h/2;
x_teljes2=[0 x2];

%kezdeti feltétel
g=(-sin(a*t));
u0=sin(x);
u02=sin(x2);

%pontos megoldas:
for n=1:N+1;
    up(:,n)=(sin(x_teljes-a*t(n)))';
    up2(:,n)=(sin(x_teljes2-a*t(n)))';
end

f1=zeros(k,1);
f1(1)=h/6+a*dt*theta/2;
f2=zeros(k,1);
f2(1)=-a*dt*(theta-1)/2-a*a*dt*dt*theta*(theta-1)/h;
f3=zeros(k,1);
```

```

f3(1)=-a*dt*theta/2-a*a*dt*dt*theta*theta/h;

u(:,1)=[g(1); u0'];
alfa=u0';

for n=1:N;
    F=(g(n)-g(n+1))*f1+g(n)*f2-g(n+1)*f3+(h/6*s1+a*dt*(theta-1)*s2+...
        a*dt*theta*s3+a*a*dt*dt*theta*(theta-1)/h*s4)*alfa;
    alfa=inv(S)*F;
    u(:,n+1)=[g(n+1); alfa];
end

%hiba
uh=u-up;

plot(x_teljes,u(:,N), x_teljes2,up2(:,N),'--')
axis([0 pi 0 1])

end

```

2. Az egydimenziós, periodikus peremfeltételes advekción feladat Petrov–Galjorkin-módszerrel való megoldására használt Matlab programkód:

```

function pgir_ppf(k,N)
%Petrov-Galjorkin módszer

%du/dt+a*du/dx=0      a=konst, a>0
%u(x,t)=u(x+2pi,t)    t:[0,T]
%u(x,0)=u0(x)         x:[0,2pi]

%u0=sin(x)

a=1;
T=1;
dt=T/N;
h=2*pi/k;
theta=0.5;

s1=diag(zeros(1,k))+diag(0.5*ones(1,k-1),1)+diag(-0.5*ones(1,k-1),-1);
s1(k,k)=0.5;
s1(1,k)=-0.5;
s2=diag(2*(ones(1,k)))+diag(-1*ones(1,k-1),1)+diag(-1*ones(1,k-1),-1);
s2(k,k)=1;
s2(1,k)=-1;
s3=diag(4*(ones(1,k)))+diag(ones(1,k-1),1)+diag(ones(1,k-1),-1);
s3(k,k)=2;
s3(1,4)=1;
s4=diag(zeros(1,k))+diag(-1*ones(1,k-1),1)+diag(ones(1,k-1),-1);
s4(k,k)=1;
s4(1,k)=1;
S=a*dt*theta*s1+a*dt*theta/2*s2+h/6*s3+h/4*s4;

sorx=1:k;

```

```

x=sorx*h;
sort=0:N;
t=sort*dt;
x_teljes=[0 x];

%kezdeti feltétel
u0=sin(x);

%pontos megoldas:
for n=1:N+1;
    up(:,n)=(sin(x_teljes-a*t(n)))';
end

u(:,1)=(sin(x_teljes))';
alfa=u0';

for n=1:N;
    F=(-a*dt*(1-theta)*s1-a*dt*(1-theta)/2*s2+h/6*s3+h/4*s4)*alfa;
    alfa=inv(S)*F;
    u(:,n+1)=[alfa(k); alfa];
end

%hiba
uh=u-up;

plot(x_teljes,u(:,N), x_teljes,up(:,N),'--')
axis([0 2*pi 0 1])

end

```

3. A kétdimenziós, periodikus peremfeltételes advekción feladat hagyományos véges elemes módszerrel való megoldására használt Matlab programkód:

```

function hvem_2D_ppf(K,J,N)
%Hagyományos véges elemes módszer

%dc/dt+u*dc/dx+v*dc/dy=0      u,v=konst, u,v>0
%c(x,y,t)=c(x+2*pi,y,t)      t:[0,T]
%c(x,y,t)=c(x,y+2*pi,t)      x:[0,X]
%c(x,y,0)=c0(x,y)=sin(x)+sin(y)  y:[0,Y]

u=2;
v=3;
T=1;
X=2*pi;
Y=2*pi;
dt=T/N;
hx=X/K;
hy=Y/J;
theta=0.5;

s1=diag(zeros(1,K))+diag(0.5*ones(1,K-1),1)+diag(-0.5*ones(1,K-1),-1);
s1(K,K)=0.5;

```

```

s1(1,K)=-0.5;
s2=diag(4*(ones(1,K)))+diag(ones(1,K-1),1)+diag(ones(1,K-1),-1);
s2(K,K)=2;
s2(1,K)=1;
S1=u*dt*theta*s1+hx/6*s2;

s3=diag(zeros(1,J))+diag(0.5*ones(1,J-1),1)+diag(-0.5*ones(1,J-1),-1);
s3(J,J)=0.5;
s3(1,J)=-0.5;
s4=diag(4*(ones(1,J)))+diag(ones(1,J-1),1)+diag(ones(1,J-1),-1);
s4(J,J)=2;
s4(1,J)=1;
S2=v*dt*theta*s3+hy/6*s4;

sorx=1:K;
x=sorx*hx;
sory=1:J;
y=sory*hy;
sort=0:N;
t=sort*dt;
x_teljes=[0 x];
y_teljes=[0 y];

%pontos megoldas:
for j=1:J+1
cp(:,j)=(sin(x_teljes-u*t(N+1))+sin(y_teljes(j)-v*t(N+1)));
end

%kezdeti feltétel
for i=1:K+1;
c(i,:)=sin(x_teljes(i))+sin(y_teljes);
end

for n=1:N;
for j=1:J;
F=(hx/6*s2-u*dt*(1-theta)*s1)*c(2:K+1,j+1);
alfa=inv(S1)*F;
c(:,j+1)=[alfa(K); alfa];
end
c(:,1)=c(:,J+1);
for i=1:K;
F=(hy/6*s4-v*dt*(1-theta)*s3)*c(i+1,2:J+1)';
alfa=inv(S2)*F;
c(i+1,:)=[alfa(J); alfa]';
end
c(1,:)=c(K+1,:);
end

%hiba
ch=c-cp;

%max hiba
b=max(max(abs(c-cp)))

mesh(x_teljes,y_teljes,cp)

end

```


Irodalomjegyzék

Chock, D. P., 1985: A comparison of numerical methods for solving advection equations – II. *Atmos. Environ.*, 19, 571-586.

Chock, D. P., 1991: A comparison of numerical methods for solving advection equations – III. *Atmos. Environ.*, 25A, 853-871.

Codina, R., 1998: Comparison of finite element methods for solving the diffusion-convection-reaction equation. *Comput. Methods Appl. Mech. Engrg.*, 156, 185-210.

Côté, J. and Staniforth, A., 1990: An Accurate and Efficient Finite-Element Global Model of the Shallow-Water Equations. *Mon. Wea. Rev.*, 118, 2702-2717.

Côté, J., Gravel, S., Méthot, A., Patoine, A., Roch, M. and Staniforth, A., 1998: The operational CMC-MRB Global Environmental Multiscale (GEM) model: Part I. Design considerations and formulation. *Mon. Wea. Rev.*, 126, 1373–1395.

Csomós, P., Faragó, I., and Havasi, Á., 2005: Weighted Sequential Splittings and Their Analysis. *Comput. Math. Appl.*, 50, 1017-1031.

Cullen, M. J. P., 1973: A simple finite element method for meteorological problems. *J. Inst. Math. Applics.*, 11, 15-31.

Ern, A. and Guermond, J.-L., 2004: *Theory and Practice of Finite Elements*. Springer, New York.

Haltiner G., J. and Williams, R., T., 1980: *Numerical Prediction and Dynamic Meteorology*. John Wiley & Sons Inc., New York.

Havasi, Á., Bartholy, J., and Faragó, I., 2001: Splitting method and its application in air pollution modeling. *Időjárás*, 105, 39-58.

Hunsdorfer, W. and Spee, E. J., 1995: An Efficient Horizontal Advection Scheme for the Modeling of Global Transport of Constituents. *Mon. Wea. Rev.*, 123, 3554-3564.

Hundsorfer, W. and Verwer, J. G., 2003: *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer, Berlin.

Kan, J. van, Vermolen, F. and Segal, A., 2005: *Numerical Methods in Scientific Computing*. VSSD, Delft.

Kreiss, O. and Olinger, J., 1972: Comparison of accurate methods for the integration of hyperbolic equations. *Tellus*, 24, 199-215.

Long, P. E. Jr and Pepper, D. W., 1981: An examination of some simple numerical schemes for calculating scalar advection. *J. Appl. Math.*, 20, 146-156.

- Mailhot, J., Sarrazin, R., Bilodeau, B., Brunet, N. and Pellerin, G., 1997: Development of the 35-km version of the operational regional forecast system. *Atmos.–Ocean*, 35, 1–28.
- Neta, B., 1992: Analysis of finite elements and finite differences for shallow water equations: A review. *Mathematics and Computers in Simulation*, 34, 141-161.
- Neta, B. and Williams, R. T., 1986: Stability and phase speed for various finite element formulations of the advection equations. *Computers and Fluids*, 34, 393-410.
- Orszag, S. A., 1971: Numerical simulation of incompressible flows within simple boundaries – I. *J. Fluid. Mech.*, 49, 75-112.
- Orszag, S. A., 1972: Comparison of pseudospectral and spectral approximations. *Stud. Appl. Math.*, 51, 253-259.
- Pepper, D. W., Kern, C. D. Long and P. E Jr, 1979: Modelling the dispersion of atmospheric pollution using cubic splines and chapeau functions. *Atmos. Environ.*, 13, 223-237.
- Riddaway, R. W., 1986: An Introduction to Numerical Methods for Weather Prediction. *Internal Report from ECMWF, Meteorological Training Course, Lecture Note No.1.4*.
- Richtmyer, R. D. and Morton, K. W., 1967: *Difference methods for initial value problems*. Interscience Publishers, New York.
- Sklarz, M. A., Spielvogel, L. Q. and Loomis, H. G., 1979: Numerical Simulation of the 29 November 1975 Island of Hawaii Tsunami by the Finite-Element Method. *Journal of Physical Oceanography*, 9, 1022-1031.
- Smiths, G. D., 1978: *Numerical solution of partial differential equations: Finite differences*. Oxford University Press, Oxford.
- Stanners, D. and Bourdeau, P., 1995: *Europe's Environment - The Dobris Assessment*. EEA (European Environment Agency)
- Strang, G., 1963: Accurate partial difference methods I: Linear Cauchy problems. *Arch. Rational Mech. Anal.*, 12, 392-402.
- Strang, G., 1968: On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.*, 5, 506-517.
- Untch, A. and Hortal, M., 2003: *A Finite-Element Scheme for the Vertical Discretization in the Semi-Lagrangian Version of the ECMWF Forecast Model*. ECMWF, Reading.
- Zlatev, Z., 1995: *Computer Treatment of Large Air Pollution Models*. Kluwer Academic Publishers.

Köszönetnyilvánítás

Ezúton szeretném megköszönni az útmutatást és a sok segítséget témavezetőmnek, Havasi Ágnesnek, és az értékes javaslatokat Faragó Istvánnak és Izsák Ferencnek.