Kvalitatív tulajdonságokat megőrző diszkretizációk

Lóczi Lajos
LLoczi@inf.elte.hu
ELTE IK, Numerikus Analízis Tanszék

TKP-workshop
2021. június 10.

NATIONAL RESEARCH, D
AND INNOVATION
HUNGAR

I. Időben megmaradó mennyiségek, konzervatív rendszerek

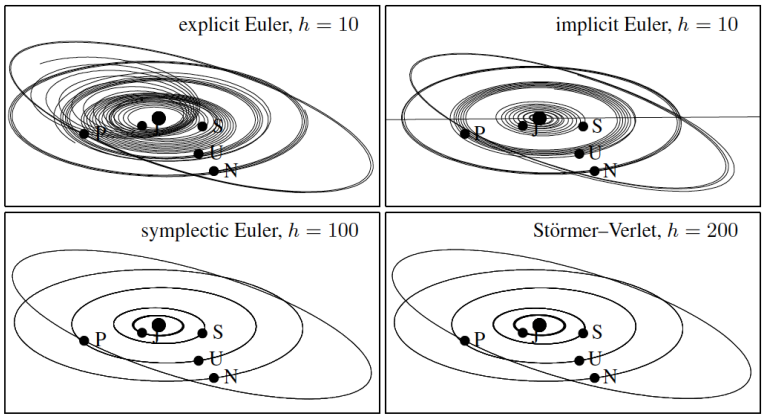


Fig. 2.4. Solutions of the outer solar system

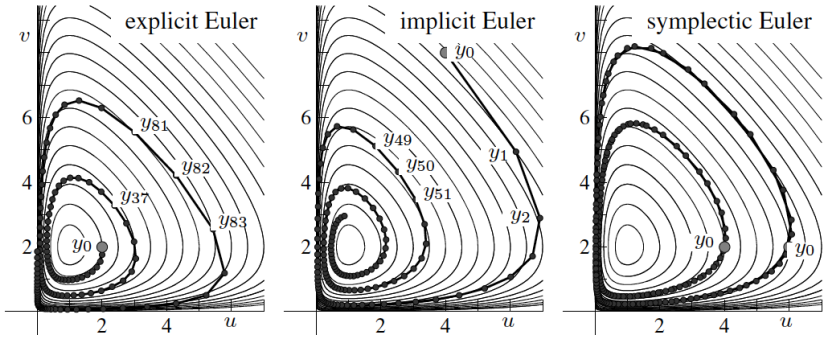


Fig. 1.2. Solutions of the Lotka–Volterra equations (1.1) (step sizes $h = 0.12$; initial values $(2, 2)$ for the explicit Euler method, $(4, 8)$ for the implicit Euler method, $(4, 2)$ and $(6, 2)$ for the symplectic Euler method)

• periodikus megoldások, zárt trajektóriák:

testek mozgása, mozgáspályák leírása (harmonikus oszcillátor, Kepler-egyenletek, energiamegmaradás, invariánsok),

járványterjedési modell, ragadozó-zsákmány modell (első integrál),

kémiai koncentrációváltozások

• *kvalitatív* szempontból “rossz” numerikus módszerek: explicit Euler, implicit Euler
• “jó” módszerek, pl. szimplektikus Euler-módszer – *structure-preserving numerical methods*, *geometric numerical integration*
• noha *kvantitatív* szempontból mindhárom módszer *ugyanúgy* viselkedik (elsőrendben pontos módszerek)
• a numerikus módszerek klasszikus hibaanalízise tehát ezekben a szituációkban nem elégséges

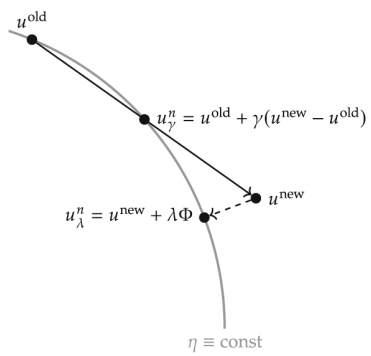
Időben csökkenő (nemnövekvő) mennyiségek, disszipatív rendszerek

- pl. súrlódásos mozgás
- entrópia

Az előbbi, közönséges differenciálegyenlet-rendszereket tartalmazó példák megoldásait elsőrendben pontos (= elsőrendű) numerikus módszerekkel közelítettük

Hogyan lehet *magasabb rendű* (= *kvantitatív* szempontból *jobb*, “pontosabb”, “hatékonyabb”) módszereket konstruálni, amelyek *kvalitatív szempontból* is megfelelőek?

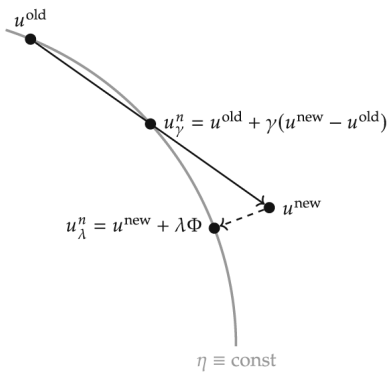
Egy megközelítés: adott numerikus módszerből kiindulva projekciós technikát alkalmazva:



- itt a numerikus módszer lépésközét *nem* változtatjuk

meg, hanem a sokaságra vetítünk

Egy másik megközelítés:
adott numerikus módszerből kiindulva **relaxációs** technikát alkalmazva



- itt a numerikus módszer lépésközét változtatjuk meg úgy, hogy a sokaságon maradjunk

A relaxációs módszertant Runge–Kutta-módszerekre (= az egylépéses numerikus módszerek gyakorlatban fontos családja) szisztematikusan kidolgozták és elemezték:

> Ketcheson D. I.: Relaxation Runge–Kutta methods: conservation and stability for inner-product norms. SIAM J. Numer. Anal. **57** (6), 2850–2870 (2019)

Kiindulási RK-módszer ⟶ relaxáció ⟶ módosított RK-módszer (már “jó” tulajdonságú)

Kiindulási RK-módszer ⟶ relaxáció ⟶ módosított RK-módszer (már “jó” tulajdonságú)

A relaxációval a kvalitatív tulajdonságokat ugyan megjavítottuk, de kérdés: a módosítás hatására a kvantitatív tulajdonságok nem romlanak el?

A fenti cikk megmutatja, hogy szerencsére *nem* (a módszer pontossága nem romlik).

Természetes kérdés: mi a helyzet más numerikus módszerekkel?
Egy *többlépéses* módszer kvantitatívan “jó”, de kvalitatív szempontból nem feltétlenül.

Kiterjeszthető-e a relaxációs módszertan *többlépéses* numerikus módszerekre is?

- lineáris többlépéses módszerek
- állandó lépésközű módszerek
- változó lépésközű (= adaptív) módszerek

Az erre vonatkozó eredményeinket az alábbi (D1-es besorolású) folyóiratban publikáltuk:

> Ranocha H., Lóczi L., Ketcheson D. I.: General relaxation methods for initial-value problems with application to multistep schemes. Numerische Mathematik **146**, 875–906 (2020)

Kiindulási többlépéses módszer ⟶ relaxáció ⟶ módosított többlépéses módszer

A módosított módszer kvantitatív és kvalitatív szempontból is “jó” tulajdonságú

- az eredeti numerikus módszer rendje megőrződik
- az eredeti numerikus módszer SSP-tulajdonsága (strong-stability preserving property) megőrződik
- a módosított módszer megőrzi a geometriai invariánst

Példák:

- Adams-, Nyström-, BDF-, eBDF-, vagy változó lépésközű SSP-módszerek

Kiindulási többlépéses módszer $\longrightarrow$ relaxáció $\longrightarrow$ módosított többlépéses-módszer

Hogyan történik a módosítás?

Iterációs lépésenként módosítjuk az eredeti numerikus módszert: lépésenként mindössze egy, nemlineáris skaláris egyenletet kell megoldani

- ez az extra számításigény elhanyagolható a fő iteráció számításigényéhez képest

$\Longrightarrow$ a relaxációs módszertan hatékony módosított módszereket eredményez.

II. Pozitív rendszerek

- egy másik fontos kvalitatív tulajdonság
- ez is megfogalmazható invarianciatulajdonságként

- ha az **input > 0**, akkor az **output > 0**
- ha a kezdeti feltétel függvénye **> 0**, akkor a differenciálegyenlet megoldása **∀t esetén > 0**

A numerikus módszerek megőrzik ezt a kvalitatív tulajdonságot? Általában nem

Parciális differenciálegyenletek, gyakran kétlépcsős diszkretizáció:

PDE térváltozó, időváltozó ⟶ térbeli szemidiszkretizáció ⟶ KDE-rendszer ⟶ idődiszkretizáció ⟶ teljesen diszkretizált rendszer

A kvalitatív tulajdonságok megőrződését mindkét fázisban meg lehet követelni, illetve mindkét fázisban megsérülhetnek.

Elmúlt néhány évtized: pozitivitási/oszcillációmentességi/kontraktivitási/monotonitási tulajdonságokat megőrző térbeli szemidiszkretizációk kifejlesztése

Elmúlt 2-3 évtized: SSP-elmélet

- hogyan lehet “jó” térbeli szemidiszkretizációkhoz magasabb rendű “jó” idődiszkretizációkat konstruálni

Ezeket összepárosítva kvalitatív szempontból megfelelő (pl. pozitivitásőrző) teljes diszkretizációkat nyerhetünk

Példa: 1D advekciós egyenlet, $\partial_t U(x, t) = \partial_x U(x, t)$

- ezek nemlineáris változatait használják többek között közlekedési modellekben (traffic flow models)

- pozitivitás: a járművek sűrűsége csak pozitív lehet

Példa: az advekciós egyenlet térbeli szemidiszkretizációi

- $\begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & -1 & 1 \\ 1 & 0 & \cdots & 0 & 0 & -1 \end{pmatrix}$ elsőrendben pontos és pozitivitástartó szemidiszkretizáció

- $\begin{pmatrix} 0 & \frac{1}{2} & 0 & \cdots & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} & \cdots & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \cdots & 0 & -\frac{1}{2} & 0 \end{pmatrix}$ másodrendben pontos centrális séma, ám sajnos *nem* pozitivitástartó szemidiszkretizáció, oka: a főátlón kívüli elemek előjele

- magasabb rendben pontos centrális sémák (4-ed, illetve 6-odrendű): kvantitatív szempontból jobbak, de ezek *sem* pozitivitástartók

- Fourier spektrális módszerek (az előzőek “határesete”, a térbeli diszkretizációs alappontok a teljes intervallumot kitöltik): szintén *nem* pozitivitástartók

Az advekciós egyenlet teljes diszkretizációi

Nem meglepő:
ha a 2-od, 4-ed, 6-odrendű centrális, illetve a spektrális szemidiszkretizációk után nyert KDE-rendszerre az explicit Euler idődiszkretizációt alkalmazzuk (akármilyen kicsi vagy nagy lépésközzel), akkor a teljes diszkretizált sem pozitivitástartó
$\longrightarrow$ az SSP-elmélet nem alkalmazható

Meglepő: a 2-od, 4-ed, 6-odrendű centrális, illetve a spektrális szemidiszkretizációk után nyert KDE-rendszerre az *implicit* Euler idődiszkretizációt alkalmazva a teljes diszkretizált már *bizonyos esetekben* pozitivitástartó

Bizonyos esetekben:

- elegendően **nagy** CFL-szám ($\frac{\Delta t}{\Delta x}$) esetén
- ha a térbeli szemidiszkretizáció alappontjainak száma **páratlan**

Mi történik tehát az explicit Euler és az implicit Euler módszer “között”? ⟶ θ-módszerek

Hadjimichael Y., Ketcheson D. I., Lóczi L.: Positivity preservation of implicit discretizations of the
advection equation (bírálat alatt, Q1-es folyóirat)

• általános eredmény megfogalmazása, amely beazonosítja, hogy a térbeli szemidiszkretizáció mely spektrális tulajdonsága elégséges ahhoz, hogy az advekciós egyenlet θ-módszerrel történő teljes diszkretizáltja pozitivitásőrző legyen

• paraméterek:

m (a teljes diszkretizáció mátrixának mérete)

ν (a CFL-szám)

θ (az idődiszkretizáció paramétere)

• $m \times m$-es, kétparaméteres mátrixsereg elemenkénti nemnegativitása:

$$\frac{1}{m}\sum_{\ell=1}^{m}\frac{\left(1-\theta(1-\theta)\nu^2\psi^2(\xi_\ell)\right)\cos((j-1)\xi_\ell)+\nu\psi(\xi_\ell)\sin((j-1)\xi_\ell)}{1+\theta^2\nu^2\psi^2(\xi_\ell)},$$

$$\psi(x) := 2\sum_{k=1}^{(N-1)/2} C_k \sin(kx)$$

• a fenti trigonometrikus összegek közötti rekurziók feltárása (Wolfram *Mathematica*), aszimptotikus eredmények