

**UNIVERZITA PAVLA JOZEFA ŠAFÁRIKA V  
KOŠICIACH**

**Prírodovedecká fakulta  
ÚSTAV FYZIKÁLNYCH VIED**



**Počítačová fyzika I**

**Milan Žukovič**

Košice 2015

# Počítačová fyzika I.

Autor: doc. RNDr. Milan Žukovič PhD.,  
Katedra teoretickej fyziky a astrofyziky, Prírodovedecká fakulta,  
Univerzita Pavla Jozefa Šafárika v Košiciach



© 2015 doc. RNDr. Milan Žukovič, PhD.

Recenzenti:

prof. Ing. Dušan Krokavec, CSc.,  
Katedra kybernetiky a umelej inteligencie,  
Fakulta elektrotechniky a informatiky, Technická univerzita v Košiciach,

doc. RNDr. Denis Horváth, PhD.,  
Centrum Interdisciplinárnych Biovied, Ústav fyzikálnych vied,  
Prírodovedecká fakulta, UPJŠ v Košiciach

Všetky práva vyhradené. Toto dielo ani jeho žiadnu časť nemožno reprodukovat', ukladať do informačných systémov alebo inak rozširovať bez súhlasu majiteľov práv. Za odbornú a jazykovú stránku tohto vysokoškolského učebného textu zodpovedá autor. Rukopis neprešiel redakčnou ani jazykovou úpravou.

Elektronický vysokoškolský učebný text pre Prírodovedeckú fakultu UPJŠ v Košiciach.

Vydavateľ: Univerzita Pavla Jozefa Šafárika v Košiciach

Umiestnenie: <http://unibook.upjs.sk/predaj-vydanych-titulov/prirodovedecka-fakulta>

Dostupné od: 03.02.2015

ISBN 978-80-8152-228-4

# Zoznam obrázkov

2.1	Vodorovný pohyb telesa upevneného k stene pomocou ideálnej pružiny so zanedbaním trenia. . . . .	21
2.2	Jednoduché kyvadlo. . . . .	22
2.3	Časový vývoj pohybu kyvadla pre počiatočné hodnoty $(\theta_0, \omega_0)$ : (a) $(1/10\pi; 0)$ , (b) $(9/10\pi; 0)$ , a (c) $(0; 2)$ . Plné čiary reprezentujú hodnoty $\theta$ a bodkované hodnoty $\omega$ . . . . .	24
2.4	Grafické znázornenie časovej zmeny stavu mikrobov $dx/dt$ ako funkcie ich stavu $x$ . Stav $\tilde{x} = 0$ predstavuje nestabilný a $\tilde{x} = b/p$ stabilný rovnovážny stav. . . . .	27
2.5	(a) Časová zmena populácií vlka (plná krivka) a zajaca (bodko-čiarkovaná krivka). (b) Fázový diagram modelu korisť-dravec, kde $x$ -ová os predstavuje populáciu dravca (vlkov) a $y$ -ová os stav populácie koristi (zajacov). . . . .	28
3.1	Grafické znázornenie Eulerovej metódy. . . . .	33
3.2	Geometrická interpretácia Heunovej metódy. . . . .	39
4.1	Diskretizácia priestorových súradníc v 2-rozmernom priestore. . . . .	53
4.2	Diskretizácia v priestore a čase pre riešenie rovnice difúzie v 1-rozmernom priestore. Počiatočné hodnoty funkcie majú vplyvom difúzie tendenciu s časom sa postupne znižovať a zaberáť stále širší interval pozdĺž osi $x$ . . . . .	54
4.3	Diskretizačná schéma pre FTCS metódu. . . . .	56
4.4	Diskretizačná schéma pre CTCS metódu. . . . .	58
4.5	Diskretizačná schéma pre BTCS metódu. . . . .	60
5.1	Schématicky zobrazený stav perkolácie na postupne sa zaplňajúcej štvorcovej mriežke keď je vytvorená cesta (hrubá lomená čiara) spájajúca najbližších susedov na mriežke preklenujúca celú mriežku. . . . .	66
5.2	Schématicky znázornený vzťah medzi teóriou, experimentom, simuláciou a realitou. . . . .	68
5.3	Uniformné rozdelenie pravdepodobnosti na intervale $[a, b]$ . . . . .	70

## Zoznam obrázkov

---

5.4	Zobrazenie spôsobu dosiahnutia nerovnomerného rozdelenia transformáciou premennej z kumulatívneho rozdelenia $P(x)$ . . . . .	75
5.5	25 000 náhodne generovaných bodov $(x, y) \in [0, 1]$ , pričom body pod (nad) krivkou (ne)spĺňajú podmienku $x^2 + y^2 \leq 1$ . . . . .	76
5.6	(a) Potenciál $V(x)/k_B = (x - 0.5)^2$ a (b) Boltzmannov faktor pre dve hodnoty teploty $T = 0, 1\text{K}$ a $0, 001\text{K}$ . . . . .	78
5.7	Dvojrozmerný systém so štyrmi atómami s diskretizovanými stavmi v piatich bodoch v každom smere. Počet všetkých kombinácií možných stavov atómov je $5^8$ . . . . .	79
5.8	Spinová konfigurácia dvojrozmerného Isingovho modelu na mriežke veľkosti $4 \times 4$ . . . . .	83

# Obsah

<b>O autorovi</b>	<b>8</b>
<b>Predslov</b>	<b>9</b>
<b>1 Počítačové riešenie fyzikálnych úloh</b>	<b>10</b>
1.1 Príprava úlohy . . . . .	10
1.2 Numerické riešenie . . . . .	11
1.3 Chyby počítačového riešenia . . . . .	12
1.3.1 Zaokrúhľovacie chyby a počítačová aritmetika . . . . .	13
1.3.2 Definície chýb a ich šírenie . . . . .	13
1.3.3 Obrátená úloha teórie chýb . . . . .	16
<b>2 Úvod do dynamických systémov</b>	<b>18</b>
2.1 Dynamické systémy . . . . .	18
2.2 Príklady dynamických systémov . . . . .	20
2.2.1 Mechanický oscilátor . . . . .	20
2.2.2 Jednoduché kývadlo . . . . .	22
2.2.3 Množenie populácie baktérií . . . . .	25
2.2.4 Dravec a korisť . . . . .	26
<b>3 Riešenie sústav obyčajných diferenciálnych rovníc</b>	<b>29</b>
3.1 Úvod . . . . .	29
3.2 Počiatočné úlohy pre obyčajné diferenciálne rovnice . . . . .	29
3.2.1 Podmienenosť . . . . .	31
3.3 Numerické metódy riešenia . . . . .	32
3.3.1 Eulerova metóda . . . . .	33
3.3.2 Chyby numerického riešenia . . . . .	34
3.4 Jednokrokové metódy . . . . .	38
3.4.1 Metóda Taylorovho typu . . . . .	38
3.4.2 Metódy typu Runge-Kutta (RK) . . . . .	38

## Obsah

---

3.5	Viacukrokové metódy . . . . .	40
3.5.1	Všeobecná lineárna $k$ -kroková metóda . . . . .	40
3.5.2	Metódy založené na numerickej kvadrature . . . . .	42
3.6	Okrajové úlohy pre obyčajné diferenciálne rovnice . . . . .	45
3.7	Numerické metódy riešenia . . . . .	45
3.7.1	Metóda sietí . . . . .	45
<b>4</b>	<b>Riešenie parciálnych diferenciálnych rovníc</b>	<b>48</b>
4.1	Úvod . . . . .	48
4.1.1	Počiatkové a hraničné podmienky . . . . .	49
4.1.2	Diferenčné metódy . . . . .	49
4.1.3	Konzistencia . . . . .	50
4.1.4	Stabilita . . . . .	50
4.1.5	Konvergencia . . . . .	50
4.2	Eliptické parciálne diferenciálne rovnice . . . . .	51
4.2.1	Poissonova rovnica v 1-rozmernom priestore . . . . .	51
4.2.2	Gaussova-Seidelova metóda . . . . .	51
4.2.3	Metóda postupnej relaxácie . . . . .	52
4.2.4	Poissonova rovnica v 2-rozmernom priestore . . . . .	52
4.3	Parabolické parciálne diferenciálne rovnice . . . . .	53
4.3.1	Numerické riešenie . . . . .	54
4.3.2	Explicitná FTCS metóda . . . . .	54
4.3.3	Konzistencia a presnosť . . . . .	55
4.3.4	Postupné mapovanie a numerická stabilita . . . . .	55
4.3.5	Efekt hraničných podmienok . . . . .	56
4.3.6	Von Neumannova stabilita . . . . .	57
4.3.7	Explicitná CTCS metóda . . . . .	58
4.3.8	Explicitná Du Fortova a Frankelova metóda . . . . .	58
4.3.9	Implicitná BTCS alebo Laasonenova metóda . . . . .	59
4.3.10	Crankova-Nicolsonova metóda . . . . .	61
4.3.11	Rovnica difúzie v 2D . . . . .	62
4.3.12	Explicitná FTCS metóda . . . . .	62
4.3.13	Implicitná BTCS metóda . . . . .	63
4.3.14	Metóda striedavých smerov (ADI) . . . . .	63
<b>5</b>	<b>Monte Carlo simulácie</b>	<b>65</b>
5.1	Úvod . . . . .	65
5.1.1	Aké problémy môžeme riešiť pomocou MC simulácií? . . . . .	66
5.1.2	Vzťah teória-experiment-simulácia-realita . . . . .	67

## Obsah

---

5.2	Základné pojmy teórie pravdepodobnosti . . . . .	68
5.3	Monte Carlo odhad . . . . .	70
5.3.1	Stredná hodnota súboru . . . . .	70
5.3.2	Rozptyl a smerodajná odchýlka odhadu strednej hodnoty súboru	71
5.4	Generátory náhodných čísiel . . . . .	73
5.4.1	Kongruenčná metóda . . . . .	73
5.4.2	Shift register (Tauswortheov) generátor . . . . .	73
5.4.3	Neuniformné (nerovnomerné) generátory náhodných čísiel . . . . .	74
5.5	Jednoduché a dôležité vzorkovanie . . . . .	75
5.5.1	Jednoduché vzorkovanie . . . . .	77
5.5.2	Dôležité vzorkovanie . . . . .	78
5.5.3	Metropolisov algoritmus . . . . .	81
5.6	Isingov model . . . . .	83
	Literatúra . . . . .	86

## O autorovi

Doc. RNDr. Milan Žukovič, PhD. je teoretický fyzik, zaoberajúci sa problematikou komplexných systémov so zameraním na spinové modely v magnetizme. Taktiež sa zaoberá niektorými netradičnými aplikáciami fyzikálne motivovaných modelov, napr. vo finančníctve, geoštatistike, enviromentálnom monitorovaní a spracovaní digitálneho obrazu. Má bohaté skúsenosti v oblasti aplikácií metód Monte Carlo na štúdium spinových systémov a modelovanie časovo-priestorových náhodných polí.

Vedecko-pedagogickej činnosti sa venuje od roku 1990. Jeho predchádzajúce akademické pôsobiská boli Slovenská akadémia vied (1990), Technická Univerzita v Košiciach (1991-1996), Univerzita Kyushu v Japonsku (1996-2002) a Technická univerzita Kréty v Grécku (2006-2009). Od roku 2009 pracuje na Prírodovedeckej fakulte Univerzity Pavla Jozefa Šafárika v Košiciach. Aktuálne zastáva funkciu docenta na Katedre teoretickej fyziky a astrofyziky PF UPJŠ, kde prednáša pre študentov bakalárskeho, magisterského a doktorandského štúdia kurzy Numerických metód, Počítačovej fyziky a Ekonofyziky. Je odborným spolugarantom študijného programu Progresívne materiály na doktorandskom stupni vzdelávania.



# Predslov

Tento učebný text nadväzuje na základný kurz numerických metód so zameraním na výklad základných princípov počítačového riešenia niektorých typických fyzikálnych úloh. Pokrýva jednak oblasť deterministických metód riešenia problémov modelovaných obyčajnými a parciálnymi diferenciálnymi rovnicami ako aj oblasť stochastických Monte Carlo simulácií. Kurz Počítačová fyzika I, ktorý je prednášaný na bakalárskom stupni študijného programu Fyzika, vytvára základ pre ďalšie štúdium pokročilejších počítačových metód prednášaných na magisterskom stupni v rámci kurzu Počítačová fyzika II.

V súčasnosti je používanie počítača ako nástroja riešenia reálnych fyzikálnych úloh, a v prípade študentov ako nástroja vypracovania ich záverečných bakalárskych či diplomových prác, takmer nevyhnutnosťou. To viedlo k vzniku množstva užívateľsky orientovaných softvérových balíkov v rôznych prostrediach pre riešenie širokej škály problémov. Avšak, bez dôsledného pochopenia fungovania jednotlivých počítačových techník sú tieto programy čiernymi skrinkami s množstvom úskalí a rizík pri ich používaní. Preto tento kurz kladie dôraz na pochopenie silných aj slabých stránok jednotlivých metód a predpokladov ich úspešného použitia. To je precvičované na cvičeniach formou vypracovania projektov implementovaných vo voľne šíriteľnom softvéri Octave, s ktorým sa študenti majú možnosť oboznámiť v predchádzajúcom štúdiu. Využitie prostredia Octave, ktoré je nenáročné na programovacie schopnosti užívateľa a tak vhodné na výučbové účely, zároveň dáva možnosť porovnať výsledky vlastne naprogramovaných algoritmov s tými zabudovanými v knižničných programoch.

Na záver by som sa chcel poďakovať recenzentom prof. Ing. Dušanovi Krokavcovi, CSc. a doc. RNDr. Denisovi Hováthovi, PhD. za starostlivé preštudovanie textu ako aj ich cenné pripomienky a návrhy, ktoré prispeli k zvýšeniu kvality tohto učebného textu po obsahovej aj formálnej stránke. Taktiež by som chcel vyjadriť vďačnosť RNDr. Michalovi Borovskému za textové korekcie.

Košice, december 2014,

Milan Žukovič

# Kapitola 1

## Počítačové riešenie fyzikálnych úloh

### 1.1 Príprava úlohy

Počítačové riešenie úlohy prostredníctvom numerických metód je len súčasťou (avšak dôležitou) viacerých štádií, ktoré je potrebné vykonať pre správne vyriešenie fyzikálneho problému.

- V prvom rade je potrebné ***jednoznačne formulovať fyzikálny problém***, ktorý chceme riešiť. Mnoho problémov vo vede a technike sa najvhodnejšie formuluje v matematických termínoch.
- To umožňuje ***sformulovať matematický model*** vo forme rovníc vystihujúcich základné vlastnosti fyzikálneho systému prostredníctvom premenných popisujúcich systém a teda obyčajne idealizujúcich skutočné fyzikálne procesy. Predpokladáme, že rovnice jednoznačne špecifikujú vzťah medzi vstupnými a výstupnými veličinami. Napríklad, matematický model modifikovaného ideálneho plynu, ktorý je opísaný tlakom  $p$ , objemom  $V$ , teplotou  $T$  a počtom častíc  $n$ , má tvar rovnice  $(p + na^2/V^2)(V - nb) = nRT$ , kde  $a, b, R$  sú konštanty. Avšak, iba najjednoduchšie problémy môžu byť vyriešené s použitím analytických prístupov. Obyčajne takéto rovnice môžu byť (a vo väčšine reálnych problémov aj sú) príliš zložité na to aby sa dali riešiť analyticky.
- Vtedy nastupuje ***použitie numerických metód***, ktoré nahradia zložitejšiu úlohu postupnosťou jednoduchších úloh len s použitím aritmetických prostriedkov, a tak umožnia získať výsledok s istou presnosťou. Teda môžeme povedať že numerické metódy sa v praxi používajú na získanie približného riešenia matematických problémov, ktoré predstavujú matematický model fyzikálnej reality.

## 1.2 Numerické riešenie

---

- Ďalšou fázou je **príprava algoritmu** zvolenej numerickej metódy, pozostávajúcej zo špecifikácie konečnej postupnosti operácií, ktorou zo vstupných údajov dostaneme požadované výsledky.
- Následuje transformácia algoritmu do vybraného programovacieho jazyka, t. j. **programovanie a ladenie programu**.
- Požadované údaje potom získame **spustením programu** a posledným štádiom je vhodné **spracovanie výstupných dát**.

## 1.2 Numerické riešenie

V ďalšom budeme predpokladať, že máme správne definovaný matematický model a sústredíme sa na jeho numerické riešenie. To predstavuje jednoznačný popis vzťahu medzi konečným počtom vstupných a výstupných dát (tzv. diskretná úloha). Kľúčové pojmy popisu numerickej metódy sú:

- **Diskretizácia úlohy** - namiesto pôvodnej (obyčajne spojitej) úlohy riešime diskretnú úlohu s možnosťou priblíženia sa presnému riešeniu pôvodnej úlohy s ľubovoľnou presnosťou.
- **Rád diskretizácie** - kladné číslo  $k$ , také, že platí  $|I - I_h| \leq ch^k$ , kde  $I$  je presný výsledok,  $I_h$  približný výsledok riešenia diskretizovanej úlohy,  $h$  je diskretizačný parameter (krok) a  $c$  je nejaká konštanta nezávislá na  $h$ .
- **Iteračný proces** - postup, pri ktorom sa opakovane používa nejaká operácia s cieľom postupne sa priblížiť k požadovanému riešeniu.
- **Podmienenosť úlohy** - vplyv relatívnej zmeny vstupných údajov na relatívnu zmenu výstupných údajov.
- **Číslo podmienenosti** - taká hodnota  $C$ , že platí  $r(y) \leq Cr(x)$ , kde  $r(x), r(y)$  sú relatívne zmeny vstupných a výstupných údajov. Ak je malé ( $C \approx 1$ ), hovoríme že úloha je dobre podmienená. V opačnom prípade je zle podmienená.
- **Stabilita** - numerická metóda je stabilná, ak je dobre podmienená, t. j. málo citlivá na poruchy v údajoch a je numericky stabilná ak je málo citlivá na vplyv zaokrúhlovacích chýb. V opačnom prípade je metóda nestabilná.
- **Konvergenca** - numerická metóda je konvergentná, ak iteračný proces konverguje k správne riešeniu.

### 1.3 Chyby počítačového riešenia

---

- **Rád konvergenzie** - nech  $\{\beta_n\}_{n=1}^{\infty}$  je postupnosť konvergujúca k nule a  $\{\alpha_n\}_{n=1}^{\infty}$  konverguje k  $\alpha$  ak existuje také  $K > 0$ , že

$$|\alpha_n - \alpha| \leq K|\beta_n|. \quad (1.1)$$

Potom rád konvergenzie  $\{\alpha_n\}_{n=1}^{\infty}$  k  $\alpha$  je  $O(\beta_n)$ . Obyčajne

$$\alpha_n = \alpha + O(1/n^p). \quad (1.2)$$

### 1.3 Chyby počítačového riešenia

Chyba predstavuje rozdiel medzi získanou a skutočnou<sup>1</sup> hodnotou vo fyzikálnom systéme. Za predpokladu, že model budeme považovať za presný a nedopustíme sa chýb z nedbanlivosti, budeme rozlišovať následovné zdroje chýb vzniknutých pri počítačovom riešení úlohy.

- **Chyby numerickej metódy** vznikajú pri aproximácií rovníc modelu s cieľom získať jednoduchšiu, numericky riešiteľnú úlohu v dôsledku diskretizácie pôvodnej matematickej úlohy, alebo náhrady nekonečného procesu konečným (napr. výpočet  $e^x$  pomocou Taylorovho polynómu, výpočet integrálu pomocou súčtu konečného počtu hodnôt). Odhad tejto chyby sa obyčajne dá urobiť analyticky a je dôležitou súčasťou riešenia.

**Príklad:** Výpočet funkčnej hodnoty  $\sin(1)$  sčítaním konečného počtu členov Taylorovho rozvoja v bode  $x = 1$  je aproximovaný podľa vzťahu

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots \quad (1.3)$$

Dá sa ukázať, že sčítaním prvých  $n$  členov postupnosti sa dopustíme chyby veľkosti nanajvyš  $1/(2n+1)!$ .

- **Chyby vstupných údajov** môžu vzniknúť v dôsledku zaokrúhľovania vstupných hodnôt, meracími chybami (konečný počet meraní) alebo stanovením empirických závislostí.
- **Zaokrúhľovacie chyby** vznikajú v dôsledku počítačovej reprezentácie čísiel, t. j. reprezentácie s konečným počtom cifier.
- **Chyby aritmetických operácií** predstavujú šírenie (propagáciu) chýb v dôsledku aritmetických operácií.

---

<sup>1</sup>Skutočná hodnota, t. j. presné číslo, je obyčajne neznáma a je chápaná ako hodnota určená na základe nekonečného počtu meraní.

### 1.3 Chyby počítačového riešenia

---

Chybami numerických metód sa budeme bližšie zaoberať pri výklade jednotlivých metód. Chyby vo vstupných údajoch a chyby vznikajúce pri reprezentácii čísel budú vykonávaním aritmetických operácií spolu interagovať a naším cieľom bude odhadnúť ich vplyv na celkovú chybu výsledku.

#### 1.3.1 Zaokrúhľovacie chyby a počítačová aritmetika

Iba malá časť reálnych čísiel môže byť reprezentovaná v počítači. *Zaokrúhľovacie chyby* sa vždy vyskytujú v počítačových výpočtoch, pretože sa používa iba približná reprezentácia reálnych čísiel. Zaokrúhľovacie chyby možno znížiť použitím aritmetiky vyšších rádov (dvojnásobná presnosť), ale nemožno ich celkom odstrániť. Počítače spracovávajú čísla pomocou *reprezentácie s pohyblivou (plávajúcou) desatinnou čiarkou*, skladajúcej sa z hlavnej časti (*mantis*) a exponentu.

Maximálne a minimálne čísla, ktoré počítač môže reprezentovať, sú fixne stanovené. Keď je vykonaný pokus prekročiť maximálne číslo, nastáva *pretečenie*. a keď je vykonaný pokus ísť pod minimálne číslo, nastáva *podtečenie*. V systémoch s plávajúcou desatinnou čiarkou sa dajú uložiť kladné čísla zhruba v rozmedzí od  $2,2251 \times 10^{-308}$  po  $1,7977 \times 10^{308}$  s relatívnou presnosťou  $2,2204 \times 10^{-16}$ . Napríklad, v programovacom prostredí Octave jednotlivé presné hodnoty dostaneme pomocou príkazov `realmin`, `realmax` a `eps`.

Ak je číslo v numerickom rozmedzí počítača, jeho reprezentácia v plávajúcej čiarky môže byť získaná *odseknutím* alebo *zaokrúhlením*. Výsledkom toho je, že ak sú pri výpočtoch použité čísla v plávajúcej čiarky, vždy dochádza k *zaokrúhľovacím chybám*.

**Príklad:** Reprezentácia čísla  $\pi$  na päť desatinných miest s použitím odseknutia je  $0.31415 \times 10^1$ , kým s použitím zaokrúhlenia dostávame  $0.31416 \times 10^1$ .

#### 1.3.2 Definície chýb a ich šírenie

Nech  $x$  je presná hodnota reálneho čísla a  $\tilde{x}$  nech je jeho aproximácia. Budeme rozlišovať následovne druhy chýb:

- *absolútna chyba*  $e$

$$e(\tilde{x}) = x - \tilde{x}, \quad (1.4)$$

- *odhad absolútnej chyby*  $\epsilon$

$$|x - \tilde{x}| \leq \epsilon(\tilde{x}), \quad (1.5)$$

### 1.3 Chyby počítačového riešenia

---

- **relatívna chyba**  $r$  (pre  $\tilde{x} \neq 0$ )

$$r(\tilde{x}) = \frac{x - \tilde{x}}{\tilde{x}}, \quad (1.6)$$

- **odhad relatívnej chyby**  $\delta$

$$\left| \frac{x - \tilde{x}}{\tilde{x}} \right| \leq \delta(\tilde{x}). \quad (1.7)$$

Pri aritmetických operáciách sa vyššie uvedené chyby šíria.

- Pre **odhad absolútnej chyby** platia nasledovné vzťahy pre:

1. **súčet**  $\tilde{u} = \tilde{x}_1 + \tilde{x}_2$

$$|e(\tilde{u})| \leq \epsilon(\tilde{x}_1) + \epsilon(\tilde{x}_2), \quad (1.8)$$

2. **rozdiel**  $\tilde{v} = \tilde{x}_1 - \tilde{x}_2$

$$|e(\tilde{v})| \leq \epsilon(\tilde{x}_1) + \epsilon(\tilde{x}_2), \quad (1.9)$$

3. **súčin**  $\tilde{w} = \tilde{x}_1 \cdot \tilde{x}_2$

$$|e(\tilde{w})| \leq |\tilde{x}_1| \epsilon(\tilde{x}_2) + |\tilde{x}_2| \epsilon(\tilde{x}_1), \quad (1.10)$$

4. **podiel**  $\tilde{z} = \frac{\tilde{x}_1}{\tilde{x}_2}$

$$|e(\tilde{z})| \leq \frac{|\tilde{x}_1| \epsilon(\tilde{x}_2) + |\tilde{x}_2| \epsilon(\tilde{x}_1)}{|\tilde{x}_2|^2}. \quad (1.11)$$

- Pre **odhad relatívnej chyby** platia nasledovné vzťahy pre:

1. **súčet**  $\tilde{u} = \tilde{x}_1 + \tilde{x}_2$

$$|r(\tilde{u})| \leq \frac{1}{|\tilde{x}_1 + \tilde{x}_2|} (|\tilde{x}_1| \delta(\tilde{x}_1) + |\tilde{x}_2| \delta(\tilde{x}_2)), \quad (1.12)$$

2. **rozdiel**  $\tilde{v} = \tilde{x}_1 - \tilde{x}_2$

$$|r(\tilde{v})| \leq \frac{1}{|\tilde{x}_1 - \tilde{x}_2|} (|\tilde{x}_1| \delta(\tilde{x}_1) + |\tilde{x}_2| \delta(\tilde{x}_2)), \quad (1.13)$$

### 1.3 Chyby počítačového riešenia

---

3. **súčin**  $\tilde{w} = \tilde{x}_1 \cdot \tilde{x}_2$

$$|r(\tilde{w})| \leq \delta(\tilde{x}_1) + \delta(\tilde{x}_2), \quad (1.14)$$

4. **podiel**  $\tilde{z} = \frac{\tilde{x}_1}{\tilde{x}_2}$

$$|r(\tilde{z})| \leq \delta(\tilde{x}_1) + \delta(\tilde{x}_2). \quad (1.15)$$

Vyššie uvedené vzťahy sú špeciálnym prípadom odhadu chýb funkčných hodnôt keď je funkcia daná sčítaním, odčítaním, násobením a delením dvoch približných hodnôt. Vzťah pre odhad chyby funkčných hodnôt v prípade všeobecnej funkcie  $f$  odvodíme nasledovne. Majme funkciu  $f$  diferencovateľnú v oblasti  $G$ , pričom  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ ,  $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$  a  $\epsilon(\tilde{x}_i) \geq |x_i - \tilde{x}_i|$  pre  $i = 1, 2, \dots, n$ .

Za predpokladu, že poznáme horné ohraničenia parciálnych derivácií

$$\left| \frac{\partial f(\mathbf{x})}{\partial x_i} \right| \leq K_i, \quad \forall \mathbf{x} \in G, \quad i = 1, 2, \dots, n, \quad (1.16)$$

**odhad absolútnej chyby funkčnej hodnoty**  $\tilde{u} = f(\tilde{\mathbf{x}})$  môžeme urobiť pomocou Lagrangeovej vety:

$$|u - \tilde{u}| = |f(x_1, x_2, \dots, x_n) - f(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)| \leq \sum_{i=1}^n K_i \epsilon(\tilde{x}_i) = \epsilon(\tilde{u}). \quad (1.17)$$

Pre **odhad relatívnej chyby funkčnej hodnoty** potom platí:

$$\delta(\tilde{u}) = \frac{\epsilon(\tilde{u})}{|u|} = \frac{1}{|u|} \sum_{i=1}^n K_i \epsilon(\tilde{x}_i). \quad (1.18)$$

Za predpokladu, že oblasť  $G$  je malá, sa v praxi namiesto odhadov  $K_i$  používajú hodnoty  $|\partial f(\tilde{\mathbf{x}})/\partial x_i|$  a namiesto hodnoty  $u$  sa používa  $\tilde{u}$ .

**Príklad:** Odhadnite hodnotu funkcie  $u = xy^2z^3$  a jej absolútnu a relatívnu chybu v bode  $x = 12, 3 \pm 0, 25$ ,  $y = 9, 32 \pm 0, 15$ ,  $z = 4, 72 \pm 0, 02$ .

*Riešenie:* Potrebujeme vypočítať  $\epsilon(\tilde{u})$ , pričom  $\tilde{\mathbf{x}} = (12, 3; 9, 32; 4, 72)$  a

$$\frac{\partial f(\tilde{\mathbf{x}})}{\partial x} = (9, 32)^2 \cdot (4, 72)^3 = 9133, 9330, \quad (1.19)$$

$$\frac{\partial f(\tilde{\mathbf{x}})}{\partial y} = 2 \cdot (12, 3) \cdot (9, 32) \cdot (4, 72)^3 = 24109, 8789, \quad (1.20)$$

$$\frac{\partial f(\tilde{\mathbf{x}})}{\partial z} = 3 \cdot (12, 3) \cdot (9, 32)^2 \cdot (4, 72)^2 = 71407, 2303. \quad (1.21)$$

### 1.3 Chyby počítačového riešenia

---

Potom  $\epsilon(\tilde{u}) \approx 9133,9330.0,25 + 24109,8789.0,15 + 71407,2303.0,02 = 7328,1097$  a teda platí

$$u = f(\tilde{\mathbf{x}}) \pm \epsilon(\tilde{u}) = 112347,3756 \pm 7328,1097. \quad (1.22)$$

Pre odhad relatívnej chyby funkčnej hodnoty platí

$$\delta(\tilde{u}) \approx \frac{\epsilon(\tilde{u})}{|f(\tilde{\mathbf{x}})|} = \frac{7328,1097}{112347,3756} \approx 0,0652, \quad (1.23)$$

teda relatívna chyba je 6,52%.

#### 1.3.3 Obrátená úloha teórie chýb

Častou úlohou v praxi je určiť odhady absolútnych chýb  $\epsilon(\tilde{x}_i)$  približných hodnôt  $\tilde{x}_i$  tak, aby absolútna chyba funkčnej hodnoty neprekročila predom stanovenú hodnotu  $\epsilon(\tilde{u})$ . Existuje viacero spôsobov ako túto požiadavku splniť podľa toho, čo predpokladáme.

1. Predpokladajme, že chceme aby všetky parciálne diferenciály mali približne rovnaký vplyv na absolútnu chybu uvažovanej funkcie  $u = f(\mathbf{x})$ , teda aby platilo

$$\left| \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_1} \right| \epsilon(\tilde{x}_1) = \left| \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_2} \right| \epsilon(\tilde{x}_2) = \dots = \left| \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_n} \right| \epsilon(\tilde{x}_n). \quad (1.24)$$

Potom

$$\epsilon(\tilde{u}) \approx \sum_{i=1}^n \left| \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i} \right| \epsilon(\tilde{x}_i) = n \left| \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i} \right| \epsilon(\tilde{x}_i), \quad i = 1, 2, \dots, n. \quad (1.25)$$

Odtiaľ dostávame

$$\epsilon(\tilde{x}_i) \approx \frac{\epsilon(\tilde{u})}{n \left| \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i} \right|}. \quad (1.26)$$

2. Predpokladajme, že chceme aby odhady absolútnych chýb  $\epsilon(\tilde{x}_i)$  mali približne rovnaký vplyv na absolútnu chybu uvažovanej funkcie  $u = f(\mathbf{x})$ , teda aby platilo

$$\epsilon(\tilde{x}_1) = \epsilon(\tilde{x}_2) = \dots = \epsilon(\tilde{x}_n). \quad (1.27)$$

Potom

$$\epsilon(\tilde{u}) \approx \sum_{i=1}^n \left| \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i} \right| \epsilon(\tilde{x}_i) = \epsilon(\tilde{x}_i) \sum_{i=1}^n \left| \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i} \right|, \quad i = 1, 2, \dots, n. \quad (1.28)$$



### 1.3 Chyby počítačového riešenia

---

Odtiaľ dostávame

$$\epsilon(\tilde{x}_i) \approx \frac{\epsilon(\tilde{u})}{\sum_{i=1}^n |\partial f(\tilde{\mathbf{x}})/\partial x_i|}. \quad (1.29)$$

3. Chceme aby odhady relatívnych chýb  $\delta(\tilde{x}_i)$  mali približne rovnaký vplyv, teda aby platilo

$$\frac{\epsilon(\tilde{x}_1)}{|\tilde{x}_1|} = \frac{\epsilon(\tilde{x}_2)}{|\tilde{x}_2|} = \dots = \frac{\epsilon(\tilde{x}_n)}{|\tilde{x}_n|} = k. \quad (1.30)$$

Potom pre  $i = 1, 2, \dots, n$  je  $\epsilon(\tilde{x}_i) = k|\tilde{x}_i|$  a

$$\epsilon(\tilde{u}) \approx \sum_{i=1}^n \left| \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i} \right| \epsilon(\tilde{x}_i) = k \sum_{i=1}^n \left| \tilde{x}_i \frac{\partial f(\tilde{\mathbf{x}})}{\partial x_i} \right|. \quad (1.31)$$

Odtiaľ dostávame

$$k \approx \frac{\epsilon(\tilde{u})}{\sum_{i=1}^n |\tilde{x}_i \partial f(\tilde{\mathbf{x}})/\partial x_i|} \quad (1.32)$$

a pre odhad  $\epsilon(\tilde{x}_i)$  platí

$$\epsilon(\tilde{x}_i) = k|\tilde{x}_i| \approx \frac{|\tilde{x}_i| \epsilon(\tilde{u})}{\sum_{i=1}^n |\tilde{x}_i \partial f(\tilde{\mathbf{x}})/\partial x_i|}. \quad (1.33)$$

Poznamenajme, že vo všetkých prípadoch sa jedná len o približné riešenie vzhľadom na to, že na presné splnenie požiadavky na neprekročenie požadovanej tolerancie je potrebné použiť namiesto uvedených parciálnych derivácií odhady  $K_i$ .

**Príklad:** Majme valec s polomerom podstavy  $R \approx 2\text{m}$  a výškou  $H \approx 3\text{m}$ . Aké môžu byť maximálne tolerancie  $\epsilon(\tilde{R})$  a  $\epsilon(\tilde{H})$ , aby objem valca spĺňal toleranciu  $\epsilon(\tilde{V}) = 0,1\text{m}^3$ ?

*Riešenie:* Pre objem valca platí  $V = \pi R^2 H$ . Označme  $\tilde{R} = 2\text{m}$ ,  $\tilde{H} = 3\text{m}$  a približnú hodnotu  $\pi$  označíme  $\tilde{p} \approx 3,14$ . Potom  $\tilde{V} = \tilde{p} \tilde{R}^2 \tilde{H}$ . Budeme predpokladať že sú splnené predpoklady prípadu (a) a platí  $n = 3$ ,  $x_1 = p$ ,  $x_2 = R$ ,  $x_3 = H$ . Potom pre jednotlivé tolerancie dostávame

$$\epsilon(\tilde{p}) = \frac{\epsilon(\tilde{V})}{3|\partial \tilde{V}/\partial p|} = \frac{0,1}{3 \cdot 12} < 0,003, \quad (1.34)$$

$$\epsilon(\tilde{R}) = \frac{\epsilon(\tilde{V})}{3|\partial \tilde{V}/\partial R|} = \frac{0,1}{3 \cdot 2 \cdot 3 \cdot 14 \cdot 2 \cdot 3} < 8,85 \cdot 10^{-4}, \quad (1.35)$$

$$\epsilon(\tilde{H}) = \frac{\epsilon(\tilde{V})}{3|\partial \tilde{V}/\partial H|} = \frac{0,1}{3 \cdot 3 \cdot 14 \cdot 2^2} < 2,6543 \cdot 10^{-3}. \quad (1.36)$$

$$(1.37)$$

# Kapitola 2

## Úvod do dynamických systémov

### 2.1 Dynamické systémy

**Dynamický systém** Pod dynamickým systémom si môžeme predstaviť v podstate akýkoľvek systém vyvíjajúci sa v čase, ktorý robí to isté znova a znova a vždy vie, čo bude robiť ďalej. Túto širokú definíciu v ďalšom zúžime na matematické dynamické systémy ako matematické modely fyzikálnej reality. Dynamický systém má dve časti: *stavový vektor*, ktorý presne popisuje stav nejakého systému, a *funkciu* (t. j. pravidlo), ktorá definuje v akom stave bude systém v budúcom časovom okamihu vzhľadom k súčasnému stavu.

**Stavový vektor** Fyzikálne systémy môžu byť popísané číslami. Napríklad, stav lopty vyhodenej kolmo hore môže byť popísaný pomocou dvoch čísiel: jeho výšky  $h$  a rýchlosti  $v$ . Dvojica čísiel  $(h, v)$  je vektor ktorý úplne popisuje stav lopty a preto sa nazýva stavový vektor systému. Typicky, stavové vektory budeme zapisovať ako stĺpcové vektory, takže presnejšie stav tohto systému je  $(h, v)^T$ . V niektorých prípadoch je možné popísať stav systému pomocou jediného čísla. Napríklad, uvažujme bankový účet otvorený s vkladom 100 EUR a s ročným úrokom vo výške 6%. Stav tohto systému v ktoromkoľvek okamihu je možné popísať jediným číslom a to zostatkom na účte. V tomto prípade, stavový vektor má iba jednu zložku. Na druhej strane, niektoré dynamické systémy vyžadujú na popis obrovské množstvo čísiel. Napríklad, dynamický systém ktorý modeluje globálne počasie môže mať milióny premenných, ktoré popisujú teplotu, tlak, rýchlosť vetra, a tak ďalej, v množstve bodov mapy celej zemegule. Napriek extrémnej zložitosti, na stav systému sa môžeme jednoducho pozeráť ako na zoznam čísiel, teda vektor, ktorý budeme označovať napríklad  $\mathbf{x}$ .

## 2.1 Dynamické systémy

---

**Stav v ďalšom okamihu - diskretný čas** Druhá časť dynamického systému je pravidlo, ktoré nám hovorí, ako sa systém mení v čase. Inými slovami, ak máme daný aktuálny stav systému, pravidlo nám hovorí, aký bude stav systému v ďalšom okamihu. V prípade účtu popísaného vyššie, bude ďalší okamih o rok neskôr, keďže úroky sú vyplácané len raz ročne, a teda čas je diskretný. Pre bankový účet je ľahké definovať dynamické pravidlo, ktoré bude meniť stav systému z jedného okamihu na stav systému v ďalšom okamihu:

$$\mathbf{x}(k+1) = 1.06\mathbf{x}(k), \quad (2.1)$$

kde  $k$  predstavuje diskretný čas a  $\mathbf{x}(k)$  stav systému v čase  $k$ . Pre úplný popis systému však potrebujeme poznať aj počiatočný stav  $\mathbf{x}(0)$  a teda, v našom prípade, úplný popis systému má tvar

$$\mathbf{x}(k+1) = 1.06\mathbf{x}(k), \quad (2.2)$$

$$\mathbf{x}(0) = 100. \quad (2.3)$$

Obyčajne začíname v čase  $k = 0$  a počiatočný stav zapisujeme  $\mathbf{x}_0 = \mathbf{x}(0)$ . Potom stav bankového účtu v ďalších rokoch môžeme vyjadriť v tvare

$$\mathbf{x}(k) = (1.06)^k \mathbf{x}_0. \quad (2.4)$$

Je zrejmé, že táto všeobecná rovnica (2.4) spĺňa obidve predchádzajúce rovnosti, teda počiatočnú podmienku ak dosadíme  $k = 0$ , t. j.  $\mathbf{x}(0) = (1.06)^0 \mathbf{x}_0 = \mathbf{x}_0$ , a taktiež pre  $k + 1$  dostávame  $\mathbf{x}(k+1) = (1.06)^{k+1} \mathbf{x}_0 = (1.06)(1.06)^k \mathbf{x}_0 = (1.06)\mathbf{x}(k)$ .

**Stav v ďalšom okamihu - spojitý čas** Na rozdiel od bankových účtov, ktoré sa menia len raz za rok alebo v iných diskretných cykloch, zmeny mnohých systémov je lepšie popísať v spojitom čase. Zoberme si už spomenutý príklad hodu loptou nahor. Jej okamžitý stav je daný jeho stavovým vektorom  $\mathbf{x} = (h, v)^T$ . Nemá však zmysel sa pýtať aký bude jej stav v ďalšom okamihu, keďže čas beží spojite. Pre spojite sa meniaci čas budeme používať označenie  $t$ , namiesto  $k$ , a  $t$  bude nezáporné reálne číslo zvyšujúce sa od hodnoty  $t = 0$ . Vzhľadom k tomu, že nemôžeme napísať pravidlo pre "ďalší" časový okamih, budeme popisovať ako sa systém mení v danom okamihu. Ak naša lopta má vertikálnu rýchlosť  $v$ , potom vieme, že  $dh/dt = v$ , čo je definícia rýchlosti. Ďalej, gravitácia ťahá loptu dole a preto máme  $dv/dt = -g$ , kde  $g$  je kladná konštanta. Teda zmeny v systéme môžeme popísať závislosťami

$$\begin{aligned} h'(t) &= v(t) \\ v'(t) &= -g, \end{aligned} \quad (2.5)$$

## 2.2 Príklady dynamických systémov

---

čo môžeme prepísať do maticového tvaru

$$\mathbf{x}' = \mathbf{f}(\mathbf{x}), \quad (2.6)$$

kde  $\mathbf{x}(t) = \begin{pmatrix} h(t) \\ v(t) \end{pmatrix}$ ,  $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ ,  $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  a  $\mathbf{b} = \begin{pmatrix} 0 \\ -g \end{pmatrix}$ .

Vyššie uvedený tvar rovnice (2.6) platí pre všetky dynamické systémy spojite sa meniace s časom. Majú stavový vektor  $\mathbf{x}(t) \in \mathbf{R}^n$  a funkcia  $\mathbf{f} : \mathbf{R}^n \rightarrow \mathbf{R}^n$  definuje ako rýchlo sa jednotlivé zložky menia, t. j.  $\mathbf{x}'(t) = \mathbf{f}(\mathbf{x}(t))$ , alebo stručnejšie  $\mathbf{x}' = \mathbf{f}(\mathbf{x})$ .

Vráťme sa k nášmu príkladu a predpokladajme, že lopta je za začiatku vo výške  $h_0$  a má vertikálnu rýchlosť  $v_0$ , teda  $\mathbf{x}_0 = \begin{pmatrix} h_0 \\ v_0 \end{pmatrix}$ . Dá sa ľahko overiť, že rovnice

$$h(t) = h_0 + v_0 t - \frac{1}{2}gt^2, \quad (2.7)$$

$$v(t) = v_0 - gt \quad (2.8)$$

popisujú pohyb lopty a teda spĺňajú diferenciálne rovnice (2.5).

Dynamické systémy budeme rozlišovať na autonómne a neautonómne. Autonómne sú také, pre ktoré platia pre diskrétny a spojitý čas rovnice

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}(k)), \quad (2.9)$$

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t)) \quad (2.10)$$

a neautonómne sú také, pre ktoré platia rovnice

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}(k), \mathbf{x}(k)), \quad (2.11)$$

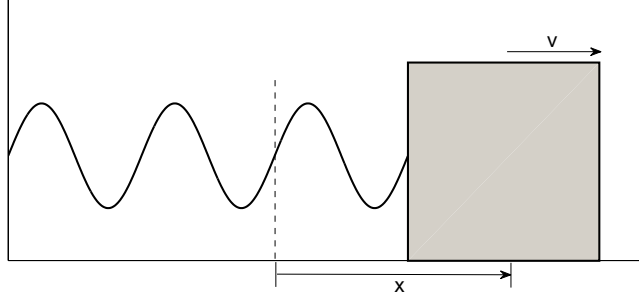
$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}(t), \mathbf{x}(t)). \quad (2.12)$$

## 2.2 Príklady dynamických systémov

### 2.2.1 Mechanický oscilátor

Prvý príklad dynamického systému v spojitom čase je vodorovný pohyb telesa upevneného k stene pomocou ideálnej pružiny (mechanický oscilátor) so zanedbaním trenia (viď obr. 2.1). Stav tohto systému je určený dvoma hodnotami: vzdialenosťou telesa od neutrálnej polohy  $x$  a jeho rýchlosťou pohybu do pravej strany  $v$ . Pre  $x = 0$  budeme predpokladať, že pružina nie je ani natiahnutá, ani stlačená a nepôsobí na teleso žiadnou silou. Keď teleso vychýlime z tejto neutrálnej polohy doprava ( $x > 0$ ), pružina bude ťahať teleso späť vľavo. Naopak, keď teleso vychýlime vľavo ( $x < 0$ ), pružina ho bude ťahať doprava. Za predpokladu, že máme k dispozícii ideálnu pružinu, sila  $F$

## 2.2 Príklady dynamických systémov



**Obr. 2.1:** Vodorovný pohyb telesa upevneného k stene pomocou ideálnej pružiny so zanedbaním trenia.

pôsobiaci na teleso v polohe  $x$  je  $-kx$ , kde  $k$  je kladná konštanta. Znamienko mínus vyjadruje fakt, že sila pôsobí proti smeru posunutia. Z Newtonovho zákona platí  $F = ma$ , kde  $m$  je hmotnosť telesa a  $a = dv/dt$  je zrýchlenie. Po dosadení  $F = -kx$  dostávame

$$v' = -\frac{k}{m}x. \quad (2.13)$$

Z definície d'ajej platí

$$x' = v. \quad (2.14)$$

Ak pre jednoduchosť položíme  $k = m = 1$  a skombinujeme vyššie uvedené rovnice, dostávame

$$\mathbf{y}' = \mathbf{A}\mathbf{y}, \quad (2.15)$$

kde  $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  a stavový vektor  $\mathbf{y} = \begin{pmatrix} x \\ v \end{pmatrix}$ . Predpokladajme, že teleso začína pohyb zo stavu  $\mathbf{y}_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ . Potom môžeme tvrdiť, že vzťah

$$\mathbf{y}(t) = \begin{pmatrix} \cos(t) \\ -\sin(t) \end{pmatrix} \quad (2.16)$$

popisuje pohyb telesa v budúcom čase. Toto tvrdenie môžeme overiť následovne. Potrebujeme overiť (A) že  $\mathbf{y}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  a (B) že  $\mathbf{y}$  spĺňa rovnice (2.15). (A) overíme tak, že jednoducho dosadíme  $t = 0$  do rovnice (2.16) a dostávame

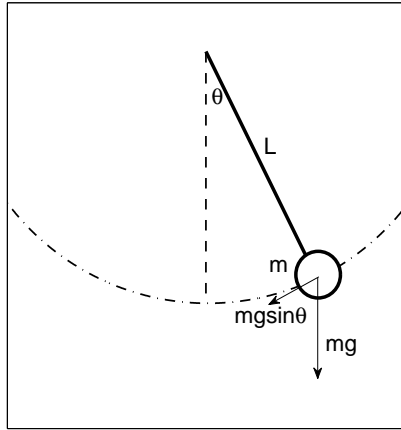
$$\mathbf{y}(0) = \begin{pmatrix} \cos(0) \\ -\sin(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \mathbf{y}_0. \quad (2.17)$$

Pre overenie (B), zderivujeme

$$\mathbf{y}'(t) = \begin{pmatrix} \cos(t) \\ -\sin(t) \end{pmatrix}' = \begin{pmatrix} -\sin(t) \\ -\cos(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \cos(t) \\ -\sin(t) \end{pmatrix} = \mathbf{A}\mathbf{y}(t). \quad (2.18)$$

## 2.2 Príklady dynamických systémov

---



Obr. 2.2: Jednoduché kyvadlo.

Z riešenia  $x(t) = \cos(t)$  vidíme, že teleso sa bude pohybovať dopredu a naspäť do nekonečna, čo samozrejme nie je fyzikálne realistické a je dôsledkom úplného zanedbania trenia.

### 2.2.2 Jednoduché kývadlo

Uvažujme ideálne kyvadlo, ako je znázornené na obr. 2.2. Nech jeho hmotnosť je  $m$  a nech je pripevnené k pevnému čapu tuhou tyčou dĺžky  $L$ . Stav tohto dynamického systému môže byť popísaný dvoma hodnotami: uhlom  $\theta$ , ktorý kyvadlo zvierá s vertikálnou osou, a rýchlosťou otáčania  $\omega$  (merané, povedzme, v radiánoch za sekundu). Z definície platí  $\omega = d\theta/dt$ . Gravitácia ťahá kyvadlo kolmo dole silou  $mg$ . Táto sila môže byť rozložená na dve zložky: jednu paralelnú k tyči a jednu kolmú. Zložka pôsobiaca paralelne s tyčou nemá vplyv na to, ako sa kyvadlo pohybuje a zložka kolmá na tyč má veľkosť  $mg \sin(\theta)$ .

V ďalšom potrebujeme dať do súvisu zrýchlenie  $a$  so stavovou premennou  $\theta$ . Keďže vzdialenosť  $s$  pozdĺž oblúka kyvadla je  $L\theta$ , a platí  $a = s''$ , dostávame  $a = (L\theta)''$ . Z toho vyplýva, že  $\omega' = a/L = (ma)/(mL) = -(mg \sin(\theta))/(mL) = -(g/L) \sin(\theta)$ , čo môžeme zhrnúť do rovníc

$$\begin{aligned}\theta'(t) &= \omega(t), \\ \omega'(t) &= -\frac{g}{L} \sin(\theta(t)).\end{aligned}\tag{2.19}$$

## 2.2 Príklady dynamických systémov

---

Poznamenanajme, že znamienko mínus odráža fakt, že keď  $\theta > 0$ , tak pôsobiaca sila sa snaží vrátiť kyvadlo naspäť do vertikálnej polohy. Ak označíme stavový vektor  $\mathbf{x} = \begin{pmatrix} \theta \\ \omega \end{pmatrix}$ , potom rovnice (2.19) môžeme zapísať v tvare

$$\mathbf{x}' = \mathbf{f}(\mathbf{x}), \quad (2.20)$$

kde  $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  je definovaná ako

$$\mathbf{f} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -\frac{g}{L} \sin(x) \end{pmatrix}. \quad (2.21)$$

Nanešťastie, na rozdiel od prípadu mechanického oscilátora, v prípade kyvadla nie je možné vyjadriť presné riešenie pre popis jeho pohybu, t. j. riešenie rovníc (2.19). Napriek tomu, predstavu o jeho správaní môžeme získať pomocou lineárnej aproximácie a numerických metód.

### Lineárna aproximácia

Problémom pre získanie presného riešenia je nelineárny tvar funkcie (2.21). Avšak v prípade, že uhlový výkyv kyvadla je veľmi malý, tak  $\sin(\theta) \approx \theta$ , keďže platí

$$\lim_{\theta \rightarrow 0} \frac{\sin(\theta)}{\theta} = 1. \quad (2.22)$$

Potom ak v rovniciach (2.19) nahradíme  $\sin(\theta)$  premennou  $\theta$ , môžeme ich prepísať v tvare

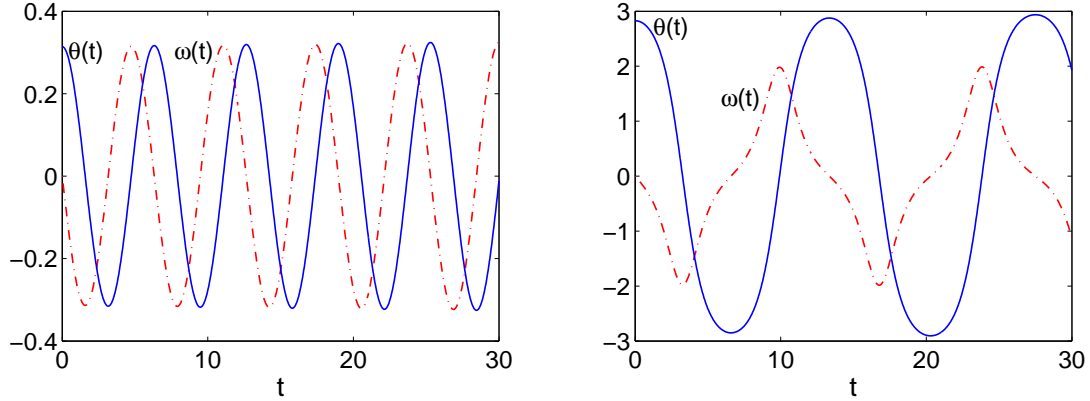
$$\begin{pmatrix} \theta \\ \omega \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -\frac{g}{L} & 0 \end{pmatrix} \begin{pmatrix} \theta \\ \omega \end{pmatrix}. \quad (2.23)$$

Ak položíme  $L = g$  (teda predpokladáme že dĺžka kyvadla je 9.8 m), potom rovnice (2.23) sú identické s rovnicami (2.18) a teda riešením bude sinusoidálna zmena uhla striedavo do jednej a do druhej strany.

### Numerické riešenie

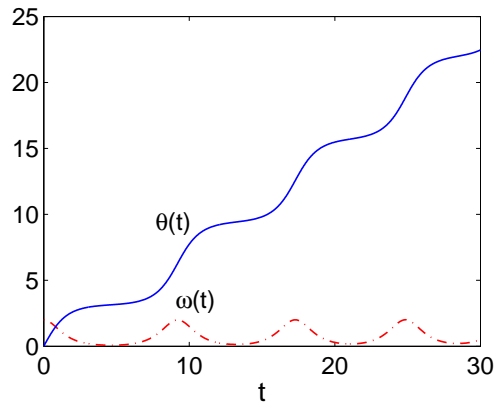
Niektoré systémy diferenciálnych rovníc (napr. rovnice (2.18)) môžu byť vyriešené presne analytickými metódami, zatiaľ čo iné (napr. rovnice (2.20)) nemôžu byť vyriešené. V takom prípade sú užitočné numerické metódy. Jednotlivými metódami riešenia systémov diferenciálnych rovníc sa budeme bližšie zaoberať v nasledujúcich kapitolách. Viaceré sofistikované metódy sú implementované v niektorých programovacích prostrediach, ako napr. Matlab, Octave, Maple, Matematika alebo Mathcad. Pre ilustráciu si uvedieme príklady numerických riešení v prostredí Octave pre niektoré

## 2.2 Príklady dynamických systémov



(a)  $\theta_0 = 1/10\pi, \omega_0 = 0$

(b)  $\theta_0 = 9/10\pi, \omega_0 = 0$



(c)  $\theta_0 = 0, \omega_0 = 2$

**Obr. 2.3:** Časový vývoj pohybu kyvadla pre počiatkové hodnoty  $(\theta_0, \omega_0)$ : (a)  $(1/10\pi; 0)$ , (b)  $(9/10\pi; 0)$ , a (c)  $(0; 2)$ . Plné čiary reprezentujú hodnoty  $\theta$  a bodkované hodnoty  $\omega$ .

počiatkové podmienky, ktoré vedú ku kvalitatívne rozdielnemu správaniu systému. Pre jednoduchosť položíme  $g = L = 9, 8$ , takže náš systém (2.20) má tvar

$$\begin{pmatrix} \theta \\ \omega \end{pmatrix}' = \begin{pmatrix} \omega \\ -\sin(\theta) \end{pmatrix}. \quad (2.24)$$

1. Začneme počiatkovou podmienkou  $\mathbf{x}_0 = \begin{pmatrix} \theta(0) \\ \omega(0) \end{pmatrix} = \begin{pmatrix} 0,1 \\ 0 \end{pmatrix}$ . Znamená to, že kyvadlo vychýlime o malý uhol od vertikálnej polohy. Výsledok je znázornený na obr. 2.3(a). Všimneme si, že výsledné krivky vyzerajú úplne rovnako ako sínusové a kosínusové vlnovky, získané lineárnou aproximáciou diskutovanou vyššie.



## 2.2 Príklady dynamických systémov

---

2. Ďalej vyskúšajme vychýliť kyvadlo o veľký uhol. V prípade  $\theta \approx 3$  je kyvadlo na počiatku v polohe takmer vertikálne nahor a riešenie je znázornené na obr. 2.3(b). Krivky sú síce periodické, avšak nie sínusovky ako v prípade (1).
3. Nakoniec, začnime pohyb z polohy kedy kyvadlo visí rovno dole ( $\theta = 0$ ), ale dostane silný počiatočný impulz ( $\omega = 2$ ). Výsledkom je situácia zobrazená na obr. 2.3(c). Prekvapujúco, kyvadlo nevykonáva periodický pohyb, ale uhol vychylenia sa neustále zväčšuje. Znamená to, že kyvadlo sa neustále otáča v rovnakom smere (všimnite si, že uhol je vždy kladný) a tak sa kyvadlo otáča stále dookola. Je zaujímavé si všimnúť, že kyvadlo strávi väčšinu svojho času v blízkosti zvislej polohy, kedy sa pohybuje najpomalšie.

### 2.2.3 Množenie populácie baktérií

Majme nádobu naplnenú výživnou zmesou v ktorej sa nachádzajú baktérie. Postupom času sa baktérie množia (delením) a odumierajú. Nech  $b$  je rýchlosť s akou sa baktérie reprodukovujú a  $p$  je rýchlosť ktorou odumierajú. Potom populácia rastie rýchlosťou  $b - p$ . To znamená, že ak v nádobe máme  $x$  baktérií, tak ich počet sa mení podľa vzťahu  $(b - p)x$ , a teda zmenu počtu baktérií môžeme vyjadriť vzťahom  $dx/dt = rx$ , kde  $r = b - p$ . Ak začneme v čase  $t = 0$  s  $x_0$  baktériami, tak riešenie tejto jednoduchéj diferenciálnej rovnice môžeme vyjadriť v tvare

$$x(t) = x_0 e^{rt}. \quad (2.25)$$

V krátkodobom horizonte to dáva zmysel. Vzorec hovorí, že na začiatku je  $x_0$  baktérií a potom ich počet rastie exponenciálne. Avšak, s plynutím času sa počet baktérií zväčší tak, že prekročí aj počet atómov vo vesmíre, čo znamená, že jednoduchý model  $dx/dt = rx$  nie je realistický z dlhočasového hľadiska. Reálnejší model je taký, v ktorom budeme uvažovať, že baktérie so zvyšujúcou sa reprodukciou si budú navzájom uberať životný priestor, budú produkovať toxické splodiny, atď., a teda má zmysel uvažovať, že ich úmrtnosť nebude konštantná, ale sa bude zvyšovať s početnosťou populácie. Opäť predpokladajme konštantnú rýchlosť rozmnožovania  $b$ , takže ak máme  $x$  baktérií, tak rýchlosť ich rozmnožovania je  $bx$ . Teraz miesto konštantnej umrtnosti predpokladajme, že je úmerná populácii, teda  $px$ . V takom prípade, ak máme  $x$  baktérií, tak rýchlosť ich odumierania je  $px^2$ . Kombináciou týchto faktorov dostávame dynamický systém

$$\frac{dx}{dt} = bx - px^2. \quad (2.26)$$

Toto je diferenciálna rovnica, ktorú vieme riešiť analyticky a jej riešenie je v tvare

$$x(t) = \frac{x_0 b e^{bt}}{(b - px_0) + px_0 e^{bt}}. \quad (2.27)$$

## 2.2 Príklady dynamických systémov

---

Nájsť analytické riešenie diferenciálnych rovníc je však často ťažké, alebo aj nemožné, a preto v ďalšom skúsme sa pozrieť, čo z rovnice (2.26) dokážeme učiť priamo, bez toho aby sme poznali jej analytické riešenie. Položme si otázku, či existuje pre tento model samoudržateľná populácia. Inými slovami, hľadáme také číslo  $\tilde{x}$ , pre ktoré platí  $b\tilde{x} - p\tilde{x}^2 = 0$ . Splnenie tejto podmienky znamená, že miery rozmnožovania a úmrtnosti sú presne v rovnováhe a (pretože pre  $\tilde{x}$  je  $dx/dt = 0$ ) populácia ani nenarastá ani neklesá, teda je samoudržateľná. Ak položíme pravú stranu rovnice (2.26) rovnú nule, dostaneme

$$\frac{dx}{dt} = bx - px^2 = 0 \Rightarrow x = 0 \text{ alebo } x = \frac{b}{p}. \quad (2.28)$$

Dostávame teda dve riešenia pre samoudržateľnú populáciu,  $\tilde{x} = 0$  a  $\tilde{x} = \frac{b}{p}$ , ktoré odpovedajú dvom koreňom kvadratickej rovnice  $bx - px^2 = 0$ . To je rovnica paraboly a jej graf je znázornený na obr. (2.4). Uvažujme najprv riešenie  $\tilde{x} = 0$ . Samozrejme, že takáto populácia je samoudržateľná, keďže nie sú žiadne baktérie a teda nemôžu sa ani rozmnožovať ani odumierať. Avšak, aj sebemenšia kontaminácia ( $x > 0$ , ale menšie ako  $\frac{b}{p}$ ) spôsobí že  $\frac{dx}{dt} = bx - px^2 > 0$  (pozrite si graf na obr. 2.4). To znamená, že sa počet baktérií začne zvyšovať, akonáhle bola nádoba kontaminovaná. Rovnovážna hodnota  $\tilde{x} = 0$  je nestabilná, aj maličké perturbácie túto rovnováhu zničia<sup>1</sup>.

Na druhej strane, uvažujme riešenie za  $\tilde{x} = \frac{b}{p}$ . Na tejto populačnej úrovni sa baktérie množia rýchlosťou  $b\tilde{x} = b(b/p) = b^2/p$  a odumierajú rýchlosťou  $p\tilde{x}^2 = p(b/p)^2 = b^2/p$ , a teda miera množenia a úmrtnosti sú presne v rovnováhe. Ale uvažujme čo sa stane v prípade, že populácia  $x$  je mierne nad alebo tesne pod úrovňou  $\tilde{x} = \frac{b}{p}$ . Ak  $x$  je mierne nad  $\frac{b}{p}$ , vidíme, že  $dx/dt$  je záporná (pozrite si graf na obr. 2.4), a preto počet baktérií klesne späť na úroveň  $\frac{b}{p}$ . Naopak, ak je  $x$  je mierne nižšia než  $\frac{b}{p}$ , môžeme vidieť, že  $dx/dt$  je kladné a teda populácia bude inklinovať k zvýšeniu späť na úroveň  $\frac{b}{p}$ . Vidíme, že  $\frac{b}{p}$  je rovnovážna hodnota a aj malé odchýlky od  $\frac{b}{p}$  sa samovoľne vrátia späť na úroveň  $\frac{b}{p}$ .

Vráťme sa ešte k analytickému riešeniu (2.27). Môžeme pozorovať že ak  $x_0$  je ľubovoľné kladné číslo, tak  $x(t) \rightarrow \frac{b}{p}$  pre  $t \rightarrow \infty$ . To len potvrdzuje naše úvahy uvedené vyššie, t. j. že systém konverguje k stabilnému pevnému bodu  $\frac{b}{p}$ .

### 2.2.4 Dravec a korisť

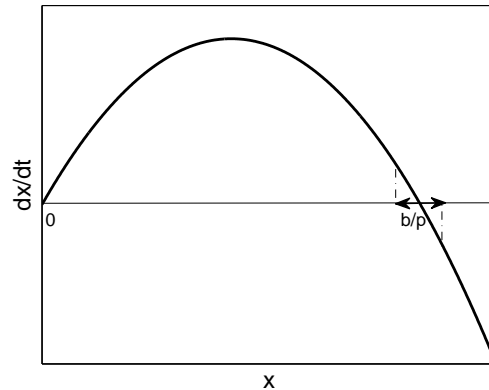
V predchádzajúcej časti sme uvažovali jednoduchý model biologického systému zahŕňajúci iba jeden druh. Teraz budeme uvažovať zložitejší model, do ktorého vstupujú

---

<sup>1</sup>Príklad toho, čo nazývame nestabilný pevný bod.

<sup>2</sup>Príklad toho, čo nazývame stabilný pevný bod.

## 2.2 Príklady dynamických systémov



**Obr. 2.4:** Grafické znázornenie časovej zmeny stavu mikróbov  $dx/dt$  ako funkcie ich stavu  $x$ . Stav  $\tilde{x} = 0$  predstavuje nestabilný a  $\tilde{x} = b/p$  stabilný rovnovážny stav.

dva druhy. Prvý (korisť) môže byť reprezentovaný napríklad nejakým bylinožravcom (povedzme zajace), ktorého populácia v čase  $t$  je  $r(t)$ . Druhý (predátor) sa živí korisťou (povedzme, že sú to vlci) a ich populácia v čase  $t$  je  $w(t)$ .

Ak je populácia zajacov izolovaná, tak sa budú reprodukovat' rýchlosťou  $dr/dt = ar$ , kde  $a$  je nejaká kladná konštanta. Na druhej strane, izolovaná populácia vlkov bude bez zajacov (svojej potravy) hladovať a ich populácia bude klesať rýchlosťou:  $dw/dt = -bw$ , pre nejaké  $b > 0$ .

Avšak, ak dáme do toho istého prostredia vlkov aj zajacov, tak vlci začnú požírať zajacov s očakávanými účinkami na každú populáciu: vlkov bude viac a zajacov menej. Predpokladajme, že máme  $w$  vlkov a  $r$  zajacov. Aká je pravdepodobnosť, že vlk zožerie zajaca? Čím je viac vlkov alebo zajacov, tým je väčšia pravdepodobnosť ich stretu. Z tohto dôvodu môžeme predpokladať, že úhyn zajacov ako aj nárast populácie vlkov bude úmerný  $rw$ . Tieto zmeny v populáciách môžeme vyjadriť nasledovne<sup>3</sup>:

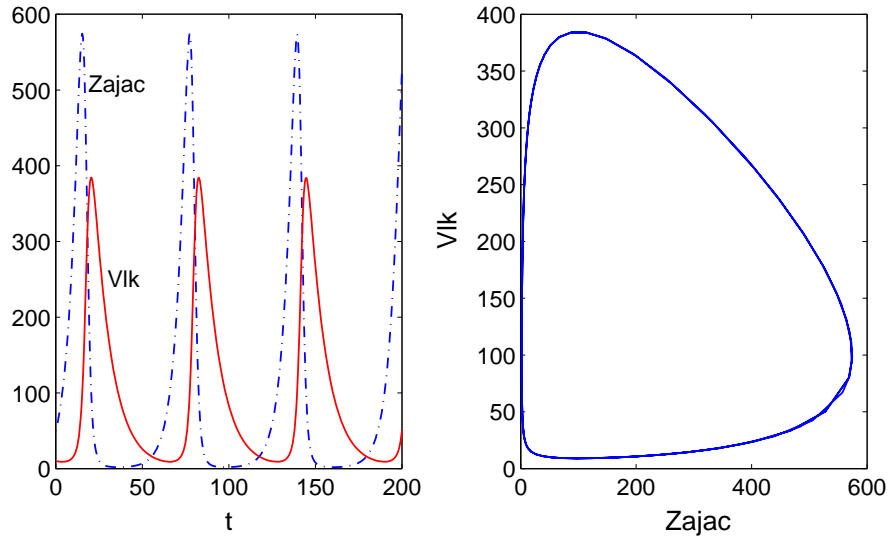
$$\begin{aligned} \frac{dr}{dt} &= ar - grw, \\ \frac{dw}{dt} &= -bw + hrw, \end{aligned} \tag{2.29}$$

kde  $a, b, g, h$  sú kladné konštanty. Tieto rovnice môžeme prepísať do tvaru  $\mathbf{x}' = \mathbf{f}(\mathbf{x})$ , kde  $\mathbf{x} = \begin{pmatrix} r \\ w \end{pmatrix}$ , a  $\mathbf{f} \begin{pmatrix} r \\ w \end{pmatrix} = \begin{pmatrix} ar - grw \\ -bw + hrw \end{pmatrix}$ .

Približné riešenie sústavy týchto diferenciálnych rovníc môžeme nájsť numericky, povedzme v prostredí Octave použitím príkazu `lsode`. Napríklad, nech  $a = 0,2; b =$

<sup>3</sup>Klasický model ekologického systému autorov Lotka a Volterra.

## 2.2 Príklady dynamických systémov



**Obr. 2.5:** (a) Časová zmena populácií vlka (plná krivka) a zajaca (bodko-čiarkovaná krivka). (b) Fázový diagram modelu koristiť-dravec, kde  $x$ -ová os predstavuje populáciu dravca (vlkov) a  $y$ -ová os stav populácie koristi (zajacov).

$0, 1; g = 0,002; h = 0,001; r_0 = 50$ , a  $w_0 = 10$ . Pri pohľade na výsledky na obr. 2.5(a) môžeme vidieť, že populácie zajaca a vlka v čase kolíšu. Ich správanie je periodické - približne každých 9 časových jednotiek sa vzor opakuje. Ak je málo vlkov, tak stúpa populácia zajacov. Tým však zároveň rastie potrava pre vlkov, ktorých populácia začne rásť. Ale ako populácia vlkov stúpa, vlci vo veľkom požívajú zajacov a ich populácia začne výrazne klesať. To spôsobí, že potrava pre vlkov začne byť nedostupná a tak pre zmenu začne klesať populácia vlkov. Nakoniec, keď je počet vlkov dostatočne malý začne sa zotavovať populácia zajacov a cyklus začína odznova.

Cyklickú povahu tohto procesu môžeme najlepšie pozorovať vykreslením správania sa populácií zajaca a vlka do jedného grafu, kde na osi  $x$  vynášame počet zajacov a na os  $y$  počet vlkov (vid' obr. 2.5(b)). Každý bod na krivke predstavuje stav systému. Krivka sa nazýva fázový diagram a v plnom rozsahu predstavuje celý príbeh o tom, ako systém postupuje. Stav systému je reprezentovaný bodom, ktorý putuje proti smeru hodinových ručičiek pozdĺž krivky.

# Kapitola 3

## Riešenie sústav obyčajných diferenciálnych rovníc

### 3.1 Úvod

Diferenciálne rovnice, kde sú podmienky uvedené iba na jednom konci oblasti nezávislej premennej nazývame počiatkové úlohy (PÚ) alebo tiež Cauchyho úlohy. PÚ sú zvyčajne spojené s časovým vývojom dynamického systému z daného počiatkového stavu. V tejto kapitole popíšeme niektoré numerické techniky používané pri riešení PÚ. Jednotlivé techniky zahŕňajú jednokrokové a viackrokové metódy, ako aj explicitné a implicitné schémy.

### 3.2 Počiatkové úlohy pre obyčajné diferenciálne rovnice

Problém PÚ pre obyčajné diferenciálne rovnice pozostáva z obyčajnej diferenciálnej rovnice premennej  $x(t)$ ,  $t \in [a, b]$ <sup>1</sup>

$$\frac{dx}{dt} = f(t, x) \quad (3.1)$$

spolu s počiatkovou podmienkou

$$x(a) = \alpha, \quad (3.2)$$

kde  $f(t, x)$  a  $\alpha$  sú dané. Počiatková podmienka slúži na jednoznačné určenie riešenia (ak existuje), ktoré je ináč určené až na integračnú konštantu.

---

<sup>1</sup>Nezávislou premennou je obyčajne čas a preto ju budeme označovať  $t$ .

### 3.2 Počiatkové úlohy pre obyčajné diferenciálne rovnice

---

Podobný problém môžeme riešiť aj pre prípad sústavy rovníc, kde  $\mathbf{x}(t) = (x_1, x_2, \dots, x_n)$  a  $\mathbf{f}(t, \mathbf{x}) = (f_1(t, \mathbf{x}), f_2(t, \mathbf{x}), \dots, f_n(t, \mathbf{x}))$

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}) \quad (3.3)$$

s počiatkovou podmienkou

$$\mathbf{x}(a) = (\alpha_1, \alpha_2, \dots, \alpha_n). \quad (3.4)$$

Tvar zápisu diferenciálnych rovníc (3.1), resp. (3.3) nazývame *dynamický tvar*.

**PÚ vyššieho rádu** môžu byť pretransformované do systému PÚ prvého rádu. Napríklad, problém stanovenia funkcie  $x(t)$  spĺňajúcej

$$\frac{d^2x}{dt^2} = -\omega^2x \quad (3.5)$$

s počiatkovými podmienkami

$$x(0) = x_0, \quad \left. \frac{dx}{dt} \right|_0 = 0 \quad (3.6)$$

môže byť pretransformovaný do nasledovného systému dvoch PÚ prvého rádu

$$\frac{dx_1}{dt} = x_2, \quad (3.7)$$

$$\frac{dx_2}{dt} = -\omega^2x_1, \quad (3.8)$$

s počiatkovými podmienkami

$$x_1(0) = x_0, \quad x_2(0) = 0. \quad (3.9)$$

**Lipschitzova podmienka** Nech  $\mathbf{f}(t, \mathbf{x})$  je reálna vektorová funkcia  $n + 1$  reálnych premenných  $t, x_1, x_2, \dots, x_n$ , ktorá je definovaná a spojitá na množine  $M = [a, b] \times R^n$  a PÚ daná v tvare

$$\mathbf{x}'(t) = \mathbf{f}(t, \mathbf{x}(t)), \quad t \in [a, b] \quad (3.10)$$

$$\mathbf{x}(a) = \mathbf{x}_0. \quad (3.11)$$

má na  $[a, b]$  práve jedno riešenie. Potom funkcia  $\mathbf{f}(t, \mathbf{x})$  spĺňa na množine  $M$  *Lipschitzovu podmienku*, ak existuje taká konštanta  $L > 0$ , že platí

$$\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq L \|\mathbf{x} - \mathbf{y}\| \quad (3.12)$$

## 3.2 Počiatkové úlohy pre obyčajné diferenciálne rovnice

pre  $\forall t \in [a, b]$  a  $\forall \mathbf{x}, \mathbf{y}$ , kde  $\|\cdot\|$  je norma v priestore  $R^n$  (pre  $n = 1$  je to absolútna hodnota).

Lipschitzova podmienka spolu so spojitosťou funkcie  $\mathbf{f}(t, \mathbf{x})$  zaručujú na množine  $M$  existenciu a jednoznačnosť riešenia PÚ na celom  $[a, b]$ . K tomu postačuje aby na  $M$  existovali spojité a ohraničené parciálne derivácie  $\frac{\partial f_i}{\partial x_j}, i, j = 1, 2, \dots, n$ .

**Príklad:** Majme PÚ

$$x' = \frac{1}{1+x^2}, \quad x(a) = x_0. \quad (3.13)$$

Ukážme, že táto PÚ má na ľubovoľnom intervale  $[a, b]$  práve jedno riešenie.

*Riešenie:* Máme teda  $f(t, x) = \frac{1}{1+x^2}$ ,  $x \in R$ . Je zrejmé, že  $f$  je spojitá. Využívajúc vetu o strednej hodnote dostávame

$$f(t, x) - f(t, y) = \frac{\partial f(t, \xi)}{\partial x} (x - y), \quad (3.14)$$

kde  $\xi \in (x, y)$  a  $\frac{\partial f}{\partial x} = -\frac{2x}{(1+x^2)^2}$ . Na splnenie Lipschitzovej podmienky stačí aby existovalo také  $L$  že  $|\partial f / \partial x| \leq L$  pre  $\forall t, x \in R$ . Vyšetrením priebehu funkcie  $|\partial f / \partial x|$  zistíme že má maximum v bode  $x_0 = 3\sqrt{3}/8$ . Funkcia  $f(t, x) = \frac{1}{1+x^2}$  teda spĺňa Lipschitzovu podmienku s konštantou  $L = 3\sqrt{3}/8$  a teda PÚ má jediné riešenie na ľubovoľnom intervale  $[a, b]$ .

### 3.2.1 Podmienenosť

Pod podmienenosťou PÚ budeme rozumieť vplyv nepresností a rôznych perturbácií vo vstupných dátach, ako napr. chyby merania počiatkovej hodnoty  $x_0$  alebo aproximáciu funkcie  $f$ , na výstupné dáta. Uvažujme jednoduchý prípad keď PÚ je daná *lineárnou* diferenciálnou rovnicou

$$x' = p(t)x + q(t), \quad (3.15)$$

kde  $p, q$  sú funkcie spojité na  $[a, b]$ . Riešenie takejto rovnice vieme vyjadriť explicitne v tvare

$$x(t) = x_0 \exp\left(\int_a^t p(r) dr\right) + \int_a^t q(r) \exp\left(\int_r^t p(s) ds\right) dr. \quad (3.16)$$

Predpokladajme, že počiatková hodnota je určená s chybou  $\Delta x_0$  a pravá strana rovnice  $f(t, x) = p(t)x + q(t)$  s chybou  $\delta(t)$ . Dostávame tak porušenú úlohu a jej riešenie v tvare

$$\tilde{x}(t) = (x_0 + \Delta x_0) \exp\left(\int_a^t p(r) dr\right) + \int_a^t [q(r) + \delta(r)] \exp\left(\int_r^t p(s) ds\right) dr. \quad (3.17)$$

### 3.3 Numerické metódy riešenia

---

V prípade, že  $\frac{\partial f(t,x)}{\partial x} = p(t) < 0$  je vplyv porúch slabý a také PÚ sa nazývajú *stabilné*. PÚ je *dobře podmienená* ak pre absolútnu odchýlku riešenia presnej a porušenej PÚ platí:

$$|x(t) - \tilde{x}(t)| \leq M|\Delta x_0| + N \int_a^t |\delta(r)| dr, \quad (3.18)$$

kde  $M \leq 1, N \leq 1$ .

V prípade že  $\frac{\partial f(t,x)}{\partial x} = p(t) \geq 0$ , chybu môžeme odhadnúť rovnakým spôsobom, avšak hodnoty  $M, N$  budú vo všeobecnosti väčšie ako 1.

Keď PÚ je daná *nelineárnou* diferenciálnou rovnicou, jej stabilita je taktiež určená znamienkom  $\frac{\partial f(t,x)}{\partial x}$ , avšak v tomto prípade už nie je len funkciou  $t$  ale aj  $x$ . Situácia je komplikovanejšia v prípade sústavy rovníc, ktorá môže mať stabilné aj nestabilné zložky. Stabilita však nemusí byť podstatná pre veľkosť relatívnej chyby, ako si to ukážeme v nasledujúcom príklade.

**Príklad:** Dá sa ľahko overiť že  $x(t) = \exp(\lambda t)$  je riešením PÚ  $x' = \lambda x, x(0) = 1$  pre  $t \in R$ . Preskúmajte vplyv poruchy v počiatočnej podmienke na stabilitu a presnosť riešenia.

*Riešenie:* Riešením porušenej úlohy  $\tilde{x}' = \lambda \tilde{x}, \tilde{x}(0) = 1 + \epsilon$  je  $\tilde{x}(t) = (1 + \epsilon) \exp(\lambda t)$ . Pre absolútnu chybu platí  $\tilde{x}(t) - x(t) = \epsilon \exp(\lambda t)$ , čo znamená že absolútna chyba sa zosilňuje pre  $\lambda > 0$  a zoslabuje pre  $\lambda < 0$ . Avšak, relatívna chyba  $(\tilde{x}(t) - x(t))/x(t) = \epsilon \exp(\lambda t)/\exp(\lambda t) = \epsilon$  je konštatná aj pre nestabilný prípad.

### 3.3 Numerické metódy riešenia

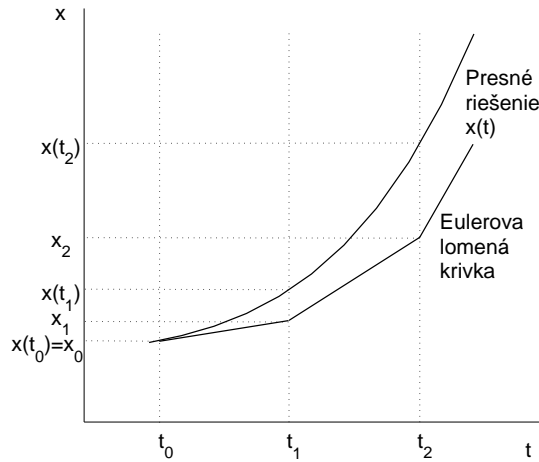
Numerické metódy riešenia PÚ sú užitočné v prípadoch, kedy analytické riešenie buď vôbec nie je možné získať alebo je príliš komplikované a nepraktické. V ďalšom sa obmedzíme na prípad riešenia jednej diferenciálnej rovnice, avšak spôsob riešenia je možné mechanicky zovšeobecniť na systém viacerých rovníc.

**Všeobecný postup riešenia PÚ** môžeme stručne naznačiť v nasledovných bodoch:

- **diskretizácia premenných** - generuje sa množina bodov  $t_0 = a, t_1, t_2, \dots$  (uzly siete) a riešenia v nich  $x(t_0), x(t_1), x(t_2), \dots$  sa aproximujú hodnotami  $x_0, x_1, x_2, \dots$ ,
- **vytvorenie siete** - voľba kroku  $h_n = t_{n+1} - t_n > 0$ , pričom ak  $h_n \equiv h$  dostávame sieť s konštantným a v opačnom prípade premenným krokom,
- **rekurentný vzťah** - aproximácia  $x_n$  v  $t_n$  sa počíta z hodnôt vypočítaných v predchádzajúcich krokoch (uzloch siete) pomocou rekurentného vzťahu,



### 3.3 Numerické metódy riešenia



Obr. 3.1: Grafické znázornenie Eulerovej metódy.

- ***k*-kroková metóda** - ak  $x_{n+1}$  je vyjadrené pomocou  $x_n, x_{n-1}, \dots, x_{n-k+1}$ , potom hovoríme o *k*-krokovej metóde,
- ***diferenčné metódy*** - metódy založené na vyššie opísanom postupe sa nazývajú *diferenčné metódy*.

#### 3.3.1 Eulerova metóda

Uvažujme PÚ

$$x'(t) = f(t, x), \quad a \leq t \leq b, \quad x(a) = \alpha. \quad (3.19)$$

Ddefinujme *sieťové uzly*  $t_n, n = 0, 1, 2, \dots, N$  rozmiestnené na  $[a, b]$  s premenným krokom  $h_n = t_{n+1} - t_n > 0$ . Hodnota  $x_{n+1}$  sa počíta extrapoláciou z hodnoty  $x_n$  v predchádzajúcom uzle a na intervale  $[t_n, t_{n+1}]$  sa riešenie aproximuje priamkou, ktorá prechádza bodom  $(t_n, x_n)$  a má smernicu  $x'_n = f(t_n, x_n)$ . Potom rovnica takejto priamky je

$$x = x_n + (t - t_n)f(t_n, x_n) \quad (3.20)$$

a pre  $t = t_{n+1}$  dostávame rekurentný vzťah

$$x_{n+1} = x_n + h_n f(t_n, x_n), \quad n = 0, 1, \dots \quad (3.21)$$

### 3.3 Numerické metódy riešenia

---

Je to jednokroková metóda a jej geometrický zmysel je nasledujúci. Riešenie  $x(t)$  je na  $[t_n, t_{n+1}]$  aproximované úsečkou so smernicou v začiatočnom bode intervalu a postupuje v tomto smere o celý krok  $h_n$  bez ohľadu na zmeny smerového poľa na tomto úseku, čo znamená nárast chyby v každom kroku. Takto postupujeme cez celý interval a dostávame tzv. Eulerovu lomenú čiaru.

Na Eulerovu metódu sa tiež môžeme pozeráť ako na aproximáciu hodnoty  $x(t_{n+1}) = x(t_n + h_n)$  pomocou Taylorovho polynómu 1. stupňa v bode  $t_n$ :

$$x(t_{n+1}) \approx x(t_n) + h_n x'(t_n) = x(t_n) + h_n f(t_n, x(t_n)). \quad (3.22)$$

Nakoniec, rekurentný vzťah môžeme tiež dostať z aproximačnej rovnice, ktorou nahradíme pôvodnú diferenciálnu rovnicu (3.19) diferenčnou s doprednou diferenciou

$$\frac{x_{n+1} - x_n}{h_n} \approx f(t_n, x_n), \quad n = 0, 1, \dots \quad (3.23)$$

#### Algoritmus Eulerovej metódy s konštantným krokom $h \equiv h_n$

**Vstup:**  $a, b, f(t, x), \alpha, N$

Položíme  $h = (b - a)/N$ ;  $t = a, x_0 = \alpha$ .

Pre  $n = 0, 1, 2, \dots, N - 1$  vypočítame

$$t_n = a + nh$$

$$x_{n+1} = x_n + hf(t_n, x_n)$$

**Výstup:**  $t_0 = a, t_1, \dots, t_N = b, x_0, x_1, \dots, x_N$ .

#### 3.3.2 Chyby numerického riešenia

Z predchádzajúceho výkladu by sa dalo intuitívne očakávať, že chyba Eulerovej metódy by sa mala so znižujúcim sa krokom  $h_n$  tiež znižovať. Ak zanedbáme zaokružovaciu chybu, tak je tomu naozaj tak. Zároveň tým však vzrastie počet uzlov siete, v ktorých potrebujeme počítať hodnoty  $f(x_n, t_n)$ . Zvýši sa tým pracnosť výpočtu a preto by sme chceli použiť také metódy, pri ktorých by celková chyba klesala so znižujúcim sa krokom  $\bar{h} = \max h_n$  čo najrychlejšie, napr. ako mocnina  $\bar{h}^p$  s dostatočne veľkým  $p$ .

#### Lokálna diskretizačná chyba

Lokálna diskretizačná chyba  $d_n$  na intervale  $[t_n, t_{n+1}]$  je nepresnosť, ktorej sa dopustíme pri výpočte  $x_{n+1}$  na jednom kroku, t. j. lokálne, za predpokladu, že všetky hodnoty  $x_j$ , potrebné k výpočtu  $x_{n+1}$  sú presné a že sa neprejavujú zaokružovacie chyby. Ináč povedané, je to nepresnosť s ktorou hodnoty teoretického riešenia spĺňajú rekurentný vzťah

$$x(t_{n+1}) = x(t_n) + h_n f(t_n, x(t_n)) + d_n. \quad (3.24)$$

### 3.3 Numerické metódy riešenia

---

Ak predpokladáme, že v  $t_n$  platí  $x(t_n) = x_n$ , potom pre lokálnu diskretizačnú chybu dostávame

$$d_n = x(t_{n+1}) - x_{n+1}. \quad (3.25)$$

#### Globálna diskretizačná chyba

Globálna diskretizačná chyba  $e_n$  je merítkom toho ako presne aproximuje postupnosť hodnôt  $x_n$  presné riešenie a teda

$$e_n = x(t_n) - x_n. \quad (3.26)$$

#### Konvergenca metódy

Pri takomto type aproximácie je žiadúce, aby približné riešenie  $x_n$  v nejakom vhodnom zmysle pri znižovaní kroku  $h$ , respektíve  $\bar{h} = \max h_n$  pri premennom kroku, konvergovalo k teoretickému riešeniu  $x$ . To by aspoň teoreticky zaručilo, že riešenie danej PÚ je možné obdržať s ľubovoľnou presnosťou, ak vynaložíme dostatočné úsilie, t. j. ak zvolíme dostatočne malé  $\bar{h}$ . Znamená to, že vyžadujeme aby platilo

$$x_n \rightarrow x(t) \text{ pre } h \rightarrow 0 \quad (3.27)$$

pri konštantnom kroku a

$$x(t, \bar{h}) \rightarrow x(t) \text{ pre } \bar{h} \rightarrow 0 \quad (3.28)$$

pri premennom kroku, kde  $x(t, \bar{h})$  je približné riešenie získane pri kroku  $\bar{h}$ . Ak sú splnené vyššie uvedené podmienky pre všetky PÚ s funkciami  $f$ , ktoré sme popísali na začiatku, potom sa jedná o konvergentnú metódu.

#### Rád metódy

Pri konvergentnej metóde očakávame, že pri každom pevnom  $n$  pre  $\bar{h} \rightarrow 0$  budú aj lokálne diskretizačné chyby  $d_n$  konvergovať k nule. Rýchlosť tejto konvergenencie udáva rád metódy. Rád diferenčnej metódy je najväčšie prirodzené číslo  $p$  také, že pre danú metódu aplikovanú na ľubovoľnú PÚ s dostatočne hladkým riešením platí pri každom pevnom  $n$  a  $h_n \rightarrow 0$  odhad

$$d_n = O(h_n^{p+1}). \quad (3.29)$$

Inými slovami, že existuje číslo  $c$ , nezávislé od  $h_n$  také, že pre malé  $h_n$  platí

$$|d_n| \leq ch_n^{p+1}, \quad (3.30)$$

### 3.3 Numerické metódy riešenia

---

a teda pre  $h_n \rightarrow 0$  sú lokálne chyby  $d_n$  rádovo veľkosti  $h_n^{p+1}$ .

Pri jednokrokových metódach sa dá ukázať, že ak je metóda aspoň 1. rádu, tak je konvergentná a pre globálnu chybu s dostatočne hladkou pravou stranou  $f$  sa dá dokázať, že pre  $\bar{h} \rightarrow 0$  platí

$$e(t, \bar{h}) = O(\bar{h})^p. \quad (3.31)$$

Z tohto dôvodu hovoríme, že metóda je  $p$ -tého rádu. Vo viackrokových metódach vysoký rád všeobecne nezaručuje ani malú globálnu chybu ani konvergenciu metódy, avšak u veľa metód (vrátane tých, ktoré si uvedieme v ďalšom) platí to, čo v jednokrokových metódach.

**Príklad:** Rád Eulerovej metódy.

Z vyššie uvedeného platí (berúc do úvahy tiež vzťah  $x'(t_n) = f(t_n, x(t_n))$ ) že diskretizačná chyba sa dá vyjadriť ako

$$d_n = x(t_{n+1}) - x(t_n) - h_n x'(t_n). \quad (3.32)$$

Ak rozvineme  $x(t_{n+1})$  pomocou Taylorovho polynómu v  $t_n$ , za predpokladu existencie druhej derivácie  $x$  na intervale  $[t_n, t_{n+1}]$ , dostávame

$$x(t_{n+1}) = x(t_n + h_n) = x(t_n) + h_n x'(t_n) + \frac{1}{2} h_n^2 x''(\xi), \quad \xi \in [t_n, t_{n+1}]. \quad (3.33)$$

Z toho vyplýva, že

$$d_n = \frac{1}{2} h_n^2 x''(\xi), \quad \xi \in [t_n, t_{n+1}]. \quad (3.34)$$

Ak je  $x''$  na  $[a, b]$  ohraničená, potom  $d_n = O(h_n^2)$  a teda Eulerova metóda je 1. rádu.

**Poznámka:** Vyššie uvedené úvahy platia pre PÚ s dostatočne hladkým riešením ( $x''$  musí existovať a byť ohraničená na  $[a, b]$ ). V opačnom prípade bude konvergencia pomalšia a rád metódy nižší.

#### Zaokružľovacie chyby

Uvažujme Eulerovu metódu s konštantným krokom. Nech  $\epsilon$  je maximálna zaokružľovacia chyba na jednom kroku. Pre skutočné hodnoty  $x_n^*$  približného riešenia  $x_n$  platí

$$x_{n+1}^* = x_n^* + hf(t_n, x_n^*) + \epsilon, \quad n = 0, 1, \dots \quad (3.35)$$

Potom celková zaokružľovacia chyba v nejakom  $t_k$  bude

$$k\epsilon = \epsilon(t_k - a)/h. \quad (3.36)$$

### 3.3 Numerické metódy riešenia

---

Z predchádzajúceho vyplýva, že globálna diskretizačná chyba Eulerovej metódy je  $|x(t) - x(t, h)| \approx Kh$ . Ak skutočne vypočítaná hodnota je  $x^*(t, h)$ , tak odhad zaokružľovacej chyby môžeme vyjadriť ako

$$|x(t, h) - x^*(t, h)| \approx \epsilon(t - a)/h \quad (3.37)$$

a celkovú chybu výpočtu ako

$$|x(t) - x^*(t, h)| \approx Kh + \epsilon(t - a)/h \equiv g(h). \quad (3.38)$$

Vyšetrovaním priebehu funkcie  $g(h)$  ľahko zistíme, že funkcia nadobúda minimum v  $h = h_0 = \sqrt{\epsilon(t - a)/K}$  a teda celková chyba je minimálna (optimálna) pre  $h_{opt} \approx h_0$ . To znamená, že pri ďalšom znižovaní kroku bude celková chyba narásť kvôli narastaniu zaokružľovacej chyby. Pri metódach vyššieho ( $p$ -tého) rádu postupujeme podobne s tým, že člen  $Kh$  nahradíme  $Kh^p$ ,  $p > 0$ . Aby sme však dosiahli vysokú celkovú presnosť, musíme dbať na to, aby aj druhý člen bol dostatočne malý a teda musíme výpočty robiť s presnosťou na dostatočne veľa desatinných miest.

#### Odhad chyby metódou polovičného kroku - Richardsonova extrapolácia

V prípade ak PÚ riešime diferenčnou metódou  $p$ -tého rádu s konštantným krokom  $h$  a úloha má dostatočne hladké riešenie, tak potom je možné globálnu chybu asymptoticky pre  $h \rightarrow 0$  vyjadriť v tvare

$$e(t, h) = a_0(t)h^p + a_1(t)h^{p+1} + \dots, \quad (3.39)$$

kde koeficienty  $a_i(t)$  nezávisia na  $h$ . Túto vlastnosť môžeme využiť na odhad globálnej diskretizačnej chyby v bode  $t$ . Využijúc toho, že  $e(t, h) = a_0(t)h^p + O(h^r)$ ,  $r > p$ , môžeme postupovať podobne ako pri Richardsonovej extrapolácii popísanej v numerických metódach pre spresnenie výpočtu určitého integrálu lichobežníkovou metódou. Konkrétne, hodnotu globálnej diskretizačnej chyby  $e(t, h/2)$  je možné vyjadriť pomocou dvoch hodnôt približného riešenia  $x(t, h)$  a  $x(t, h/2)$  ako

$$e(t, h/2) = \frac{x(t, h/2) - x(t, h)}{2^p - 1} + O(h^r), \quad r > p, \quad (3.40)$$

čo znamená, že pre dostatočne malé  $h$

$$e(t, h/2) \approx \frac{x(t, h/2) - x(t, h)}{2^p - 1}. \quad (3.41)$$

Teda pre odhad chyby potrebujeme len dva odhady riešenia pre rôzne kroky a rád metódy.

### 3.3 Numerické metódy riešenia

---

## 3.4 Jednokrokové metódy

V predchádzajúcom odseku sme popísali jednoduchú Eulerovu metódu, ktorá je jednokroková, 1. rádu a lineárna vzhľadom na  $f(t_n, h_n)$ . Je zrejmé, že lepšiu aproximáciu dostaneme, ak sa vzdáme linearity na  $f(t_n, h_n)$ . Potom rekurentný vzťah pre takúto všeobecnú jednokrokovú metódu bude mať tvar

$$x_{n+1} = x_n + h_n \Phi(t_n, x_n, h_n, f), \quad n = 0, 1, \dots, \quad (3.42)$$

kde  $\Phi(t_n, x_n, h_n, f)$  je vo všeobecnosti nejaká prírastková funkcia.

### 3.4.1 Metóda Taylorovho typu

Ako jednu z interpretácií Eulerovej metódy sme si uviedli aproximáciu Taylorovým polynómom 1. stupňa. Ak vezmeme aj vyššie stupne, tak pre dostatočne hladkú funkciu  $x$  platí

$$\begin{aligned} x(t_{n+1}) &= x(t_n + h_n) = x(t_n) + h_n x'(t_n) + \frac{h_n^2}{2!} x''(t_n) + \dots \\ &\quad + \frac{h_n^p}{p!} x^{(p)}(t_n) + \frac{h_n^{p+1}}{(p+1)!} x^{(p+1)}(\xi), \quad \xi \in [t_n, t_{n+1}], \end{aligned} \quad (3.43)$$

kde  $x'(t) = f(t, x(t))$ ,  $x''(t) = f_t + f_x x' = f_t + f_x f, \dots$ ,  $x^{(r+1)}(t) = f^{(r)}(t, x(t))$ .

Po dosadení do vzťahu (3.43) dostávame

$$x(t_{n+1}) = x(t_n) + h_n f(t_n, x(t_n)) + \dots + \frac{h_n^p}{p!} f^{(p-1)}(t_n, x(t_n)) + h_n^{p+1} R_{p+1}(t_n), \quad (3.44)$$

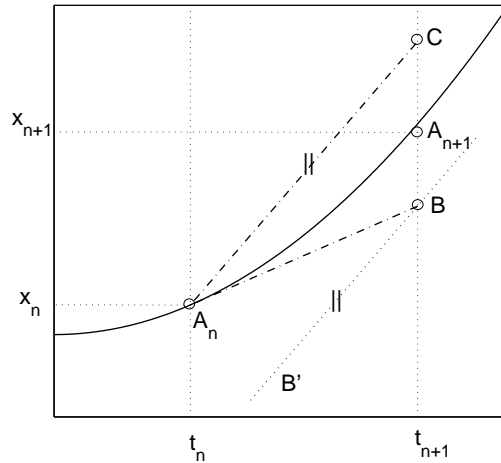
kde  $R_{p+1}(t_n) = x^{(p+1)}(\xi_n)/(p+1)!$ . Zanedbaním zvyšku  $R$  dostávame rekurentný vzťah pre približné hodnoty. Nevýhodou je však pracnosť počítania derivácií vyšších rádo  $f^r$  a ich hodnôt.

### 3.4.2 Metódy typu Runge-Kutta (RK)

Na rozdiel od metódy Taylorovho typu, metódy typu Runge-Kutta nevyžadujú explicitné počítanie derivácií  $f^r$  a ich hodnôt, ale aproximujú ich kombináciou hodnôt  $f$  v niekoľkých strategických bodoch na intervale  $[t_n, t_{n+1}]$ . Všeobecne sú dané rekurentným vzťahom

$$x_{n+1} = x_n + h_n \sum_{i=1}^r \alpha_i k_i, \quad n = 0, 1, \dots, \quad (3.45)$$

### 3.3 Numerické metódy riešenia



**Obr. 3.2:** Geometrická interpretácia Heunovej metódy.

kde  $k_1 = f(t_n, x_n)$  a  $k_i = f(t_n + \lambda_i h_n, x_n + \mu_i h_n k_{i-1})$  pre  $i > 1$ . Hodnoty  $k_i$  je treba počítať v každom kroku a teda náročnosť výpočtu je  $r$  vyčíslení funkčných hodnôt  $f$  v každom kroku.

#### Metóda Runge-Kutta 2. rádu (RK2)

Uvažujme dve možnosti voľby parametrov:

- **Modifikovaná Eulerova metóda:** Pre  $r = 2, \alpha_1 = 0, \alpha_2 = 1, \lambda_2 = \mu_2 = 1/2$  dostávame rekurentný vzťah:

$$x_{n+1} = x_n + h_n k_2, \text{ kde} \\ k_2 = f\left(t_n + \frac{h_n}{2}, x_n + \frac{h_n}{2} k_1\right) \text{ a } k_1 = f(t_n, x_n).$$

- **Heunova metóda:** Pre  $r = 2, \alpha_1 = \alpha_2 = 1/2, \lambda_2 = \mu_2 = 1$  dostávame rekurentný vzťah:

$$x_{n+1} = x_n + h_n \frac{k_1 + k_2}{2}, \text{ kde} \\ k_2 = f(t_{n+1}, x_n + h_n k_1) \text{ a } k_1 = f(t_n, x_n).$$

Geometrický význam Heunovej metódy môžeme vidieť na obr. 3.2. Priamka  $A_n B$  má smernicu  $k_1 = f(t_n, x_n)$  a teda  $B = [t_{n+1}, x_n + h_n k_1]$ . Priamka  $BB'$  je dotyčnica k riešeniu rovnice  $x' = f(t, x)$ , ktorá prechádza bodom  $B$ , a priamka  $A_n C \parallel BB'$  má smernicu  $k_2 = f(t_{n+1}, x_n + h_n k_1)$ . Potom  $C = [t_{n+1}, x_n + h_n k_2]$  a výsledný bod  $A_{n+1}$  je stredom úsečky  $BC$ . Hodnotu  $f$  počítame dva-krát a to v bodoch  $A_n$  a  $B$ .

### 3.3 Numerické metódy riešenia

---

#### Metóda Runge-Kutta 4. rádu (RK4)

Rekurentný vzťah metódy RK4 je

$$x_{n+1} = x_n + h_n \frac{k_1 + 2k_2 + 2k_3 + k_4}{6}, \quad (3.46)$$

kde

$$\begin{aligned} k_1 &= f(t_n, x_n), \\ k_2 &= f\left(t_n + \frac{h_n}{2}, x_n + \frac{h_n}{2}k_1\right), \\ k_3 &= f\left(t_n + \frac{h_n}{2}, x_n + \frac{h_n}{2}k_2\right), \\ k_4 &= f(t_{n+1}, x_n + h_n k_3). \end{aligned}$$

Odvodenie vzťahov pre RK metódy je technicky pomerne zložité, ale princíp je následovný. Snažíme sa nájsť také hodnoty parametra  $r$  a koeficientov  $\alpha_i$ ,  $i = 1, 2, \dots, r$ , a  $\lambda_i, \mu_i$ ,  $i = 2, 3, \dots, r$ , aby prvých  $p + 1$  členov Taylorovho rozvoja výrazu (3.45) po dosadení hodnôt presného riešenia  $x$  za  $x_n$  a  $x_{n+1}$  súhlasilo s prvými  $p + 1$  členmi vo vzťahu (3.44). Takto dostaneme metódu rovnakého rádu ako v metóde Taylorovho typu (teda  $p$ ) bez počítania derivácií  $f^{(r)}$  a ich vyčísľovania.

**Odhad chyby aproximácie:** Pre vyššie uvedené metódy 2. a 4. rádu je dokázané, že asymptoticky pre  $h \rightarrow 0$  sa dajú globálne chyby vyjadriť v tvare

$$e(t, h) = a_0(t)h^p + a_1(t)h^{p+1} + \dots, \quad (3.47)$$

s  $p = 2$  a  $4$  a metódou polovičného kroku dostaneme

$$e(t, h/2) \approx \frac{x(t, h/2) - x(t, h)}{2^p - 1} \left( = \frac{x(t, h/2) - x(t, h)}{15} \text{ pre RK4} \right). \quad (3.48)$$

Nevýhodou je značná pracnosť vyčísľovania odhadov riešení pre obidva kroky, t. j.  $x(t, h/2)$  aj  $x(t, h)$ , čo je možné obísť použitím iných metód typu RK, napr. Fehlbergove metódy, kde  $k_i$  nezávisí len od  $k_{i-1}$  ale tiež od  $k_{i-2}, \dots, k_1$ .

## 3.5 Viackrokové metódy

### 3.5.1 Všeobecná lineárna $k$ -kroková metóda

Pre výpočet hodnoty  $x_{n+1}$  sa využijú nielen hodnoty v  $t_n$ , teda  $x_n$ , ale aj hodnoty v predchádzajúcich uzloch siete  $x_{n-1}, x_{n-2}, \dots, f_{n-1}, f_{n-2}, \dots$ . V ďalšom sa pre jednoduchosť obmedzíme na konštantný krok  $h$ .  $k$ -krokové metódy sú založené na rekurentnom



### 3.3 Numerické metódy riešenia

---

vzťahu

$$x_{n+1} = \sum_{i=1}^k a_i x_{n+1-i} + h \sum_{i=0}^k b_i f_{n+1-i}, \quad (3.49)$$

kde  $k$  je prirodzené číslo,  $a_i, i = 1, 2, \dots, k$ ,  $b_i, i = 0, 1, \dots, k$ , sú konštanty a buď  $a_n \neq 0$  alebo  $b_k \neq 0$ . Eulerova metóda ( $k = 1$ ) je špeciálnym prípadom, kde  $x_{n+1} = x_n + hf_n$ . K zahájeniu výpočtu potrebujeme poznať hodnoty  $x_0, x_1, \dots, x_{k-1}$  a z nich vypočítané  $f_0, f_1, \dots, f_{k-1}$ , napr. pomocou jedнокrokovovej RK metódy. Potom hodnotu  $x_{n+1}$  vypočítame zo známych hodnôt  $x_n, x_{n-1}, \dots, x_{n+1-k}, f_n, f_{n-1}, \dots, f_{n+1-k}$ . Viac-krokové metódy rozlišujeme na explicitné a implicitné.

- **Explicitné metódy:** vo vzťahu (3.49) platí  $b_0 = 0$  a teda dostávame priamu rekurenciu, podobnú ako v Eulerovej metóde.
- **Implicitné metódy:** vo vzťahu (3.49) platí  $b_0 \neq 0$  a teda  $x_{n+1}$  nie je možné vypočítať z rovnice priamo a po odštartovaní je potrebné v každom kroku riešiť rovnicu

$$x_{n+1} = hb_0 f(t_{n+1}, x_{n+1}) + g, \quad (3.50)$$

kde  $g = \sum_{i=1}^k a_i x_{n+1-i} + hb_i f_{n+1-i}$  je daná už vypočítanými hodnotami. Ak je riešená PÚ lineárna, tak aj rovnica (3.50) je lineárna a teda riešenie je bezproblémové. Ak je však riešená PÚ nelineárna, tak sa za splnenia predpokladov kladených na PÚ, uvedených v odstavci 3.2, dá ukázať, že pre  $h < 1/(L|b_0|)$ , kde  $L$  je Lipschitzova konštanta funkcie  $f$ , má rovnica práve jedno riešenie, ktoré je možné nájsť, napr. metódou prostej iterácie. Implicitné metódy sú teda pracnejšie, ale pri danom  $k$  dosahujú väčšej presnosti než explicitné a navyše majú výhodnejšie vlastnosti z hľadiska stability výpočtu.

*Rád nelineárnej k-krokovovej metódy* je možné určiť z koeficientov  $a_i, b_i$  a lokálna diskreditačná chyba je  $d_n = O(h^{p+1})$ . Metóda, ktorá je aspoň 1. rádu, sa nazýva *konzistentná*, čo je nevyhnutná podmienka na to aby bola metóda konvergentná. Pre  $k > 1$  to však nie je postačujúce, pretože na globálnu diskretizačnú chybu majú vplyv okrem lokálnej chyby aj ďalšie faktory, ako napr. chyba v stanovení počiatočných hodnôt  $x_0, x_1, \dots, x_{k-1}$  a pri implicitných metódach tiež chyba riešenia rovnice (3.50) pre  $x_{n+1}$ . Za predpokladu, že chyba v stanovení  $x_{n+1}$  je zanedbateľná voči lokálnej diskretizačnej chybe a počiatočné hodnoty pre metódu  $p$ -tého rádu sú stanovené s presnosťou  $O(h^p)$ , je lineárna  $k$ -kroková metóda konvergentná a globálna chyba pre dostatočne malé  $h$  a úlohy s dostatočne hladkým riešením sú veľkosti  $O(h^p)$ .

### 3.3 Numerické metódy riešenia

---

Pri viackrokových metódach je okrem aproximovaného teoretického riešenia aj možnosť vzniku parazitných zložiek, ktoré nemajú s riešením rovnice nič spoločné. Ak pre  $h \rightarrow 0$  parazitné zložky vymiznú, hovoríme že metóda je *stabilná*. O stabilite metódy sa dá rozhodnúť na základe znalosti koeficientov metódy. V prípade nestabilných lineárnych  $k$ -krokových metód  $p$ -tého rádu, hodnoty  $\{x_n\}$  pre  $h \rightarrow 0$  všeobecne neaproximujú teoretické riešenie aj keď je lokálna chyba veľkosti  $O(h^{p+1})$ .

#### 3.5.2 Metódy založené na numerickej kvadratúre

##### Adamsove-Bashforthove (AB) metódy

Integrovaním diferenciálnej rovnice  $x'(t) = f(t, x(t))$  v medziach od  $t_{n+1-s}$  po  $t_{n+1}$  dostaneme

$$x(t_{n+1}) = x(t_{n+1-s}) + \int_{t_{n+1-s}}^{t_{n+1}} f(t, x(t)) dt. \quad (3.51)$$

Integrál na pravej strane aproximujeme nejakým interpolačným kvadratúrnym vzorcom s uzlami v niekoľkých bodoch siete  $t_0, t_1, \dots$ , kde  $t_i = t_0 + ih$ . Použitím Newtonových-Cotesových vzorcov, voľbou  $s = 2$  a aproximáciou integrálu obdĺžnikovým pravidlom dostávame

$$\int_{t_{n-1}}^{t_{n+1}} f(t, x(t)) dt \approx 2hf(t_n, x(t_n)). \quad (3.52)$$

Ak vzťah (3.52) dosadíme do (3.51) pre  $s = 2$  a zameníme  $x(t_i)$  hodnotami  $x_i$ , dostaneme rekurentný vzťah

$$x_{n+1} = x_{n-1} + 2hf_n, \quad (3.53)$$

čo nie je nič iné ako obdĺžniková metóda.

Viac sa využívajú  $k$ -krokové metódy, založené na kvadratúrnych vzorcoch pre  $s = 1$ :

$$x(t_{n+1}) = x(t_{n+1-s}) + \int_{t_n}^{t_{n+1}} f(t, x(t)) dt, \quad (3.54)$$

pričom sa použijú kvadratúrne vzorce s uzlami  $t_{n+1}, t_n, t_{n-1}, \dots, t_{n+1-k}$ , teda aj uzly čiastočne ležiace mimo oblasť integrácie. Predpokladajme, že v bodoch  $t_i = t_0 + ih$ ,  $i = 0, 1, \dots, n$  sme už vypočítali  $x_0, x_1, \dots, x_n$  a  $f_i = f(t_i, x_i)$  sú aproximáciami  $f(t, x(t))$ . Zostrojíme interpolačný polynóm  $p_{k-1}$  ( $k-1$ )-vého stupňa k dátam  $\{(t_i, f_i), i = n, n-1, \dots, n+1-k\}$ ,  $n+1 \geq k$  a použijeme ho ako aproximáciu funkcie  $f(t, x(t))$  na

### 3.3 Numerické metódy riešenia

---

$[t_n, t_{n+1}]$ . Takto získame explicitnú lineárnu  $k$ -krokovú metódu (keďže  $f_{n+1}$  nebude vystupovať a teda vlastne do  $t_{n+1}$  extrapolujeme) a dostávame

$$x_{n+1} = x_n + \int_{t_n}^{t_{n+1}} p_{k-1}(t) dt. \quad (3.55)$$

Ak  $p_{k-1}$  vyjadríme ako Newtonov interpolačný polynóm so spätnými diferenciami  $N_{k-1}^-(t)$ , dopracujeme sa k rekurentnému vzťahu

$$x_{n+1} = x_n + h(b_1^{(k)} f_n + b_2^{(k)} f_{n-1} + \dots + b_k^{(k)} f_{n+1-k}), \quad (3.56)$$

kde koeficienty  $b_i^{(k)}$  závisia na  $k$  a, napr. pre  $k = 1$ , dostávame Eulerovu metódu a pre Adamsovu-Bashforthovu metódu tretieho rádu ( $p = 3$ ) dostávame rekurentný vzťah

$$x_{n+1} = x_n + \frac{h}{12}(23f_n - 16f_{n-1} + 5f_{n-2}). \quad (3.57)$$

Je to explicitná metóda a dá sa ukázať že jej rád je  $k$ .

#### Adamsove-Moultonove (AM) metódy

Sú to implicitné metódy, ktoré namiesto extrapolácie polynómu využívajú interpoláciu na  $[t_n, t_{n+1}]$  a teda vstupujú tu aj hodnoty  $t_{n+1}, f_{n+1}$ . Na rozdiel od AB metódy, funkciu  $f(t, x(t))$  na  $[t_n, t_{n+1}]$  aproximujeme polynómom  $q_k$   $k$ -tého stupňa

$$x_{n+1} = x_n + \int_{t_n}^{t_{n+1}} q_k(t) dt. \quad (3.58)$$

Podobne  $q_k$  vyjadríme ako Newtonov interpolačný polynóm so spätnými diferenciami  $N_k^-(t)$  a dostaneme rekurentný vzťah

$$x_{n+1} = x_n + h(c_0^{(k)} f_{n+1} + c_1^{(k)} f_n + c_2^{(k)} f_{n-1} + \dots + c_k^{(k)} f_{n+1-k}). \quad (3.59)$$

V špeciálnych prípadoch pre  $k = 0$  je  $c_0^{(0)} = 1$  a dostávame implicitnú Eulerovu metódu a ak  $k = 1$  potom  $c_0^{(1)} = c_1^{(1)} = 1/2$  a dostávame lichobežníkové pravidlo.

Keďže je to implicitná metóda, riešime rovnicu

$$x_{n+1} = hc_0^{(k)} f(t_{n+1}, x_{n+1}) + g, \quad (3.60)$$

kde  $g = x_n + h(c_1^{(k)} f_n + c_2^{(k)} f_{n-1} + \dots + c_k^{(k)} f_{n+1-k})$ , metódou prostej iterácie (pričom  $g$  sa nemení). Postačujúcou podmienkou existencie, jednoznačnosti a konvergencie metódy je  $hL < 1/c_0^{(k)}$ , kde  $L$  je Lipschitzova konštanta. Vo všeobecnosti sa dá ukázať, že rád

### 3.3 Numerické metódy riešenia

---

$k$ -krokovej AM metódy je pre  $k > 1$  rovný  $k + 1$ , teda o jednotku vyšší ako pri AB metóde. To však nie je až taká výhoda, keďže pridaním jednej štartovacej hodnoty navyše, a teda použitím  $(k + 1)$ -krokovej metódy, môžeme rád AB metódy vyrovnáť s rádom AM metódy. Väčšou výhodou implicitnej AM metódy je to, že aj pri rovnakom ráde dáva presnejšie výsledky než AB metóda a navyše kladie oveľa menšie obmedzenia na krok  $h$  z dôvodu numerickej stability pri opakovanom vyčíslňovaní rekurencie metódy s pevným  $h$  a  $n \rightarrow \infty$ .

#### Metóda prediktor-korektor

Je to všeobecný algoritmus riešenia PÚ metódou AM (implicitnou), pričom k výpočtu  $x_{n+1}$  použijeme dve Adamsove metódy:

1. Najprv pomocou explicitnej AB metódy vypočítame počiatočnú hodnotu  $x_{n+1}^{[s]}$  pre  $s = 0$ , t. j.  $x_{n+1}^{[0]}$ . (prediktor P)
2. Vypočítame hodnotu pravej strany  $f_{n+1}^{[s]} = f(t_{n+1}, x_{n+1}^{[s]})$ . (vyhodnotenie E)
3. Pomocou implicitnej AM metódy vykonávame prostú iteráciu podľa vzorca (korektor C)

$$x_{n+1}^{[s+1]} = hc_0^{(k)} f(t_{n+1}, x_{n+1}^{[s]}) + g, \quad s = 0, 1, \dots \quad (3.61)$$

4. Ak je  $|x_{n+1}^{[s+1]} - x_{n+1}^{[s]}|$  väčšia ako zadaná tolerancia, zväčšíme  $s$  o jedna a pokračujeme v iterácii 2. Ináč  $x_{n+1} = x_{n+1}^{[s+1]}$  a prejdeme k ďalšiemu bodu siete.

V praxi namiesto požadovanej tolerancie sa ako kritérium ukončenia iterácii berú fixné hodnoty  $N \leq 2$ , čím dostávame rôzne algoritmy:  $P(EC)^N$ , resp.  $P(EC)^N E$ , kde posledné E vypočíta hodnotu  $f_{n+1} = f(t_{n+1}, x_{n+1}^{[N]})$  potrebnú pre výpočet  $x_{n+2}$ . Analýzy ukazujú že  $P(EC)^1 E$  je pri pevnom  $h$  a  $n \rightarrow \infty$  stabilnejší ako  $P(EC)^2 E$  a teda PECE je najrozumnejší algoritmus pre Adamsove metódy.

#### Stabilita metódy

Riešenie je nestabilné, ak akákoľvek chyba vzniknutá pri  $x_n$  vnáša do hľadaného riešenia nežiaducu zložku, ktorej vplyv sa postupne zväčšuje, až napokon celý výpočet znehodnotí.

**Príklad:** Nestabilita kôli zlej podmienenosti. Riešme pre  $t > 1$  počiatočnú úlohu

$$x' = x - \frac{1}{t^2} - \frac{1}{t} - 1, \quad x(1) = 2. \quad (3.62)$$

Presné riešenie je  $x(t) = 1/t + 1$ , avšak všeobecné riešenie je  $x(t) = A \exp(t) + 1/t + 1$ . Vplyvom zaokružovania a chýb metódy sa behom výpočtu objaví člen  $A \exp(t)$  s

### 3.5 Numerické metódy riešenia

---

malou ale nenulovou hodnotou  $A$ . Pre rastúce  $t$  člen  $A \exp(t)$  postupne preváži nad riešením.

## 3.6 Okrajové úlohy pre obyčajné diferenciálne rovnice

Obmedzme sa na lineárnu obyčajnú diferenciálnu rovnicu

$$a_2(t)x''(t) + a_1(t)x'(t) + a_0(t)x(t) = g(t), \quad (3.63)$$

kde  $a_2, a_1, a_0$  sú spojité funkcie a  $x$  je aspoň dvakrát diferencovateľná funkcia  $t$ . Riešme rovnicu (3.63) na intervale  $[a, b]$ , kde  $a, b$  sú kladné reálne čísla a  $a < b$  s okrajovými podmienkami

$$\alpha x(a) + \beta x'(a) = x_a, \quad (3.64)$$

$$\gamma x(b) + \delta x'(b) = x_b, \quad (3.65)$$

kde  $\alpha, \beta, \gamma, \delta, x_a, x_b$  sú dané reálne čísla, pričom  $|\alpha| + |\beta| \neq 0$  a  $|\gamma| + |\delta| \neq 0$ .

Takáto úloha je okrajová úloha (OÚ) a okrajové podmienky (3.64) a (3.65) môžu byť typu:

1. Dirichletove okrajové podmienky (podmienky 1. druhu)

$$x(a) = x_a, \quad x(b) = x_b, \quad (3.66)$$

2. Neumannove okrajové podmienky (podmienky 2. druhu)

$$x'(a) = x_a, \quad x'(b) = x_b, \quad (3.67)$$

3. Newtonove okrajové podmienky (podmienky 3. druhu) sú podmienky (3.64), (3.65) pre  $\alpha, \beta, \gamma, \delta \neq 0$ .

4. Zmiešané okrajové podmienky sú v oboch okrajoch intervalu iné, napr.  $x(a) = x_a, \quad x'(b) = x_b$ .

## 3.7 Numerické metódy riešenia

### 3.7.1 Metóda sietí

Medzi najjednoduchšie metódy patrí *metóda sietí* alebo *diferenčná metóda*. Riešenie hľadáme nie na celom intervale  $[a, b]$  ale v uzloch  $t_i$ . Pre *ekvidistančné uzly* platí  $h =$

### 3.5 Numerické metódy riešenia

---

$(b-a)/n$  a  $t_i = a + ih$ ,  $i = 0, 1, \dots, n$ . Body  $t_0 = a$  a  $t_n = b$  sú hraničné uzly, ostatné sú uzly vnútorné. Podstata prístupu spočíva v nahradení jednotlivých derivácií v rovnici diferenciami.

#### Aproximácia derivácií 1. rádu

Z Taylorovho rozvoja pre dostatočne hladkú funkciu  $x(t)$ ,  $t \in [a, b]$  platí:

$$x(t+h) = x(t) + x'(t)h + x''(t)\frac{h^2}{2} + x'''(\eta)\frac{h^3}{3!}, \quad (3.68)$$

$$x(t-h) = x(t) - x'(t)h + x''(t)\frac{h^2}{2} - x'''(\mu)\frac{h^3}{3!}. \quad (3.69)$$

Po odčítaní rovníc dostávame

$$x'(t) = \frac{x(t+h) - x(t-h)}{2h} + \frac{(x'''(\eta) + x'''(\mu))h^2}{3!}. \quad (3.70)$$

Nech  $t = t_i$ ,  $t+h = t_{i+1}$ ,  $t-h = t_{i-1}$ , potom pre dostatočne malé  $h$  a ohraničenú  $x'''$  dostávame

$$x'_i \approx \frac{x_{i+1} - x_{i-1}}{2h}, \quad (3.71)$$

teda derivácia v bode  $t_i$  je aproximovaná centrálnou diferenciou s chybou  $O(h^2)$ . Inú definíciu môžeme analogicky vytvoriť z rovnice (3.68) pri aproximácii derivácie doprednou diferenciou

$$x'_i \approx \frac{x_{i+1} - x_i}{h}, \quad (3.72)$$

kedy sa dopustíme chyby radu  $O(h)$ .

#### Aproximácia derivácií 2. rádu

Urobme rozvoj funkcie do Taylorovho radu až po 4. deriváciu. Potom, podobným spôsobom ako vyššie, dostávame

$$x''(t) = \frac{x(t+h) - 2x(t) + x(t-h)}{h^2} + \frac{(x^{(4)}(\eta) + x^{(4)}(\mu))h^2}{4!}. \quad (3.73)$$

Potom platí

$$x''_i \approx \frac{x_{i+1} - 2x_i + x_{i-1}}{h^2}, \quad (3.74)$$

### 3.5 Numerické metódy riešenia

---

a chyba aproximácie je  $O(h^2)$ .

**Príklad:** Numericky riešme obyčajnú diferenciálnu rovnicu 2. rádu s konštantnými koeficientami a Dirichletovými podmienkami

$$x'' + 2x' + x = t^3 + 6t^2 + 1; t \in (0, 2), \quad (3.75)$$

$$x(0) = 1, x(2) = 5. \quad (3.76)$$

*Riešenie:* Interval  $(0, 2)$  rozdelíme na  $n + 1$  častí veľkosti  $h = 2/(n + 1)$  a riešenie hľadáme v bodoch  $t_1 = h, t_2 = 2h, \dots, t_n = 2 - h$ , pri splnení Dirichletových podmienok v  $t_0 = 0$  a  $t_{n+1} = 2$ , t. j.  $x_0 = 1, x_{n+1} = 5$ . Pre ostatné uzly použijeme diferenčné rovnice

$$\frac{x_{i+1} - 2x_i + x_{i-1}}{h^2} + 2\frac{x_{i+1} - x_{i-1}}{2h} + x_i = t_i^3 + 6t_i^2 + 1. \quad (3.77)$$

Po úprave dostávame

$$\left(\frac{1}{h^2} - \frac{1}{h}\right)x_{i-1} + \left(1 - \frac{2}{h^2}\right)x_i + \left(\frac{1}{h^2} + \frac{1}{h}\right)x_{i+1} = t_i^3 + 6t_i^2 + 1 \quad (3.78)$$

a v maticovom tvare

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{F} \quad (3.79)$$

kde  $\mathbf{F} = [F_1, F_2, \dots, F_n]$  je vektor pravej strany a platí

$$F_i = t_i^3 + 6t_i^2 + 1, \quad i = 2, 3, \dots, n - 1 \quad (3.80)$$

ale

$$F_1 = t_1^3 + 6t_1^2 + 1 - \frac{x_0}{h^2} + \frac{x_0}{h} = t_1^3 + 6t_1^2 + 1 - \frac{1}{h^2} + \frac{1}{h}, \quad (3.81)$$

$$F_n = t_n^3 + 6t_n^2 + 1 - \frac{x_{n+1}}{h^2} - \frac{x_{n+1}}{h} = t_n^3 + 6t_n^2 + 1 - \frac{5}{h^2} - \frac{5}{h}, \quad (3.82)$$

príčom matica  $\mathbf{A}$  je tridiagonálna a má tvar

$$\mathbf{A} = \begin{bmatrix} 1 - \frac{2}{h^2} & \frac{1}{h^2} + \frac{1}{h} & 0 & \dots & 0 \\ \frac{1}{h^2} - \frac{1}{h} & 1 - \frac{2}{h^2} & \frac{1}{h^2} + \frac{1}{h} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \frac{1}{h^2} - \frac{1}{h} & 1 - \frac{2}{h^2} & \frac{1}{h^2} + \frac{1}{h} \\ 0 & \dots & 0 & \frac{1}{h^2} - \frac{1}{h} & 1 - \frac{2}{h^2} \end{bmatrix}.$$

Teda numerické riešenie úlohy dostaneme riešením sústavy lineárnych rovníc (3.79), napr. použitím niektorej z iteračných metód, ako sú Jacobiho alebo Gaussova-Seidelova.

# Kapitola 4

## Riešenie parciálných diferenciálnych rovníc

### 4.1 Úvod

Parciálne diferenciálne rovnice (PDR) sa vyskytujú takmer vo všetkých fyzikálnych situáciach kde sa veličiny menia v priestore alebo v priestore a čase, ako napr. v hydrodynamike, elektromagnetickom vlnení, difúzii, kvantovej mechanike, atď. V praxi sa len veľmi málo PDR dá riešiť analyticky. V ďalšom sa obmedzíme na numerické riešenie vo fyzike najbežnejších PDR 2. rádu:

- **parabolické** (Schrödingerova rovnica, rovnica difúzie)

**Príklad:** nestacionárna difúzna rovnica v 3D:

$$\frac{\partial f}{\partial t} = \kappa \left( \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right), \quad (4.1)$$

kde  $\kappa$  je difúzny koeficient.

- **eliptické** (Laplaceova a Poissonova rovnica)

**Príklad:** Poissonova rovnica v 3D:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = g(x, y, z), \quad (4.2)$$

príčom v prípade že  $g(x, y, z) = 0$  dostávame Laplaceovu rovnicu  $\Delta f = 0$ , kde  $\Delta = \nabla^2$ .

- **hyperbolické** (vlnová rovnica)

**Príklad:** vlnová rovnica v 3D:

$$\frac{\partial^2 f}{\partial t^2} = c^2 \left( \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right). \quad (4.3)$$



## 4.1 Úvod

---

### 4.1.1 Počiatočné a hraničné podmienky

Vo všeobecnosti sú dané podmienkou

$$\mathbf{f}(\mathbf{x}, t = 0) = \mathbf{F}(\mathbf{x}), \quad (4.4)$$

kde  $\mathbf{x} = (x, y, z)$ ,  $\mathbf{f}, \mathbf{F}$  sú vektory, pričom  $\mathbf{F}$  je známa funkcia. V takom prípade sa počet podmienok rovná dimenzií  $\mathbf{f}$  (pre každú zložku). Podobne ako u ODR, hraničné podmienky budeme rozlišovať na Dirichletove, Neumannove a zmiešané.

### 4.1.2 Diferenčné metódy

Myšlienka diferečných metód spočíva vo vytvorení siete (mriežky) bodov v doméne riešenia, vzdialených od seba o nejakú vzdialenosť  $\Delta \mathbf{x} = (\Delta x, \Delta y, \Delta z)$ , v ktorých sa bude aproximácia riešenia hľadať. V prípade časovej závislosti sa vytvorí postupnosť takýchto sietí, oddelených vzájomne o časový krok  $\Delta t$ . Sieť môže byť vytvorená v karteziánskych aj v iných krivočiarych súradniciach, v závislosti od geometrie domény riešenia (napr. ak hľadáme riešenie vo vnútri gule, potom budú hraničné podmienky popísané vo sférických súradniciach s počiatkom v strede gule, atď.). V ďalšom sa obmedzíme na karteziánske súradnice kvôli algebraickej jednoduchosti a lepšej numerickej stabilite. Po vypočítaní riešení v uzloch môžeme riešenie v ostatných bodoch odhadnúť interpoláciou alebo extrapoláciou. Pri diferenčných metódach budeme (podobne ako pri ODR) derivácie nahrádzať vhodnými diferenciami:

- *dopredná 1. diferenciacia*

$$f'_i \approx \frac{f_{i+1} - f_i}{h}, \quad (4.5)$$

- *spätná 1. diferenciacia*

$$f'_i \approx \frac{f_i - f_{i-1}}{h}, \quad (4.6)$$

- *centrálna 1. diferenciacia*

$$f'_i \approx \frac{f_{i+1} - f_{i-1}}{2h}, \quad (4.7)$$

- *centrálna 2. diferenciacia*

$$f''_i \approx \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2}. \quad (4.8)$$

Po substitúcii dostávame systém lineárnych alebo nelineárnych algebraických rovníc, viažúci premenné  $f_i$ .

## 4.1 Úvod

---

### 4.1.3 Konzistencia

Presnosť diferenčných metód závisí od veľkosti krokov v priestore  $\Delta \mathbf{x}$  a čase  $\Delta t$ . Sú to riadiace parametre numerického výpočtu. Ak sa postupným znižovaním  $\Delta \mathbf{x}$  a  $\Delta t$  (súčasne alebo ľubovoľným iným spôsobom) postupne zväčšuje presnosť priblíženia riešenia diferenčnej úlohy k pôvodnej PDR, potom hovoríme že metóda je *konzistentná*.

#### Overenie konzistencie

Predpokladáme, že všetky premenné v diferenčnej rovnici sú spojité funkcie priestoru a času a rozvineme ich do Taylorovho radu vo vybranom bode siete a časovom okamihu. Tým dostaneme tzv. *modifikovanú diferenciálnu rovnicu* (MDR). Ak pre  $\Delta \mathbf{x} \rightarrow \mathbf{0}$  a  $\Delta t \rightarrow 0$  platí že  $\text{MDR} \rightarrow \text{PDR}$ , potom metóda je konzistentná. Ináč povedané, ak je metóda konzistentná, tak rozdiel medzi riešeniami MDR a PDR je rádovo úmerný mocninám veľkosti  $\Delta \mathbf{x}$  a  $\Delta t$  a exponenty týchto mocnín udávajú *řád numerickej chyby metódy*. U štandardných diferenčných metód je konzistencia garantovaná.

### 4.1.4 Stabilita

Nech presné riešenie PDR s počiatočnou podmienkou nerastie s časom, ale je konštantné alebo klesá v každom bode. Potom, ak diferenčná metóda reprodukuje toto správanie bez artefaktných oscilácií, potom je stabilná. Nech presné riešenie PDR s počiatočnou podmienkou rastie s časom a ak diferenčná metóda rastie rýchlosťou rovnakou alebo menšou, tak je stabilná.

#### Overenie stability

1. *Von Neumannova metóda* - jednoduchá ale neberie do úvahy vplyv hraničných podmienok
2. *metóda projekčnej matice*

Bližší popis fungovania týchto metód si uvedieme v ďalšom pri parabolických PDR.

### 4.1.5 Konvergencia

Pri konvergentnej metóde pre  $\Delta \mathbf{x} \rightarrow \mathbf{0}$  a  $\Delta t \rightarrow 0$  numerické riešenie konverguje k presnému riešeniu.

**Laxova teória ekvivalencie:** garantuje, že ak numerické riešenie *lineárnej* PDR pomocou *konzistentnej* diferenčnej metódy je *stabilné*, potom v limite  $\Delta \mathbf{x} \rightarrow \mathbf{0}$  a  $\Delta t \rightarrow 0$  numerické riešenie konverguje k presnému riešeniu. Teda konzistencia a stabilita

## 4.2 Eliptické parciálne diferenciálne rovnice

---

zaručujú konvergenciu a naopak. U *nelineárnych* PDR je overovanie komplikovanejšie ale zo skúsenosti vyplýva že ak je diferenčná metóda konzistentná a lokálne stabilná tak je konvergentná.

## 4.2 Eliptické parciálne diferenciálne rovnice

### 4.2.1 Poissonova rovnica v 1-rozmernom priestore

Pre jednoduchosť si najprv uveďme Poissonovu rovnicu v 1D, ktorá má tvar

$$\frac{\partial^2 f}{\partial x^2} = g(x). \quad (4.9)$$

Aplikáciou diferenčnej metódy s krokom  $\Delta x$  dostávame aproximáciu

$$\frac{f_{i+1} - 2f_i + f_{i-1}}{(\Delta x)^2} = g_i, \quad (4.10)$$

kde  $f_i = f(x_i)$ ,  $g_i = g(x_i)$ ,  $i = 1, \dots, N - 1$ . Pre hraničné podmienky  $f(x_0) = f_0$  a  $f(x_N) = f_N$  dostávame

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -2 & 1 \\ 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f(x_0) \\ f(x_1) \\ f(x_2) \\ \vdots \\ f(x_{N-1}) \\ f(x_N) \end{bmatrix} = \begin{bmatrix} f_0 \\ (\Delta x)^2 g_1 \\ (\Delta x)^2 g_2 \\ \vdots \\ (\Delta x)^2 g_{N-1} \\ f_N \end{bmatrix} \quad (4.11)$$

alebo v maticovom zápise  $\mathbf{A}\mathbf{f} = \mathbf{B}$ . Formálne dostávame riešenie  $\mathbf{f} = \mathbf{A}^{-1}\mathbf{B}$ , avšak vzhľadom na to, že v praktických situáciách potrebujeme riešiť sústavu veľkého počtu lineárnych rovníc, je potrebné použiť iteračné numerické riešenie. S metódami riešenia sústav lineárnych rovníc sme sa bližšie oboznámili v kurze Numerické metódy, takže v ďalšom si len heslovite pripomenieme dve z nich.

### 4.2.2 Gaussova-Seidelova metóda

Riešime iteračne rovnicu (4.10) v každom uzle pomocou Gaussovej-Seidelovej metódy

$$f_i^{(m+1)} = \frac{f_{i+1}^{(m)} + f_{i-1}^{(m+1)} - g_i(\Delta x)^2}{2}. \quad (4.12)$$

**Postup:**

## 4.2 Eliptické parciálne diferenciálne rovnice

---

1. Priradíme počiatočné hodnoty každému  $f_i$  (napr. nulové) vnútri a hraničné hodnoty na hranici oblasti riešenia.
2. Prechádzame sieťou postupne, napr. od  $f_1$ , a aplikujeme rovnicu (4.12) pre každé  $f_i$ .
3. Po veľa prejdení zistíme, že sa hodnoty  $f_i$  medzi jednotlivými prejdeniami už veľmi nemenia a vtedy môžeme výpočet ukončiť, pričom posledné hodnoty vezmeme ako riešenia.

### 4.2.3 Metóda postupnej relaxácie

Gaussova-Seidelova iteratívna metóda je dosť neefektívna. Jej modifikovaním dostávame relaxačnú (alebo superrelaxačnú) metódu založenú na iteračnom vzťahu

$$f_i^{(m+1)} = \alpha \hat{f}_i + (1 - \alpha) f_i^{(m)}, \quad (4.13)$$

kde  $\hat{f}_i$  je vypočítané z rovnice (4.12) a  $m$  je iteračný index. Parameter  $\alpha$  sa nazýva relaxačný parameter a metóda konverguje ak platí  $0 < \alpha < 2$ , pričom v prípade  $0 < \alpha < 1$  hovoríme o podrelaxácii a v prípade  $1 < \alpha < 2$  o nadrelaxácii. Optimálna hodnota  $\alpha$  závisí od konkrétneho problému a nie je jednoduché ho stanoviť analyticky.

### 4.2.4 Poissonova rovnica v 2-rozmernom priestore

Riešme rovnicu

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = g(x, y) \quad (4.14)$$

s nejakými okrajovými podmienkami.

**Postup:**

- Najprv zostrojíme sieť  $N \times N$  bodov (uzlov) navzájom oddelených o  $\Delta x$  v smere osi  $x$  a  $\Delta y$  v smere osi  $y$  (vid'. obr.4.1).
- Nahradením druhých parciálnych derivácií príslušnými centrálnymi diferenciami dostávame diferenčnú rovnicu

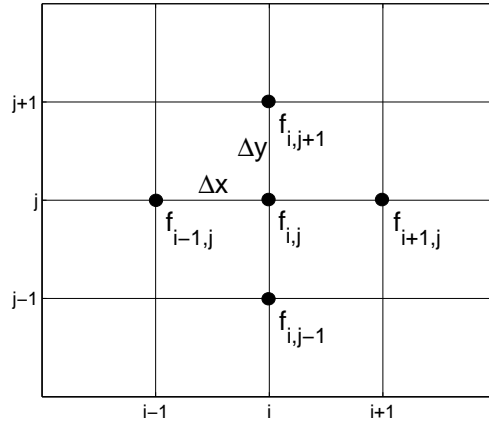
$$\frac{f_{i+1,j} - 2f_{i,j} + f_{i-1,j}}{(\Delta x)^2} + \frac{f_{i,j+1} - 2f_{i,j} + f_{i,j-1}}{(\Delta y)^2} = g_{i,j}, \quad (4.15)$$

kde  $f_{i,j} = f(x_i, y_j)$ ,  $g_{i,j} = g(x_i, y_j)$ ,  $i = 1, \dots, N-1$ . Po preusporiadaní dostávame

$$f_{i,j} = \frac{-g_{i,j}(\Delta x)^2(\Delta y)^2 + (\Delta x)^2(f_{i,j+1} + f_{i,j-1}) + (\Delta y)^2(f_{i+1,j} + f_{i-1,j})}{2((\Delta x)^2 + (\Delta y)^2)}. \quad (4.16)$$

### 4.3 Parabolické parciálne diferenciálne rovnice

---



Obr. 4.1: Diskretizácia priestorových súradníc v 2-rozmernom priestore.

- Pri použití Gaussovej-Seidelovej metódy:
  1. Položíme  $f_{i,j} = 0$  pre všetky  $i, j$ .
  2. Iterujeme  $f_{i,j}^{(m+1)} = \frac{-g_{i,j}(\Delta x)^2(\Delta y)^2 + (\Delta x)^2(f_{i,j+1}^{(m)} + f_{i,j-1}^{(m)}) + (\Delta y)^2(f_{i+1,j}^{(m)} + f_{i-1,j}^{(m)})}{2((\Delta x)^2 + (\Delta y)^2)}$ .

### 4.3 Parabolické parciálne diferenciálne rovnice

Na rozdiel od eliptických PDR, pri ktorých sa udávajú okrajové (hraničné) podmienky, pri parabolických PDR sa udáva informácia o funkcii v počiatočnom čase a zisťujeme ako sa riešenie vyvíja v čase. Typické parabolické rovnice vo fyzike zahŕňajú:

- Nestacionárna difúzia (v 1D)

$$\frac{\partial f}{\partial t} - \kappa \frac{\partial^2 f}{\partial x^2} = 0. \quad (4.17)$$

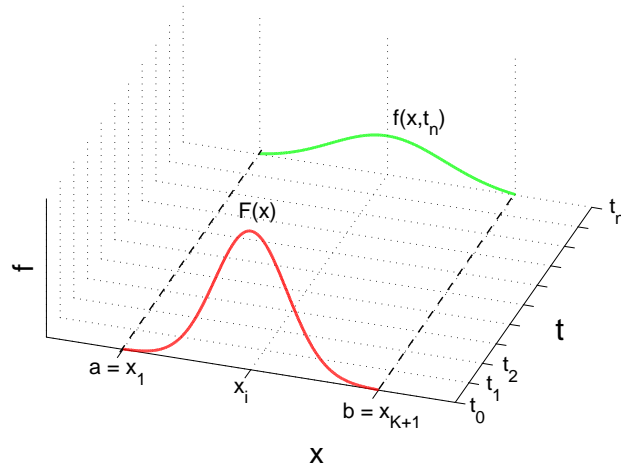
- Časová Schrödingerova rovnica

$$i\hbar \frac{\partial f}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 f}{\partial x^2} + Vf. \quad (4.18)$$

V ďalšom sa obmedzíme na rovnicu difúzie v 1D

$$\frac{\partial f}{\partial t} = \kappa \frac{\partial^2 f}{\partial x^2}, \quad (4.19)$$

### 4.3 Parabolické parciálne diferenciálne rovnice



**Obr. 4.2:** Diskretizácia v priestore a čase pre riešenie rovnice difúzie v 1-rozmernom priestore. Počiatočné hodnoty funkcie majú vplyvom difúzie tendenciu s časom sa postupne znižovať a zaberáť stále širší interval pozdĺž osi  $x$ .

s počiatočnou podmienkou  $f(x, t = 0) = F(x)$ , kde  $F(x)$  je známa funkcia a  $\kappa$  je difúzna konštanta, a homogennými podmienkami pre  $x \rightarrow \pm\infty$ , t. j.  $f(x = \pm\infty, t) = 0$ . V tomto prípade je známe aj analytické riešenie, ktoré má tvar

$$f(x, t) = \frac{1}{\sqrt{4\kappa\pi t}} \int_{-\infty}^{\infty} F(x + u) \exp\left(-\frac{u^2}{4\kappa t}\right) du. \quad (4.20)$$

#### 4.3.1 Numerické riešenie

**Postup:**

1. Vytvoríme sieť (vid' obr. 4.2) a hľadáme riešenie v uzloch siete  $(x_i, t_i)$  na oblasti  $[a, b] \times [0, \infty)$ .
2. Predpokladáme, že mimo pásu  $[a, b]$  je funkcia  $f$  nulová, teda uvažujeme hraničné podmienky  $f_1^n = 0, f_{K+1}^n = 0$ , pre  $n = 0, 1, \dots, \infty$ .

#### 4.3.2 Explicitná FTCS metóda

FTCS (z angl.: forward time centered space) využíva doprednú časovú a centrálnu priestorovú diferenciu pre aproximáciu príslušných derivácií, pričom dostávame diferenčnú

### 4.3 Parabolické parciálne diferenciálne rovnice

---

rovniciu

$$\frac{f_i^{n+1} - f_i^n}{\Delta t} + O(\Delta t) = \kappa \frac{f_{i+1}^n - 2f_i^n + f_{i-1}^n}{(\Delta x)^2} + O((\Delta x)^2). \quad (4.21)$$

Po úprave máme

$$f_i^{n+1} = \alpha f_{i-1}^n + (1 - 2\alpha) f_i^n + \alpha f_{i+1}^n, \quad (4.22)$$

kde  $\alpha = \kappa \Delta t / (\Delta x)^2$  je difúzne číslo.

#### 4.3.3 Konzistencia a presnosť

Ľahko sa presvedčíme, že rozvojom členov rovnice (4.22) do Taylorovho radu v  $(x_i, t^n)$  dostaneme modifikovanú diferenciálnu rovnicu (MDR)

$$f_t + \frac{1}{2} f_{tt} \Delta t + O((\Delta t)^2) = \kappa f_{xx} + \frac{1}{12} \kappa f_{xxxx} (\Delta x)^2 + O((\Delta x)^4), \quad (4.23)$$

kde dolné indexy označujú parciálne derivácie podľa  $t$  alebo  $x$ . Pre  $t \rightarrow 0, \Delta x \rightarrow 0$  a tak dostávame pôvodnú PDR, teda rovnica je *konzistentná*. Rovnica (4.22) má vo všeobecnosti presnosť 1. rádu v čase a 2. rádu v priestore, avšak dá sa ukázať, že pre  $\alpha = 1/6$  je aj v čase presnosť 2. rádu. Keďže rovnica (4.22) dáva hodnotu  $f$  v časovej úrovni  $n + 1$  pomocou troch priestorových bodov na predchádzajúcej časovej úrovni  $n$ , potom takúto rovnicu môžeme klasifikovať ako *explicitnú dvojúrovňovú*.

#### 4.3.4 Postupné mapovanie a numerická stabilita

Rovnicu (4.22) môžeme zapísať aj maticovo a zohľadniac okrajové podmienky  $f_1^n = f_{K+1}^n = 0$  dostávame

$$\mathbf{f}^{n+1} = \mathbf{B} \mathbf{f}^n, \quad (4.24)$$

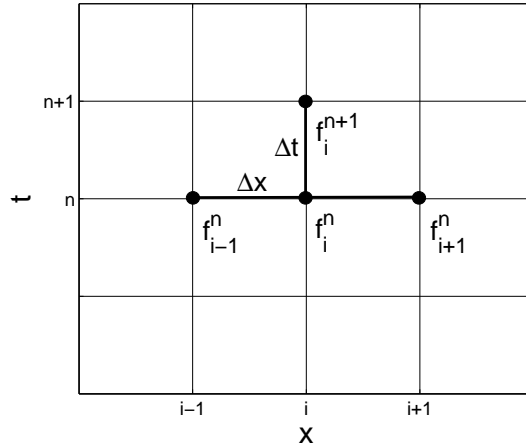
kde matica  $\mathbf{B}$  je tridiagonálna matica rádu  $(K - 1) \times (K - 1)$ :

$$\mathbf{B} = \begin{bmatrix} 1 - 2\alpha & \alpha & 0 & \dots & 0 \\ \alpha & 1 - 2\alpha & \alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \alpha & 1 - 2\alpha & \alpha \\ 0 & \dots & 0 & \alpha & 1 - 2\alpha \end{bmatrix}.$$

Matica  $\mathbf{B}$  je takzvaná *projekčná matica* a udáva vzťah medzi časovaním a postupným mapovaním. Správanie sa projekčnej matice závisí na jej spektrálnom polomere  $\rho(\mathbf{B})$ ,

### 4.3 Parabolické parciálne diferenciálne rovnice

---



**Obr. 4.3:** Diskretizačná schéma pre FTCS metódu.

definovanom ako maximálna hodnota normy vlastných hodnôt matice  $\mathbf{B}$ . Metóda je numericky stabilná ak  $\rho(\mathbf{B}) < 1$ . Využijúc faktu, že matica  $\mathbf{B}$  je tridiagonálna a Toeplitzova (konštantné prvky na diagonálach), jej vlastné hodnoty môžeme vyjadriť v tvare

$$\lambda_m = 1 - 4\alpha \sin^2\left(\frac{m\pi}{2K}\right), \quad (4.25)$$

pre  $m = 1, \dots, K - 1$ . Spektrálny polomer je menší než jedna pre akúkoľvek hodnotu  $K$  len ak  $\alpha < 1/2$ . To znamená že metóda FTCS je len *podmienečne stabilná*.

#### 4.3.5 Efekt hraničných podmienok

Zmeňme hraničné podmienky z Dirichletových na von Neumannove. Nech  $\partial f / \partial x = g$  v  $x = a$  a ponecháme homogénne Dirichletové podmienky v  $x = b$ . Aby sme aplikovali von Neumannovu podmienku s presnosťou 2. rádu rozšírme doménu riešenia za hranicu  $x = a$ , do bodu  $x_0 = x_1 - \Delta x$  a pri použití centrálnej diferencie dostávame  $f_2^n - f_0^n = 2\alpha g \Delta x$ . Potom môžeme aplikovať diferenčnú rovnicu (4.22) aj v prvom uzle  $x_1$  a dostávame

$$\mathbf{f}^{n+1} = \mathbf{B}\mathbf{f}^n + \mathbf{b}, \quad (4.26)$$



### 4.3 Parabolické parciálne diferenciálne rovnice

---

kde projekčná matica  $B$  je opäť tridiagonálna matica ale už radu  $K \times K$ :

$$\mathbf{B} = \begin{bmatrix} 1-2\alpha & 2\alpha & 0 & \dots & 0 \\ \alpha & 1-2\alpha & \alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \alpha & 1-2\alpha & \alpha \\ 0 & \dots & 0 & \alpha & 1-2\alpha \end{bmatrix}$$

a

$$\mathbf{b} = \begin{bmatrix} -2\alpha g \Delta x \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

Prítomnosť vektora  $\mathbf{b}$  neovplyvní dôležitosť matice  $\mathbf{B}$  pre stabilitu riešenia, avšak  $\mathbf{B}$  už nie je Toeplitzova a jej vlastné hodnoty sa už nedajú vyjadriť analyticky.

#### 4.3.6 Von Neumannova stabilita

Táto metóda je jednoduchšia než výpočet cez projekčnú maticu, avšak je aplikovateľná len za určitých podmienok (lineárne PDR s konštantnými koeficientami, periodickými hraničnými podmienkami, najviac dvoma nezávislými premennými a nie viac ako dvoma časovými úrovňami). Využijúc linearity, uvažujme riešenie v tvare súčiny časovej a priestorovej zložky:

$$f(x, t) = A(t) \exp(ikx), \quad (4.27)$$

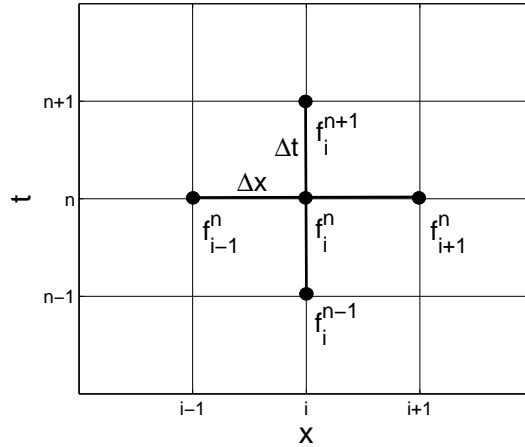
kde  $k = 2\pi/L$  je vlnové číslo a  $L$  je vlnová dĺžka. Pre diskretný tvar dostávame

$$\begin{aligned} f_l^n &= A^n \exp(ikx_l) = A^n \exp\left(\frac{i2\pi l \Delta x}{L}\right), \\ f_l^{n+1} &= A^{n+1} \exp(ikx_l) = A^{n+1} \exp\left(\frac{i2\pi l \Delta x}{L}\right), \end{aligned} \quad (4.28)$$

kde  $x_l = l\Delta x$ . Koeficient  $\xi = \frac{A^{n+1}}{A^n}$  sa nazýva faktor zosilnenia alebo rastu. Platí, že ak  $|\xi| < 1$  potom je riešenie stabilné. Ak si označíme  $\Theta = 2\pi\Delta x/L$  a vzťahy (4.28) dosadíme do rovnice (4.22), ľahko odvodíme vzťah

$$\xi = \frac{A^{n+1}}{A^n} = 1 + 2\alpha(\cos \Theta - 1) = 1 - 4\alpha \sin^2 \frac{\Theta}{2}. \quad (4.29)$$

### 4.3 Parabolické parciálne diferenciálne rovnice



**Obr. 4.4:** Diskretizačná schéma pre CTCS metódu.

Teda aby platilo  $|\xi| < 1$  musí byť splnená podmienka  $\alpha < 1/2$ , t. j. rovnaká podmienka akú sme dostali pomocou metódy projekčnej matice. Keďže  $\alpha = \frac{\kappa \Delta t}{(\Delta x)^2}$ , potom musí platiť  $\Delta t \leq \frac{(\Delta x)^2}{2\kappa}$ . To znamená, že podmienka stability vyžaduje dostatočne malý časový krok  $\Delta t$ , čo môže byť výpočtovo veľmi zaťažujúce.

#### 4.3.7 Explicitná CTCS metóda

V CTCS (z angl.: centered time centered space) je časová derivácia nahradená centrálnou diferenciou, čím sa dosiahne presnosť 2. rádu aj v čase:

$$\frac{f_i^{n+1} - f_i^{n-1}}{2\Delta t} + O((\Delta t)^2) = \kappa \frac{f_{i+1}^n - 2f_i^n + f_{i-1}^n}{(\Delta x)^2} + O((\Delta x)^2). \quad (4.30)$$

Dá sa však ukázať, že metóda je bezpodmienečne nestabilná a teda bez praktického úžitku. Je to explicitná trojúrovňová (v čase) metóda.

#### 4.3.8 Explicitná Du Fortova a Frankelova metóda

Du Fortova a Frankelova metóda je CTCS metóda modifikovaná za účelom zlepšenia stability a udržania presnosti 2. rádu v čase aj priestore. Modifikácia spočíva v nahradení funkčnej hodnoty v  $i$ -tom uzle na časovej úrovni  $n$  aritmetickým priemerom

### 4.3 Parabolické parciálne diferenciálne rovnice

---

funkčných hodnôt na časových úrovniach  $n - 1$  a  $n + 1$ :

$$\frac{f_i^{n+1} - f_i^{n-1}}{2\Delta t} + O((\Delta t)^2) = \kappa \frac{f_{i+1}^n - 2[\frac{1}{2}(f_i^{n+1} + f_i^{n-1})] + f_{i-1}^n}{(\Delta x)^2} + O((\Delta x)^2). \quad (4.31)$$

Pre  $\alpha$  z predchádzajúceho, po úprave máme

$$f_i^{n+1} = \frac{1 - 2\alpha}{1 + 2\alpha} f_i^{n-1} + \frac{2\alpha}{1 + 2\alpha} (f_{i-1}^n + \alpha f_{i+1}^n). \quad (4.32)$$

**Von Neumannova stabilita:** Dá sa ukázať, že faktor zosilnenia  $\xi$  spĺňa rovnicu

$$(1 + 2\alpha)\xi^2 + 4\xi\alpha \cos \Theta - 1 + 2\alpha = 0, \quad (4.33)$$

korene ktorej spĺňajú nerovnosť  $|\xi| < 1$  pre ľubovoľnú hodnotu  $\alpha$  a teda metóda je *bezpodmienečne stabilná*.

Vzhľadom na to, že sa urobila ad-hoc modifikácia v CTCS diferenčnej rovnici, je ešte potrebné overiť konzistentnosť metódy. Podobným postupom ako pri metóde FTCS dospejeme k modifikovanej diferenciálnej rovnici

$$f_t = \kappa f_{xx} - \kappa \left( \frac{\Delta t}{\Delta x} \right)^2 f_{tt}. \quad (4.34)$$

Je zrejmé, že ani pre ľubovoľne malé hodnoty  $\Delta t, \Delta x$  posledný člen na pravej strane rovnice nevymizne a teda metóda nie je konzistentná. Je to explicitná trojúrovňová metóda.

#### 4.3.9 Implicitná BTCS alebo Laasonenova metóda

BTCS (z angl.: backward time centered space) je implicitná metóda a bude vyžadovať riešenie sústavy rovníc. Využíva spätnú časovú a centrálnu priestorovú diferenciáciu pre aproximáciu príslušných derivácií, pričom dostávame diferenčnú rovnicu

$$\frac{f_i^{n+1} - f_i^n}{\Delta t} + O(\Delta t) = \kappa \frac{f_{i+1}^{n+1} - 2f_i^{n+1} + f_{i-1}^{n+1}}{(\Delta x)^2} + O((\Delta x)^2). \quad (4.35)$$

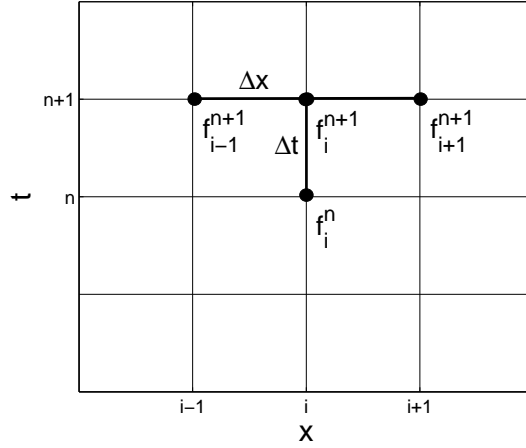
Po úprave máme

$$-\alpha f_{i-1}^{n+1} + (1 + 2\alpha) f_i^{n+1} - \alpha f_{i+1}^{n+1} = f_i^n, \quad (4.36)$$

a maticovo pre homogénne Dirichletove okrajové podmienky

$$\mathbf{A} \mathbf{f}^{n+1} = \mathbf{f}^n, \quad (4.37)$$

### 4.3 Parabolické parciálne diferenciálne rovnice



**Obr. 4.5:** Diskretizačná schéma pre BTCS metódu.

kde matica  $\mathbf{A}$  je tridiagonálna matica

$$\mathbf{A} = \begin{bmatrix} 1 + 2\alpha & -\alpha & 0 & \dots & 0 \\ -\alpha & 1 + 2\alpha & -\alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & -\alpha & 1 + 2\alpha & -\alpha \\ 0 & \dots & 0 & -\alpha & 1 + 2\alpha \end{bmatrix}.$$

Riešenie teda môžeme formálne vyjadriť v tvare

$$\mathbf{f}^{n+1} = \mathbf{A}^{-1}\mathbf{f}^n. \quad (4.38)$$

Vďaka tomu že vlastné hodnoty matice  $\mathbf{A}^{-1}$  sú prevrátenými hodnotami vlastných hodnôt matice  $\mathbf{A}$ , dostávame pre ich hodnoty vzťah

$$\lambda_m = \left[ 1 + 4\alpha \sin^2 \left( \frac{m\pi}{2K} \right) \right]^{-1}, \quad (4.39)$$

pre  $m = 1, \dots, K - 1$ . Spektrálny polomer je vždy menší než jedna pre akúkoľvek hodnotu  $K$  a teda metóda BTCS je *bezpodmienečne stabilná*. Nezávislé overenie bezpodmienečnej stability môžeme získať pomocou von Neumannovej metódy. Podobným postupom aký sme si popísali vyššie dostávame pre faktor zosilnenia

$$|\xi| = \left| \frac{1}{1 + 2\alpha(1 - \cos \Theta)} \right| = \left| \frac{1}{1 + 4\alpha \sin^2 \frac{\Theta}{2}} \right| < 1, \quad (4.40)$$

pre ľubovoľné  $\Theta$  a  $\alpha$ . Presnosť je však len 1. rádu v čase.

### 4.3 Parabolické parciálne diferenciálne rovnice

#### 4.3.10 Crankova-Nicolsonova metóda

Ponúka vylepšenie presnosti do 2. rádu v čase aj priestore pri udržaní bezpodmienečnej stability. V FTCS sme aplikovali diferenčnú metódu v bode  $x_i$  a čase  $t^n$ , v BTCS v bode  $x_i$  a čase  $t^{n+1}$  a teraz ju aplikujeme v bode  $x_i$  a čase  $t^{n+1/2}$ . Je to to isté ako keď vezmeme aritmetický priemer rovníc pre FTCS a BTCS, pričom lineárne členy  $O(\Delta t)$  na ľavej strane rovnice sa vyrušia a týmto dosiahneme presnosť  $O((\Delta t)^2)$  v čase a  $O((\Delta x)^2)$  v priestore. Výsledná diferenčná rovnica má tvar

$$\frac{f_i^{n+1} - f_i^n}{\Delta t} = \kappa \frac{1}{2} \left( \frac{f_{i+1}^{n+1} - 2f_i^{n+1} + f_{i-1}^{n+1}}{(\Delta x)^2} + \frac{f_{i+1}^n - 2f_i^n + f_{i-1}^n}{(\Delta x)^2} \right). \quad (4.41)$$

Po úprave máme

$$-\alpha f_{i-1}^{n+1} + 2(1 + \alpha)f_i^{n+1} - \alpha f_{i+1}^{n+1} = \alpha f_{i-1}^n + 2(1 - \alpha)f_i^n + \alpha f_{i+1}^n, \quad (4.42)$$

alebo v maticovom tvare pre homogénne Dirichletove okrajové podmienky

$$\mathbf{A}\mathbf{f}^{n+1} = \mathbf{B}\mathbf{f}^n, \quad (4.43)$$

kde matice  $\mathbf{A}$  a  $\mathbf{B}$  sú tridiagonálne matice v tvare

$$\mathbf{A} = \begin{bmatrix} 2(1 + \alpha) & -\alpha & 0 & \dots & 0 \\ -\alpha & 2(1 + \alpha) & -\alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & -\alpha & 2(1 + \alpha) & -\alpha \\ 0 & \dots & 0 & -\alpha & 2(1 + \alpha) \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 2(1 - \alpha) & \alpha & 0 & \dots & 0 \\ \alpha & 2(1 - \alpha) & \alpha & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \alpha & 2(1 - \alpha) & \alpha \\ 0 & \dots & 0 & \alpha & 2(1 - \alpha) \end{bmatrix}.$$

Riešenie môžeme formálne vyjadriť v tvare

$$\mathbf{f}^{n+1} = \mathbf{A}^{-1}\mathbf{B}\mathbf{f}^n, \quad (4.44)$$

kde  $\mathbf{A}^{-1}\mathbf{B}$  je projekčná matica. Dá sa ukázať, že jej spektrálny polomer je menší ako 1 a teda, že metóda je *bezpodmienečne stabilná*. Overením pomocou von Neumannovej metódy dostávame pre faktor zosilnenia

$$|\xi| = \left| \frac{1 - \alpha(1 - \cos \Theta)}{1 + \alpha(1 - \cos \Theta)} \right| = \left| \frac{1 - 2\alpha \sin^2 \frac{\Theta}{2}}{1 + 2\alpha \sin^2 \frac{\Theta}{2}} \right| < 1, \quad (4.45)$$

### 4.3 Parabolické parciálne diferenciálne rovnice

---

pre ľubovoľné  $\Theta$  a  $\alpha$ .

**Poznámka:** Metóda môže byť ponímaná ako výsledok dvoch krokov s  $\Delta t/2$ . Najprv aplikujeme explicitnú FTCS metódu pomocou diferenčnej rovnice

$$\frac{f_i^{n+1/2} - f_i^n}{\Delta t/2} = \kappa \frac{f_{i+1}^n - 2f_i^n + f_{i-1}^n}{(\Delta x)^2} \quad (4.46)$$

a potom implicitnú BTCS metódu pomocou diferenčnej rovnice

$$\frac{f_i^{n+1} - f_i^{n+1/2}}{\Delta t/2} = \kappa \frac{f_{i+1}^{n+1} - 2f_i^{n+1} + f_{i-1}^{n+1}}{(\Delta x)^2}. \quad (4.47)$$

Ich sčítaním dostávame rovnicu (4.41).

#### 4.3.11 Rovnica difúzie v 2D

Riešime parabolickú parciálnu diferenciálnu rovnicu

$$\frac{\partial f}{\partial t} = \kappa \left( \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \right), \quad (4.48)$$

s počiatočnou podmienkou  $f(\mathbf{x}, t = 0) = \mathbf{F}(\mathbf{x})$ , kde  $\mathbf{x} = (x, y)$ ,  $\mathbf{F}(\mathbf{x})$  je známa funkcia a  $\kappa$  je difúzna konštanta. Postup, ktorý sme aplikovali v 1D je možné priamo rozšíriť do 2D.

#### 4.3.12 Explicitná FTCS metóda

Nadradením parciálnych derivácií príslušnými diferenciami (viď diskretizačnú schému v 2D na obr. 4.1) dostávame diferenčnú rovnicu

$$\frac{f_{i,j}^{n+1} - f_{i,j}^n}{\Delta t} = \kappa \left( \frac{f_{i+1,j}^n - 2f_{i,j}^n + f_{i-1,j}^n}{(\Delta x)^2} + \frac{f_{i,j+1}^n - 2f_{i,j}^n + f_{i,j-1}^n}{(\Delta y)^2} \right), \quad (4.49)$$

kde pravá strana rovnice predstavuje diskretnú aproximáciu Laplaciánu v 2D. Po úprave máme

$$f_{i,j}^{n+1} = f_{i,j}^n + \alpha_x (f_{i+1,j}^n + f_{i-1,j}^n) + [1 - 2(\alpha_x + \alpha_y)] f_{i,j}^n + \alpha_y (f_{i,j+1}^n + f_{i,j-1}^n), \quad (4.50)$$

kde  $\alpha_x = \kappa \Delta t / (\Delta x)^2$  a  $\alpha_y = \kappa \Delta t / (\Delta y)^2$  sú difúzne čísla v smere osi  $x$  a  $y$ .

Skúmaním von Neumannovej stability zistíme, že metóda je stabilná pre  $\alpha_x + \alpha_y < 1/2$ , čo je dosť silné obmedzenie pre  $\Delta t$  čo sa týka efektivity výpočtu.

## 4.3 Parabolické parciálne diferenciálne rovnice

---

### 4.3.13 Implicitná BTCS metóda

Nahradením časovej parciálnej derivácie spätnou diferenciou po úprave dostávame

$$-\alpha_x(f_{i+1,j}^{n+1} + f_{i-1,j}^{n+1}) + [1 + 2(\alpha_x + \alpha_y)]f_{i,j}^{n+1} - \alpha_y(f_{i,j+1}^{n+1} + f_{i,j-1}^{n+1}) = f_{i,j}^n. \quad (4.51)$$

Metóda je bezpodmienečne stabilná s presnosťou 1. rádu v  $t$  a 2. rádu v  $x$  a  $y$ . Potrebujeme však riešiť sústavu rovníc s päťdiagonálnou maticou, čo je výpočtovo oveľa náročnejšie než to bolo v prípade tridiagonálnej matice.

### 4.3.14 Metóda striedavých smerov (ADI)

Redukuje výpočtové nároky predchádzajúcich metód. Je založená na rozdelení časového vývoja do dvoch krokov a kombinácie implicitnej a explicitnej metódy pre priestorové derivácie, aplikovanú striedavo v smere osi  $x$  a  $y$ :

1. **krok**: implicitný v smere osi  $x$  a explicitný v smere osi  $y$

$$\frac{f_{i,j}^{n+1/2} - f_{i,j}^n}{\Delta t/2} = \kappa \left( \frac{f_{i+1,j}^{n+1/2} - 2f_{i,j}^{n+1/2} + f_{i-1,j}^{n+1/2}}{(\Delta x)^2} + \frac{f_{i,j+1}^n - 2f_{i,j}^n + f_{i,j-1}^n}{(\Delta y)^2} \right), \quad (4.52)$$

2. **krok**: implicitný v smere osi  $y$  a explicitný v smere osi  $x$

$$\frac{f_{i,j}^{n+1} - f_{i,j}^{n+1/2}}{\Delta t/2} = \kappa \left( \frac{f_{i+1,j}^{n+1/2} - 2f_{i,j}^{n+1/2} + f_{i-1,j}^{n+1/2}}{(\Delta x)^2} + \frac{f_{i,j+1}^{n+1} - 2f_{i,j}^{n+1} + f_{i,j-1}^{n+1}}{(\Delta y)^2} \right), \quad (4.53)$$

kde  $n + 1/2$  je časový medziúsek. Smery osí sú alternatívne striedané po každom kroku. Celková presnosť je 2. rádu v čase aj priestore. Po úprave dostávame dvoj krokový implicitný algoritmus

$$\alpha_x f_{i-1,j}^{n+1/2} - 2(1 + \alpha_x) f_{i,j}^{n+1/2} + \alpha_x f_{i+1,j}^{n+1/2} = -\alpha_y f_{i,j-1}^n - 2(1 - \alpha_y) f_{i,j}^n - \alpha_y f_{i,j+1}^n \quad (4.54)$$

$$\alpha_y f_{i,j-1}^{n+1} - 2(1 + \alpha_y) f_{i,j}^{n+1} + \alpha_y f_{i,j+1}^{n+1} = -\alpha_x f_{i-1,j}^{n+1/2} - 2(1 - \alpha_x) f_{i,j}^{n+1/2} - \alpha_x f_{i+1,j}^{n+1/2} \quad (4.55)$$

To znamená, že v každom časovom kroku je potrebné riešiť dve sústavy tridiagonálnych rovníc, čo je možné efektívne urobiť pomocou Thomasovho algoritmu.

### 4.3 Parabolické parciálne diferenciálne rovnice

---

Vyšetovaním von Neumannovej stability dostávame

$$\begin{aligned}\xi &= \frac{[1 + \alpha_x(\cos \Theta_x - 1)][1 + \alpha_y(\cos \Theta_y - 1)]}{[1 - \alpha_x(\cos \Theta_x - 1)][1 - \alpha_y(\cos \Theta_y - 1)]} \\ &= \frac{(1 - 2\alpha_x \sin^2 \frac{\Theta_x}{2})(1 - 2\alpha_y \sin^2 \frac{\Theta_y}{2})}{(1 + 2\alpha_x \sin^2 \frac{\Theta_x}{2})(1 + 2\alpha_y \sin^2 \frac{\Theta_y}{2})}.\end{aligned}\tag{4.56}$$

Pre ľubovoľné  $\alpha_x, \alpha_y$  a  $\Theta_x, \Theta_y$  platí  $|\xi| < 1$  a teda metóda je *bezpodmienečne stabilná*.



# Kapitola 5

## Monte Carlo simulácie

### 5.1 Úvod

Vo všeobecnosti môžeme charakterizovať metódu Monte Carlo (MC) ako počítačovú metódu využívajúcu náhodne čísla. MC metóda sa snaží sledovať časový vývoj (závislosť) modelu, pre ktorý rast alebo zmena nepostupuje rigorózne definovaným spôsobom (ako napr. u Newtonových rovníc) ale náhodným (stochastickým) spôsobom, závisiacim na nejakej postupnosti náhodných čísiel. To znamená, že rôzne postupnosti náhodných čísiel dávajú rôzne výsledky, avšak vrámci nejakej štatistickej chyby.

#### Príklady využitia vo fyzike:

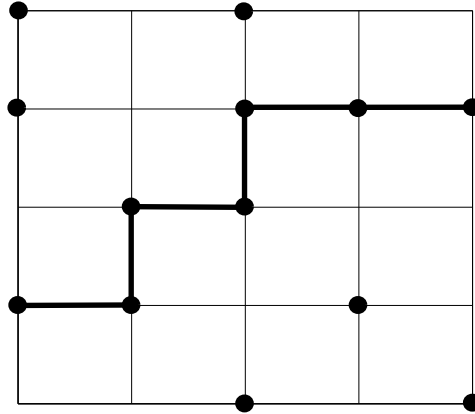
- **Perkolácia na postupne sa zaplňajúcej mriežke** - ak na počiatočne prázdnu mriežku začneme náhodným spôsobom postupne umiestňovať častice, po istom počte umiestnených častíc dospeje systém do stavu keď sa vytvorí cesta spájajúca najbližších susedov na mriežke preklenujúca celú mriežku, t. j. dochádza k perkolácii. Počet častíc potrebných na vytvorenie takejto cesty je náhodné číslo (pri každom pokuse bude vo všeobecnosti iný) a jej stredná hodnota je tzv. perkolačný prah.

#### Príklady fyzikálnych aplikácií problému perkolácie:

- ◇ vodivosť náhodných zmesí
  - ◇ tok cez pôrovité materiály
  - ◇ magnety zriadené nemagnetickými prímiesami
- **Riešenie problémov v štatistickej mechanike** - odhad istých vlastností modelu vzorkovaním oblasti fázového priestoru. Na rozdiel od exaktných riešení sa nebudeme pohybovať pozdĺž rovnakej trajektórie ale mnohých náhodných,

## 5.1 Úvod

---



**Obr. 5.1:** Schématicky zobrazený stav perkolácie na postupne sa zaplňajúcej štvorcovej mriežke keď je vytvorená cesta (hrubá lomená čiara) spájajúca najbližších susedov na mriežke preklenujúca celú mriežku.

ktoré povedú k rovnovážnemu stavu systému, a následne vypočítame priemer teplotným stredovaním. MC metóda berie správne do úvahy štatistické fluktuácie a ich vplyv na mnohočasticový interagujúci systém. Čím lepšie vzorkovanie fázového priestoru (t. j. dlhšie simulácie, väčšie súbory vzoriek), tým presnejšie výsledky môžeme dosiahnuť.

### 5.1.1 Aké problémy môžeme riešiť pomocou MC simulácií?

- prirodzene diskkrétne alebo aproximatívne diskretizované systémy
- pohyb jednotlivých atómov
- rastové javy včetně makroskopických objektov, rast koloidných častíc
- pohyb kvapalín - aproximácia blokmi jednotlivých častíc
- množstvo klasických a kvantových interagujúcich systémov
- skúmanie rovnovážnych vlastností rôznych magnetických materiálov, kovových zliatín, jednoduchých aj komplexných kvapalín, polymérov a pod.
- ceny akcií na burze
- výpočet viacrozmerných integrálov, hlavne v štatistickej fyzike

## 5.2 Základné pojmy teórie pravdepodobnosti

---

### Problémy a limitácie použitia MC simulácií

- obmedzený CPU čas a pamäť
- niektoré problémy nie je možné alebo praktické riešiť pomocou MC simulácií kvôli príliš veľkým nárokom na výpočtový CPU čas
- vďaka rýchlemu rastu výpočtových kapacít vznikajú nové perspektívy pre použitie MC techník

### Chyby MC metódy

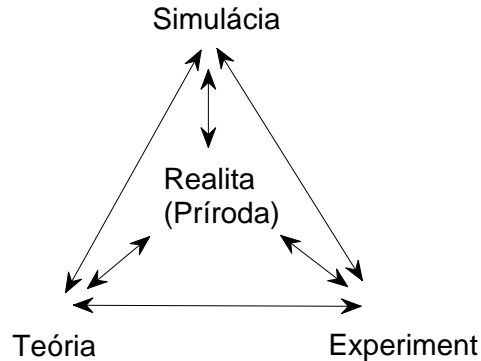
- *zaokružľovacie chyby* (viď sekciu 1.3)
- *štatistické chyby* - vznikajú nevyhnutne kvôli konečnému počtu vzoriek (volí sa kompromis medzi parametrami simulácie a rozumným CPU časom)
- *systematické chyby* - nie je možné správne riešiť isté problémy kvôli konečnému počtu vzoriek
- je dôležité *otestovať simulácie* pre známe hodnoty parametrov so známymi (už riešenými) výsledkami, ak existujú, a to najprv s malým počtom vzoriek a taktiež *optimalizovať algoritmus* pre čo najlepšiu efektívnosť výpočtu

### 5.1.2 Vzťah teória-experiment-simulácia-realita

- *Simulácia*  $\leftrightarrow$  *realita* - overovanie teórie, pre ktorú (ešte) neexistujú experimentálne realizácie alebo ich nechceme robiť lebo sú príliš drahé alebo nebezpečné (napr. výbuch reaktora).
- *Simulácia*  $\leftrightarrow$  *experiment* - reálne fyzikálne systémy sú príliš zložité na to aby sme skúmali zvlášť jednotlivé interakcie. Simuláciou môžeme separovať jednotlivé interakcie a lepšie pochopiť ich účinky. Pomocou simulácie tiež môžeme overiť správnosť hamiltoniánu porovnaním výsledkov simulácie s experimentálnymi výsledkami. Avšak, cieľom simulácie nie je fitovanie experimentálnych dát ale pochopenie fyzikálnych vlastností.

## 5.2 Základné pojmy teórie pravdepodobnosti

---



Obr. 5.2: Schématicky znázornený vzťah medzi teóriou, experimentom, simuláciou a realitou.

## 5.2 Základné pojmy teórie pravdepodobnosti

- **Náhodná premenná** - ak pri opakovaných pokusoch pri identických podmienkach merania dostávame rôzne výsledky. Napr. pri hode kockou je náhodná premenná  $x \in \{1, 2, \dots, 6\}$ .
- **Funkcia hustoty pravdepodobnosti** - definuje pravdepodobnosť  $p(x)$  že pri vzorkovaní (pokuse) dostaneme výsledok  $x$ . Pre veľký počet pokusov platí  $p(x) = (\text{počet vzoriek s výsledkom } x) / (\text{celkový počet vzoriek})$ . Napr. pri hode kockou je  $p(1) = p(2) = \dots = p(6) = 1/6$ .
- **Normalizačná podmienka**

$$\sum_k p(x_k) = 1. \quad (5.1)$$

- **Stredná (očakávaná) hodnota náhodnej premennej**

$$E(x) = \langle x \rangle = \sum_k x_k p(x_k). \quad (5.2)$$

Napr. pri hode kockou je  $\langle x \rangle = 1 \times \frac{1}{6} + \dots + 6 \times \frac{1}{6} = 3,5$ .

- **Spojité náhodné premenné** -  $x$  môže nadobúdať spojité hodnoty (napr. pozícia atómu pozdĺž osi  $x$ ). Pravdepodobnosť, že hodnota bude v rozmedzí  $x$  a  $x + dx$  je  $p(x)dx$ .

## 5.2 Základné pojmy teórie pravdepodobnosti

---

- *Normalizačná podmienka:*

$$\int_{-\infty}^{\infty} p(x)dx = 1. \quad (5.3)$$

- *Stredná hodnota*

$$E(x) = \langle x \rangle = \int_{-\infty}^{\infty} xp(x)dx. \quad (5.4)$$

- *Stredná hodnota funkcie náhodnej premennej*

$$E(f(x)) = \langle f(x) \rangle = \int_{-\infty}^{\infty} f(x)p(x)dx. \quad (5.5)$$

**Príklad:** Nech  $x$  má uniformné (rovnorné) rozdelenie na intervale  $[a, b]$ . Potom z normalizačnej podmienky ľahko overíme že  $p(x) = 1/(b - a)$ . Ďalej platí

$$\langle x \rangle = \int_a^b x \frac{1}{b-a} dx = \frac{1}{b-a} \left[ \frac{x^2}{2} \right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}. \quad (5.6)$$

Nech  $f(x) = x^2$ , potom

$$\langle x^2 \rangle = \int_a^b x^2 \frac{1}{b-a} dx = \frac{1}{b-a} \left[ \frac{x^3}{3} \right]_a^b = \frac{b^3 - a^3}{2(b-a)} = \frac{a^2 + ab + b^2}{3}. \quad (5.7)$$

- *Rozptyl a smerodajná odchýlka náhodnej premennej*

Rozptyl je daný vzťahom

$$Var(x) = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 - 2x \langle x \rangle + \langle x \rangle^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2. \quad (5.8)$$

Napr. pre uniformné rozdelenie dostávame

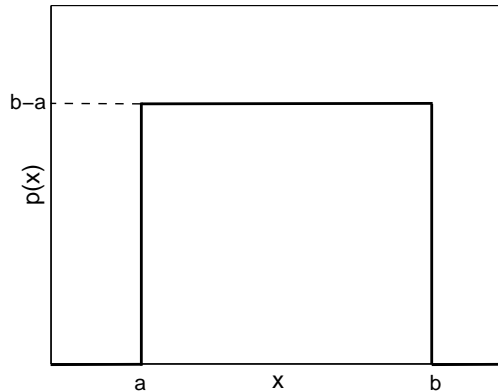
$$Var(x) = \frac{a^2 + ab + b^2}{3} - \left( \frac{a+b}{2} \right)^2 = \frac{(a-b)^2}{12}. \quad (5.9)$$

Smerodajná (alebo štandardná) odchýlka náhodnej premennej je potom

$$Std(x) = \sqrt{Var(x)} = \frac{|a-b|}{2\sqrt{3}}. \quad (5.10)$$

## 5.3 Monte Carlo odhad

---



Obr. 5.3: Uniformné rozdelenie pravdepodobnosti na intervale  $[a, b]$ .

- **Niektoré vlastnosti**

$$E[ax \pm by] = aE(x) \pm bE(y), \quad (5.11)$$

$$\text{Var}[ax \pm by] = a^2\text{Var}(x) + b^2\text{Var}(y), \quad (5.12)$$

kde  $x, y$  sú náhodné premenné a  $a, b$  sú konštanty.

Podrobnejší výklad k problematike náhodnej premennej a pravdepodobnosti, aj s niektorými aplikáciami v prostredí Octave/Matlab, je možné nájsť napr. v skriptách od autorov Buša a kol.

## 5.3 Monte Carlo odhad

### 5.3.1 Stredná hodnota súboru

Výpočet (jednorozmerného) integrálu je možné urobiť pomocou MC odhadu strednej hodnoty funkcie. Nech  $p(x)$  je nejaká funkcia hustoty pravdepodobnosti a je nenulová na nejakom intervale  $[a, b]$ .

$$F(x) = \int_a^b f(x)dx = \int_a^b \frac{f(x)}{p(x)}p(x)dx = \left\langle \frac{f(x)}{p(x)} \right\rangle. \quad (5.13)$$

To znamená, že deterministické integrovanie sme nahradili výpočtom strednej hodnoty cez náhodné vzorky. Ak  $p(x) = 1/(b - a)$ , potom

$$\int_a^b f(x)dx = (b - a) \langle f(x) \rangle. \quad (5.14)$$

### 5.3 Monte Carlo odhad

---

Pre odhad strednej hodnoty funkcie pomocou MC simulácii generujeme postupnosť uniformných náhodných čísiel  $x = r_n, n = 1, \dots, M$  na intervale  $[a, b]$ , pričom platí

$$\int_a^b f(x)dx = (b - a) \langle f(x) \rangle \approx (b - a) \frac{1}{M} \sum_{n=1}^M f(r_n). \quad (5.15)$$

#### Efekt konečného počtu vzoriek

V limitnom prípade nekonečného počtu vzoriek platí  $\langle f(x) \rangle = \int_{-\infty}^{\infty} f(x)p(x)dx$ . Avšak pre konečný počet vzoriek dostávame odhad strednej hodnoty v tvare aritmetického priemeru

$$\bar{f}(x) = \frac{1}{M} \sum_{n=1}^M f(r_n). \quad (5.16)$$

Je zrejmé že pre výberové súbory generované pomocou rôznych  $M$ -tíc náhodných čísiel  $r_n, n = 1, \dots, M$  z rozdelenia  $p(x)$  je aj odhad  $\bar{f}(x)$  náhodná premenná, ktorá má tiež strednú hodnotu a rozptyl.

**Neskreslený odhad** - ak veličina odhadnutá z  $M$  vzoriek konverguje ku skutočnej hodnote populácie (základného súboru).

**Teoréma:** Stredná hodnota  $M$  vzoriek náhodnej premennej  $\bar{f}(x)$  konverguje k hodnote  $\langle f(x) \rangle$ .

$$E\{\bar{f}(x)\} = \frac{1}{M} \sum_{n=1}^M E\{f(r_n)\} = \frac{1}{M} \sum_{n=1}^M \langle f(x) \rangle = \frac{1}{M} M \langle f(x) \rangle = \langle f(x) \rangle. \quad (5.17)$$

Tým sme overili platnosť teorémy, pričom sme využili vlastnosť (5.11) a dôležitú rovnosť  $E\{f(r_n)\} = \langle f(x) \rangle$ , ktorá vyplýva z nezávislosti vzoriek.

#### 5.3.2 Rozptyl a smerodajná odchýlka odhadu strednej hodnoty súboru

Analogicky ako pre strednú hodnotu, pre rozptyl dostávame

$$\begin{aligned} \text{Var}\{\bar{f}(x)\} &= \frac{1}{M^2} \sum_{n=1}^M \text{Var}\{f(r_n)\} = \frac{1}{M^2} \sum_{n=1}^M \text{Var}\{f(x)\} \\ &= \frac{1}{M^2} M \text{Var}\{f(x)\} = \frac{\text{Var}\{f(x)\}}{M}. \end{aligned} \quad (5.18)$$

## 5.4 Generátory náhodných čísiel

Podobne ako vyššie, využili sme vlastnosť (5.12) a rovnosť  $\text{Var}\{f(r_n)\} = \text{Var}\{f(x)\}$ , ktorá vyplýva z nezávislosti vzoriek. Odmocnením dostávame

$$\text{Std}\{\bar{f}(x)\} = \frac{\text{Std}\{f(x)\}}{\sqrt{M}}, \quad (5.19)$$

čo znamená že chyba MC odhadu strednej hodnoty sa znižuje s počtom vzoriek ako  $1/\sqrt{M}$ .

### Neskreslený odhad rozptylu a štandardnej odchýlky základného súboru

Nech  $S_M$  je rozptyl odhadnutý z  $M$  vzoriek

$$S_M = \frac{1}{M} \sum_{n=1}^M f^2(r_n) - \left[ \frac{1}{M} \sum_{n=1}^M f(r_n) \right]^2. \quad (5.20)$$

$S_M$  je tiež náhodná premenná, ktorej stredná hodnota je

$$\begin{aligned} \langle S_M \rangle &= \frac{1}{M} \sum_{n=1}^M \langle f^2(x) \rangle - \frac{1}{M^2} \sum_{m=1}^M \sum_{n=1}^M \langle f(r_m) f(r_n) \rangle \\ &= \langle f^2(x) \rangle - \frac{1}{M^2} \sum_{m=1}^M \langle f^2(x) \rangle - \frac{1}{M^2} \sum_{m=1}^M \sum_{n=1, n \neq m}^M \langle f(r_m) f(r_n) \rangle \\ &= \left(1 - \frac{1}{M}\right) \langle f^2(x) \rangle - \left(1 - \frac{1}{M}\right) \langle f(x) \rangle^2 \\ &= \left(1 - \frac{1}{M}\right) \text{Var}\{f(x)\}. \end{aligned} \quad (5.21)$$

Vyššie sme využili, že pre nekorelované (a nezávislé) náhodné premenné platí  $\langle f(r_m) f(r_n) \rangle = \langle f(x) \rangle \langle f(x) \rangle$ . To znamená, že neskresleným odhadom rozptylu základného súboru  $\text{Var}\{f(x)\}$  je  $\frac{M}{M-1} S_M$ . Dosadením do rovnice (5.18) dostávame

$$\begin{aligned} \text{Var}\{\bar{f}(x)\} &= \frac{1}{M-1} S_M = \frac{1}{M-1} \left\{ \frac{1}{M} \sum_{n=1}^M f^2(r_n) - \left[ \frac{1}{M} \sum_{n=1}^M f(r_n) \right]^2 \right\} \\ &\rightarrow \frac{\text{Var}\{f(x)\}}{M}. \end{aligned} \quad (5.22)$$

**Centrálna teoréma MC vzorkovania** sumarizuje vyššie odvodené vzťahy

$$\int f(x) p(x) dx \approx \bar{f} \pm \sqrt{\frac{\bar{f}^2 - (\bar{f})^2}{M-1}}, \quad (5.23)$$

kde  $\bar{f} = \frac{1}{M} \sum_{n=1}^M f(r_n)$ , a  $r_n$  je náhodná premenná z  $p(x)$ .



### 5.4 Generátory náhodných čísiel

Pre MC výpočty je dôležité generovať rýchlo a efektívne postupnosti náhodných čísiel. Fyzikálne procesy (napr. biely šum v elektrických obvodoch) nie sú efektívne a preto sa obyčajne generujú *pseudonáhodné čísla* pomocou softvéru s využitím rôznych algoritmov. Je preto dôležité otestovať kvalitu jednotlivých generátorov, pretože použitie zlých generátorov sa môže prejaviť v systematických chybách v MC výpočtoch. Vo všeobecnosti potrebujeme postupnosti uniformných nekorelovaných náhodných čísiel s čo najdlhšou periódou opakovania. Obyčajne sa generujú celé čísla, ktoré sa potom prenормujú tak, aby boli z intervalu  $[0, 1]$ .

#### 5.4.1 Kongruenčná metóda

Je založená na jednoduchom generovaní postupnosti veľkých kladných čísiel  $X_n$  pomocou rekurentného vzťahu

$$X_n = (cX_{n-1} + a_0) \pmod{N_{max}}, \quad (5.24)$$

kde  $c$  je multiplikátor,  $a_0$  je nejaká konštanta (obyčajne nula),  $N_{max}$  je dostatočne veľké a  $X_0$  je počiatočná hodnota (seed) ktorú je najlepšie voliť ako nepárne číslo. Populárne hodnoty pre  $(N_{max}, c) = (2^{31} - 1, 7^5)$  pre 32-bitový procesor. Potom

$$r_n = X_n/N_{max} \in [0, 1]. \quad (5.25)$$

Zistilo sa však, že takýto jednoduchý generátor vykazuje korelácie medzi trojicami po sebe idúcich čísiel. Vylepšením môžu byť rôzne kombinácie viacerých kongruenčných generátorov s rôznymi voľbami hodnôt  $c$  a  $X_0$ . Napríklad, jeden generátor vytvára tabuľku náhodných čísiel a druhý z nej náhodne vyberá členy.

#### 5.4.2 Shift register (Tauswortheov) generátor

Najprv vygenerujeme tabuľku náhodných čísiel a potom nové náhodné čísla sú generované z čísiel z tabuľky pomocou operácie .XOR. (exkluzívny OR operátor):

$$X_n = X_{n-p} \cdot \text{XOR} \cdot X_{n-q}, \quad (5.26)$$

kde  $p, q$  sú vhodne zvolené, napr. tak, aby spĺňali  $X^p + X^q + 1 =$  primitívne číslo. Príklady voľby dvojíc  $(p, q) = (98, 27), (250, 103), \dots$ . Zovšeobecnením je *Fibonacciho generátor*, založený na vzťahu

$$X_n = X_{n-p} * X_{n-q}, \quad (5.27)$$

## 5.4 Generátory náhodných čísel

---

kde  $*$  je nejaká operácia.

Kvalitu generátora náhodných čísel je možné jednoducho overiť napr. pomocou testu uniformity, v ktorom si rozdelíme interval  $[0, 1]$  na niekoľko malých priechinkov a po vygenerovaní veľkého počtu náhodných čísel a ich zaradení do priechinkov (vytvorení histogramu) skontrolujeme uniformitu, alebo pri zobrazení v  $m$ -rozmernom priestore hľadáme vzory, štruktúry, atď., ktoré by sa pri kvalitných generátoroch nemali vyskytovať.

### 5.4.3 Neuniformné (nerovnomerné) generátory náhodných čísel

Úlohou bude generovať náhodné čísla z neuniformného rozdelenia  $p(x)$  na intervale  $[a, b]$ .

- **Von Neumannovo odmietanie - algoritmus:**

1. Generuj vektor  $(x_t, w_t) \in [a, b] \times [0, w_{sup}]$ ,  
 $x_t = a + (b - a)r_{1t}$ ,  $w_t = w_{sup}r_{2t}$ , kde  
 $r_{1t} \sim U(0, 1)$ ,  $r_{2t} \sim U(0, 1)$ ,  $w_{sup} = \sup_{x \in [a, b]} p(x)$  a  $x \sim U(0, 1)$  znamená že náhodná premenná  $x$  pochádza z uniformného rozdelenia na intervale  $[0, 1]$ .
2. Ak  $w_t > p(x_t)$  potom odmietame  $x_t$  a pokračujeme v kroku 1.; ináč  $x_t$  akceptujeme a jej hodnotu priradíme premennej  $\xi$ . Potom množina všetkých  $\xi$  bude mať na  $[a, b]$  rozdelenie  $p(x)$ .

- **Nerovnomernosť dosiahnutá transformáciou premennej:**

Aj je požadované neuniformné rozdelenie  $p(x)$  na intervale  $[a, b]$ , potom kumulatívne rozdelenie (pravdepodobnosť, že  $x \leq x'$ ) je dané

$$P(x') = \int_a^{x'} p(x) dx. \quad (5.28)$$

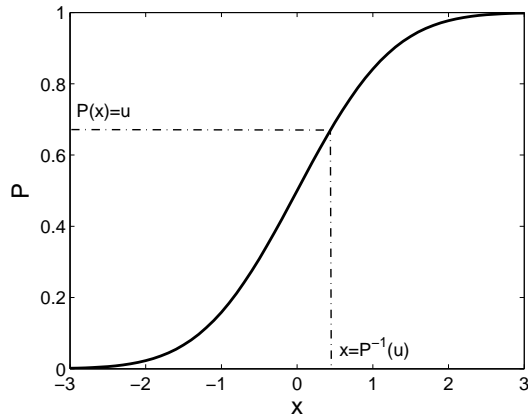
$P(x')$  je tiež náhodná premenná a to z intervalu  $[0, 1]$ :  $u = P(x)$  a potom  $x = P^{-1}(u)$ .

**Príklad:** Nech  $p(x) = \exp(-x)$  pre  $x \geq 0$ . Potom

$$P(x') = \int_0^{x'} \exp(-x) dx = [-\exp(-x)]_0^{x'} = 1 - \exp(-x'). \quad (5.29)$$

Teda máme  $u = 1 - \exp(-x)$ , odkiaľ  $x = -\ln(1 - u)$ . Ak  $u \sim U(0, 1)$ , potom  $x = \exp(-u)$ .

## 5.5 Jednoduché a dôležité vzorkovanie



**Obr. 5.4:** Zobrazenie spôsobu dosiahnutia nerovnomerného rozdelenia transformáciou premennej z kumulatívneho rozdelenia  $P(x)$ .

- **Boxova-Mullerova metóda pre generovanie gaussovského rozdelenia:** Nech  $x_1, x_2 \sim U(0, 1)$ , potom pre

$$y_1 = (-2 \log x_1)^{1/2} \cos(2\pi x_2) \quad (5.30)$$

$$y_2 = (-2 \log x_1)^{1/2} \sin(2\pi x_2) \quad (5.31)$$

platí  $y_1, y_2 \sim N(0, 1)$ , t. j. že sú z gaussovského (normálneho) rozdelenia so strednou hodnotou 0 a štandardnou odchýlkou 1.

## 5.5 Jednoduché a dôležité vzorkovanie

### Monte Carlo integrovanie

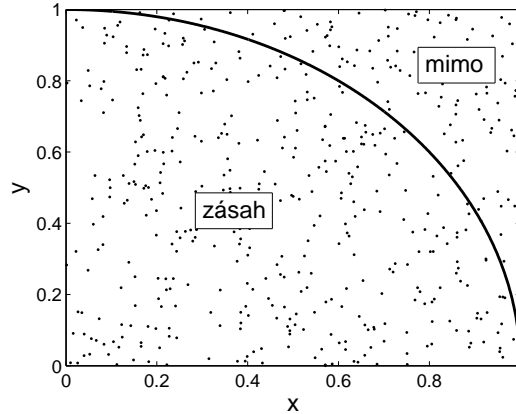
Jedným z najjednoduchších spôsobov numerického integrovania (kvadratury) je odhad pomocou obdĺžnikovej metódy (viď kurz numerických metód):

$$\int_a^b f(x) dx \approx \sum_{n=1}^M f(x_n) \Delta x, \quad (5.32)$$

kde  $f(x_n)$  sú funkčné hodnoty funkcie v uzloch siete  $x_n, n = 1, \dots, M$  a  $\Delta x = x_{n+1} - x_n$  je konštantný krok delenia intervalu  $[a, b]$ . Pri Monte Carlo integrovaní vygenerujeme postupnosť náhodných čísiel  $r_n \in [a, b]$  a

$$\int_a^b f(x) dx \approx \sum_{n=1}^M \frac{b-a}{M} f(r_n). \quad (5.33)$$

## 5.5 Jednoduché a dôležité vzorkovanie



**Obr. 5.5:** 25 000 náhodne generovaných bodov  $(x, y) \in [0, 1]$ , pričom body pod (nad) krivkou (ne)splňajú podmienku  $x^2 + y^2 \leq 1$

Teda, rozdiel je len v tom že pri MC integrovaní sú uzly siete náhodné čísla. Majme však  $N$ -rozmernú funkciu  $f(x_1, \dots, x_N)$  v jednotkovej hyperkocke  $[0, 1]^N$ . Potom integrál

$$\int_0^1 \int_0^1 \dots \int_0^1 f(\mathbf{x}) dx_1 dx_2 \dots dx_N \quad (5.34)$$

je veľmi problematické riešiť štandardnou numerickou metódou. Napríklad, len pre  $M = 10$  musíme funkciu  $f(x_1, \dots, x_N)$  vyhodnocovať  $10^N$ -krát, čo je nezvládnutelné pre veľké  $N \approx 10^6$ .

Ukázali sme si že chyba MC odhadu klesá s počtom vzoriek ako  $1/\sqrt{M}$ , a teda, pre dostatočne veľké  $M$  môžeme pomocou MC integrovania dostať dostatočne presné riešenie.

**Príklad:** Výpočet hodnoty  $\pi$  pomocou MC integrovania

Majme štvorec:  $\{(x, y) | 0 \leq x < 1, 0 \leq y < 1\}$ , ktorého obsah je  $S_{st} = 1$  a vo vnútri štvorca kruhový výsek:  $\{(x, y) | 0 \leq x < 1, 0 \leq y < 1, x^2 + y^2 \leq 1\}$ , ktorého obsah je  $S_{kv} = \pi 1^2/4 = \pi/4$  (vid' obr. 5.5). MC prístup je založený na generovaní  $N_{pok}$  párov náhodných čísel  $(x, y) \in [0, 1]$  a testovaní podmienky  $x^2 + y^2 \leq 1$ . Ak je splnená tak zvýšime počet zásahov  $N_{zas}$ . Potom plochu kruhového výseku môžeme odhadnúť ako  $S_{kv} \approx \frac{N_{zas}}{N_{pok}} \times S_{st} = \frac{N_{zas}}{N_{pok}}$ , odkiaľ vyplýva  $\pi \approx 4 \frac{N_{zas}}{N_{pok}}$ .

### Aplikácia na štatistickú mechaniku

Uvažujme kocku s dĺžkou hrany  $L$ , ktorá obsahuje  $N$  častíc. Štatistickú fyziku zaujíma dlhooasový priemer rôznych veličín popisujúcich taký systém  $\langle A(\{\mathbf{r}_i\}) \rangle$ , kde  $\{\mathbf{r}_i =$

## 5.5 Jednoduché a dôležité vzorkovanie

$(x_i, y_i, z_i) | 0 < x_i, y_i, z_i < L, i = 1, 2, \dots, N$ , ktorý je daný integrálom

$$\langle A(\{\mathbf{r}_i\}) \rangle = \frac{\int \int \int \dots \int \exp\left(-\frac{V(\{\mathbf{r}_i\})}{k_B T}\right) A(\{\mathbf{r}_i\}) d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N}{\int \int \int \dots \int \exp\left(-\frac{V(\{\mathbf{r}_i\})}{k_B T}\right) d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N}, \quad (5.35)$$

kde  $T$  je teplota,  $k_B = 1.38 \times 10^{-16}$  erg/K je Boltzmannova konštanta a  $V(\{\mathbf{r}_i\})$  je potenciálna energia systému. Príkladom veličiny  $\langle A(\{\mathbf{r}_i\}) \rangle$  môžu byť vzdialenosti medzi párami častíc a ak majú častice náboj  $e$  tak potenciálna energia je daná ako

$$V(\{\mathbf{r}_i\}) = \sum_{i < j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (5.36)$$

Je zrejmé že pri špičkových simuláciach, ktoré obsahujú až 100 miliónov častíc, výpočet 300 miliónov-rozmerného integrálu pomocou klasického numerického integrovania nie je možný.

### 5.5.1 Jednoduché vzorkovanie

MC simulácie založené na jednoduchom vzorkovaní (z angl.: simple sampling) by generovali náhodný bod v 300 miliónov-rozmernom priestore generovaním 300 miliónov náhodných čísel z intervalu  $[0, L]$  a potom vyhodnotilo funkciu v tomto bode. To by však bol veľmi neúčinný spôsob, vhládom na to že, hlavne pri nízkych teplotách, okrem blízkeho okolia bodu kde  $V(\{\mathbf{r}_i\})$  nadobúda minimum by hodnota výrazu  $\exp\left(-\frac{V(\{\mathbf{r}_i\})}{k_B T}\right)$  (tzv. Boltzmannov faktor) bola takmer nulová a teda aj ich príspevok k celkovému integrálu.

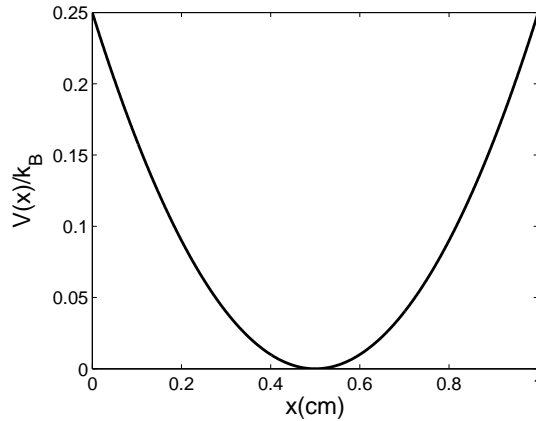
**Príklad:** Uvažujme jednu časticu v 1-rozmernom priestore dĺžky 1 cm a potenciáli  $V(x) = ck_B(x - 0.5)^2$ , kde  $c = 1\text{K}/\text{cm}^2$  (obr. 5.6).

Pre najnižšiu teplotu  $T=0.001\text{K}$  intervaly  $[0, 0.4]$  a  $[0.6, 1]$  prispievajú k integrovaniu zanedbateľne. Teda uniformným generovaním náhodných čísel márnime 80% času. Avšak, pri  $N = 100$  časticach a rovnakom potenciáli len  $0.2^{100} = 1.27 \times 10^{-70}$  z celkového počtu náhodných čísel prispieva a teda len zhruba každý  $10^{70}$ -tý pokus je užitočný.

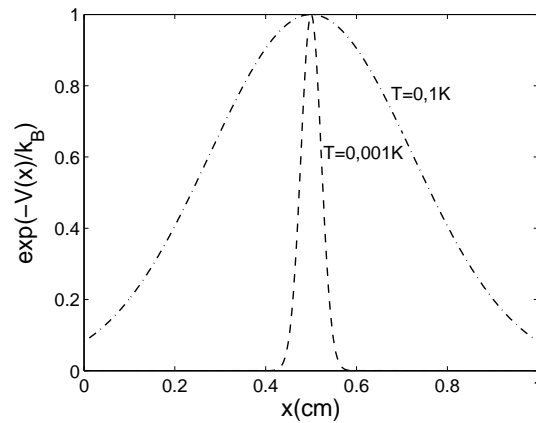
Efektívna MC metóda by teda mala generovať náhodné čísla v blízkosti minima potenciálu, teda dôležité body. Metóda na vzorkovanie takýchto dôležitých bodov sa nazýva dôležité vzorkovanie (z angl.: importance sampling).

## 5.5 Jednoduché a dôležité vzorkovanie

---



(a)



(b)

**Obr. 5.6:** (a) Potenciál  $V(x)/k_B = (x - 0.5)^2$  a (b) Boltzmannov faktor pre dve hodnoty teploty  $T = 0,1\text{K}$  a  $0,001\text{K}$ .

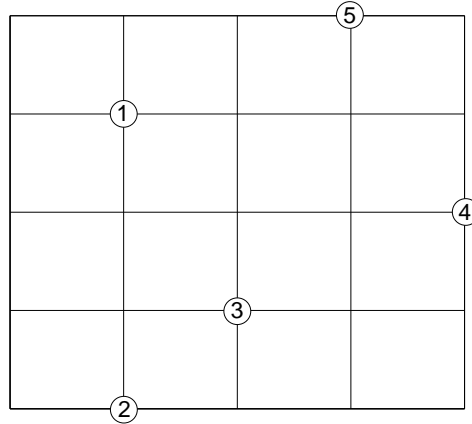
### 5.5.2 Dôležité vzorkovanie

V predchádzajúcej sekcii sme predpokladali že generujeme postupnosť náhodných bodov takých, že ich populácia je z danej hustoty pravdepodobnosti  $p(x)$ . Dôležité vzorkovanie vyberá náhodné čísla z hustoty pravdepodobnosti  $p(x)$  tak aby ich celková populácia, a nie každý bod, bola z danej hustoty pravdepodobnosti.

Uvažujme konečnostavový systém, ktorý v ľubovoľnom momente môže byť v jednom zo stavov  $\{\Gamma_1, \Gamma_2, \dots, \Gamma_N\}$ . Ako príklad si môžeme predstaviť  $n$  atómov v kocke, v ktorej sú súradnice v každom smere diskretizované tak, že môžu nadobúdať len  $m$

## 5.5 Jednoduché a dôležité vzorkovanie

---



**Obr. 5.7:** Dvozmerný systém so štyrmi atómami s diskretizovanými stavmi v piatich bodoch v každom smere. Počet všetkých kombinácií možných stavov atómov je  $5^8$ .

diskrétnych hodnôt. Stav systému je potom charakterizovaný  $3n$  diskrétnymi bodmi, pričom existuje  $m^{3n}$  kombinácií možných stavov (viď príklad v 2D na obr. 5.7). Hustota pravdepodobnosti je  $p(\Gamma_1, \Gamma_2, \dots, \Gamma_N)$ .

### Markovov reťazec

Markovov reťazec je postupnosť pokusov, ktorá spĺňa dve podmienky:

1. Výsledok každého pokusu patrí do konečnej množiny všetkých pokusov  $\{\Gamma_1, \Gamma_2, \dots, \Gamma_N\}$ , ktorú nazývame stavový priestor;
2. Výsledok každého pokusu závisí len od výsledku pokusu ktorý mu bezprostredne predchádzal (systém má veľmi krátku pamäť).

### Matica pravdepodobnosti prechodu

Ľubovoľné dva stavy  $\Gamma_n$  a  $\Gamma_m$  sú prepojené pravdepodobnosťou prechodu  $\pi_{mn}$ , ktorá je podmienenou pravdepodobnosťou, že ďalší stav bude  $\Gamma_m$  za predpokladu, že súčasný stav je  $\Gamma_n$ , t. j. pravdepodobnosť prechodu  $n \rightarrow m$ . Každá taka pravdepodobnosť pre-

## 5.5 Jednoduché a dôležité vzorkovanie

---

chodu  $\pi_{mn}$  je prvkom matice pravdepodobnosti prechodu

$$\Pi = \begin{bmatrix} \pi_{11} & \pi_{12} & \dots & \pi_{1N} \\ \pi_{21} & \pi_{22} & \dots & \pi_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ \pi_{N1} & \dots & \dots & \pi_{NN} \end{bmatrix}.$$

**Normalizácia:** Za predpokladu že je systém v súčasnosti v stave  $\Gamma_n$ , následujúci stav musí byť jeden z  $N$  možných stavov a teda musí platiť normalizačná podmienka

$$\sum_{m=1}^N \pi_{mn} = 1. \quad (5.37)$$

Predpokladajme, že začneme so 100-percentnou pravdepodobnosťou z jedného zo stavov, povedzme  $n$ , potom následujúci stav je náhodný s pravdepodobnosťou

$$\rho^{(1)} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_N \end{pmatrix} = \Pi \rho^{(0)} = \Pi \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix} n. \quad (5.38)$$

Rozdelenie pravdepodobnosti v časovom kroku  $\tau$  je dané postupným násobením matíc

$$\rho^{(\tau)} = \Pi \rho^{(\tau-1)} = \Pi^2 \rho^{(\tau-2)} = \dots = \Pi^\tau \rho^{(0)}. \quad (5.39)$$

Následujúca teoréma nám poskytne predstavu ako je Markovov reťazec použitý na dosiahnutie populácie so žiadanou hustotou pravdepodobnosti.

### Perronova-Frobeniova teoréma

Matica pravdepodobnosti prechodu má jednu vlastnú hodnotu rovnú jednej a prisluchajúci vlastný vektor je limitným rozdelením reťazca. Všetky ostatné vlastné hodnoty sú menej ako jedna.

**Dôsledok 1:**  $\Pi^\tau$  pre veľké  $\tau$  konverguje k matici, ktorú keď aplikujeme na ľubovoľné počiatkové rozdelenie pravdepodobnosti,  $\rho^{(0)}$ , konverguje k vlastnému vektoru prisluchajúcemu jednotkovej vlastnej hodnote. Všetky ostatné vlastné vektory, ktorých vlastné hodnoty sú menej ako jedna, sú odfiltrované postupným násobením matíc ( $\rho^\infty = \lim_{\tau \rightarrow \infty} \Pi^\tau \rho^{(0)}$ ).

**Dôsledok 2:** Aby sme dosiahli rozdelenie stavov z danej hustoty pravdepodobnosti, je potrebné nájsť takú maticu hustoty pravdepodobnosti (existencia ktorej je garantovaná vyššie uvedenou normalizačnou podmienkou), ktorej vlastný vektor odpovedajúci jednotkovej vlastnej hodnote je rovný žiadanej hustote pravdepodobnosti.



## 5.5 Jednoduché a dôležité vzorkovanie

---

### 5.5.3 Metropolisov algoritmus

Teraz sa úloha generovania žiadaneho rozdelenia pravdepodobnosti zredukovala na nájdenie takej matice hustoty pravdepodobnosti, ktorej vlastný vektor odpovedajúci jednotkovej vlastnej hodnote sa rovná žiadanej funkcii hustoty pravdepodobnosti. Aby sme to dosiahli, spomeňme si že pre vlastný vektor odpovedajúci jednotkovej vlastnej hodnote platí

$$\Pi\rho = \rho \text{ alebo } \sum_{n=1}^N \pi_{mn}\rho_n = \rho_m. \quad (5.40)$$

Problémom je, ako pre dané  $\rho$  nájsť  $\Pi$  ktoré spĺňa vyššie uvedenú podmienku. Môžeme to dosiahnuť požadovaním splnenia nasledovnej silnejšej podmienky.

**Podmienka detailnej rovnováhy:** Pre ľubovoľný pár stavov vyžadujeme aby bolo splnené

$$\pi_{mn}\rho_n = \pi_{nm}\rho_m. \quad (5.41)$$

Ľavá strana predstavuje tok pravdepodobnosti že aktuálny stav je  $n$  a nasledujúci stav je  $m$  a pravá strana zasa tok z  $m$  do  $n$ . Pre ľubovoľný pár stavov musí byť splnená mikroskopická rovnováha týchto tokov.

Aby sme dokázali, že podmienka detailnej rovnováhy je postačujúcou podmienkou pre splnenie rovnice (5.40), urobme sumáciu obidvoch strán cez  $n$ . Potom pre pravú stranu platí

$$\sum_{n=1}^N \pi_{nm}\rho_n = \sum_{n=1}^N (\pi_{nm})\rho_m = 1\rho_m = \rho_m \quad (5.42)$$

a pre ľavú stranu dostávame z rovnice (5.40) rovnaký výsledok, t. j.,

$$\sum_{n=1}^N \pi_{mn}\rho_n = \rho_m. \quad (5.43)$$

**Metropolisova matica pravdepodobnosti prechodu** je definovaná:

$$\pi_{mn} = \begin{cases} \alpha_{mn} & \rho_m \geq \rho_n, m \neq n, \\ (\rho_m/\rho_n)\alpha_{mn} & \rho_m < \rho_n, m \neq n, \\ 1 - \sum_{m' \neq n} \pi_{m'n} & m = n, \end{cases} \quad (5.44)$$

## 5.5 Jednoduché a dôležité vzorkovanie

kde  $\alpha$  je symetrická matica pokusov ( $\alpha_{mn} = \alpha_{nm}$ ). To znamená, že Metropolisova metóda je istý typ odmietacej metódy. Najprv sa pokúsí o prechod zo stavu  $n$  do  $m$  s pravdepodobnosťou  $\alpha_{mn}$ . Ak je pravdepodobnosť výsledného stavu väčšia než pravdepodobnosť súčasného stavu, potom je pokus bezpodmienečne akceptovaný. V opačnom rípane, ak je pravdepodobnosť výsledného stavu menšia než pravdepodobnosť súčasného stavu, potom je pokus akceptovaný s pravdepodobnosťou  $\rho_m/\rho_n$ . Posledná rovnosť vo vzťahu (5.44) garantuje splnenie normalizačnej podmienky pre maticu pravdepodobnosti prechodu.

**Dôkaz splnenia podmienky detailnej rovnováhy:** Predpokladajme, že  $\rho_m < \rho_n$  (podobný dôkaz môžeme urobiť aj pre opačný prípad).  $\pi_{nm}\rho_m = \alpha_{nm}\rho_m = \alpha_{mn}\rho_m$  kôli symetrickému pokusu. Na druhej strane,  $\pi_{mn}\rho_n = (\rho_m/\rho_n)\alpha_{mn}\rho_n = \alpha_{mn}\rho_m$ .

Zjavnou výhodou Metropolisovho algoritmu je to, že namiesto ťažko počítateľnej absolútnej pravdepodobnosti  $\rho_n$  stačí spočítať oveľa jednoduchšiu relatívnu hustotu pravdepodobnosti  $\rho_m/\rho_n$ . Napríklad, uvažujme hustotu pravdepodobnosti v štatistickej mechanike

$$P(\mathbf{r}_i) = \frac{\exp(-V(\mathbf{r}_i)/k_B T)}{\int \int \dots \int \exp(-V(\mathbf{r}_i)/k_B T) d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N}, \quad (5.45)$$

ktorú dostaneme až po integrácii exponenciálneho faktoru cez celý konfiguračný priestor, čo je prakticky neriešiteľná úloha. Naproti tomu, Metropolisova metóda vyžaduje len výpočet relatívnej pravdepodobnosti

$$\frac{P(\mathbf{r}'_i)}{P(\mathbf{r}_i)} = \frac{\exp(-V(\mathbf{r}'_i)/k_B T)}{\exp(-V(\mathbf{r}_i)/k_B T)} = \exp(-(V(\mathbf{r}'_i) - V(\mathbf{r}_i))/k_B T), \quad (5.46)$$

kde nečiarkované  $\mathbf{r}_i$  reprezentujú súradnice súčasného (pred pokusom o zmenu) a čiarkované  $\mathbf{r}'_i$  súradnice nového (po pokuse o zmenu) stavu. Môžeme si všimnúť, že problematický menovateľ sa vykrátil a teda na výpočet relatívnej pravdepodobnosti nám stačí vypočítať rozdiel potenciálnych energií po pokuse a pred pokusom o zmenu stavu.

**Pozor!** Systém zostáva v tom istom stave v prípade, že pokus o zmenu zlyhá podľa tretieho kritéria. Keď počítame štatistický priemer nejakej veličiny, aj takéto nový (aj keď nezmenený) stav sa musí brať do úvahy, okrem už započítaného predchádzajúceho stavu. Nasledujúci pseudo-program sumarizuje celý algoritmus.

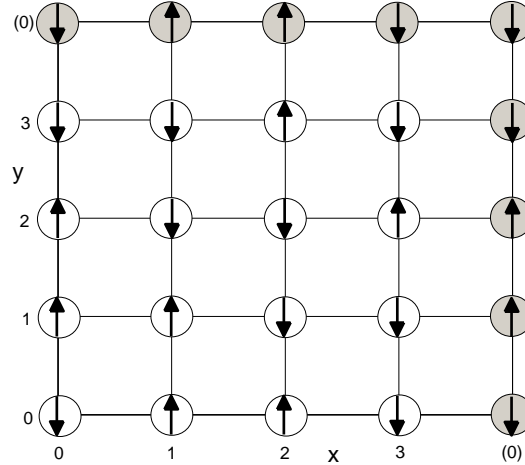
### Metropolisov algoritmus pre štatistickú fyziku

*Sum*  $A = 0$  (*Vynulovanie hodnoty veličiny A*)

Pre  $n = 1$  do  $N$  *pok* (*Opakovanie cyklu po max. počet pokusov N\_pok*)

$X' \leftarrow X + dX$  (*Pokus o zmenu stavu z X na X'*)

## 5.6 MC simulácie Isingovho modelu



**Obr. 5.8:** Spinová konfigurácia dvojrozmerného Isingovho modelu na mriežke veľkosti  $4 \times 4$ .

Vypočítaj  $dV = V(X') - V(X)$  (Výpočet zmeny potenciálnej energie)

Ak  $dV < 0$  potom

Akceptuj  $X \leftarrow X'$  (Bezpodmienečne akceptuj zmenu stavu)

Ináč ak  $\text{rand}() < \exp(-dV/k_B T)$  potom

Akceptuj  $X \leftarrow X'$  (Akceptuj zmenu stavu s pravdep.  $\rho_m/\rho_n$ )

Ináč odmietni zmenu stavu (Zachovaj pôvodný stav)

$\text{Sum}_A = \text{Sum}_A + A(X)$  (Kumulatívne spočítavanie hodnoty  $A$ )

Koniec cyklu

$\text{Avg}_A = \text{Sum}_A / N_{\text{pok}}$  (Odhad strednej hodnoty  $A$  aritmetickým priemerom)

## 5.6 Isingov model

Isingov model je štatisticko-fyzikálny model, ktorý bol pôvodne zavedený za účelom študovania správania sa magnetických častíc v magnetickom poli. Model pozostáva zo súboru magnetických častíc, tzv. spinov, lokalizovaných v uzloch mriežky (viď. obr.5.8). Každý taký spin môže nadobúdať buď hodnotu  $+1$  (šípka hore) alebo  $-1$  (šípka dole). Potenciálna energia (hamiltonián) takéhoto systému je daný vzťahom

$$V(\mathbf{s}) = -J \sum_{(i,j)} s_i s_j - H \sum_i s_i, \quad (5.47)$$

kde  $s_i = \pm 1$ ,  $(i, j)$  sú indexy najbližších susedov na mriežke,  $H$  je vonkajšie magnetické pole a  $J$  je výmenná interakcia. Pre  $J > 0$  prvý člen nadobúda minimum keď sú všetky

## 5.6 MC simulácie Isingovho modelu

---

spiny orientované rovnakým smerom a pre  $H > 0$  druhý člen, ak sú všetky spiny orientované smerom hore. V štatistickej mechanike pravdepodobnosť stavov s energiou  $V$  pri teplote  $T$  je úmerná Boltzmannovmu faktoru  $\exp(-V/k_B T)$ .

Pre dvojrozmernú mriežku veľkosti  $L \times L$ , každý index spinu pozostáva z páru celých čísiel  $(i, j)$ ,  $0 \leq i, j \leq L - 1$ . Štyria susedia (sever, juh, východ, západ) mriežkového bodu  $(i, j)$  sú  $(i \pm 1, j)$  a  $(i, j \pm 1)$ .

**Periodické okrajové podmienky** sa obyčajne aplikujú pri simulácií veľkých mriežok ( $L \rightarrow \infty$ ) za účelom eliminácie povrchových efektov. Na obr.5.8 je to znázornené pomocnými spinmi šedej farby v riadku (0) a stĺpci (0), ktoré majú vždy rovnakú hodnotu ako spiny v riadku 0 a stĺpci 0. Tým je zabezpečené, že každý spin, vrátane tých okrajových po obvode mriežky, má rovnaký počet (štyroch) susedov.

### MC algoritmus pre Isingov model

Následujúci pseudo-program popisuje MC simuláciu dvojrozmerného Isingovho modelu. Simulácia pozostáva z postupnosti pokusov o preklopenie vybraného spinu  $(i, j)$  a výpočtu odpovedajúcej zmeny energie. Preklopenie je buď akceptované alebo zamietnuté podľa Metropolisovho algoritmu. V každom MC kroku je náhodne vybraný jeden spin a celkovo je realizovaných *max\_krok* krokov.

#### Algoritmus pre Isingov model

Inicializuj hodnoty spinov  $s_{i,j}$  pre všetky  $0 \leq i, j \leq L - 1$

$Sum_A = 0$  (*Vynulovanie hodnoty veličiny A*)

Pre  $n = 1$  do *max\_krok* (*Opakovanie cyklu po max. počet krokov max\_krok*)

Náhodne vyber uzol mriežky  $(i, j)$

Vypočítaj  $dV = V(\mathbf{s}') - V(\mathbf{s})$  pri preklopení spinu  $(i, j)$

Ak  $dV < 0$  potom bezpodmienečne akceptuj preklopenie,  $s_{i,j} \leftarrow -s_{i,j}$

Ináč ak  $\text{rand}() < \exp(-dV/k_B T)$  potom akceptuj preklopenie

Ináč ponechaj starý stav (bez preklopenia spinu)

$Sum\_A = Sum\_A + A(\mathbf{s})$  (*Kumulatívne spočítavanie hodnoty A*)

$Avg\_A = Sum\_A / max\_krok$  (*Po max\_krok krokoch odhad strednej hodnoty A aritmetickým priemerom*)

#### **Praktické poznámky:**

Zmenu energie,  $dV(\mathbf{s}) = V(\dots, s'_k, \dots) - V(\dots, s_k, \dots)$ , spojenej s preklopením jedného spinu  $s_k \rightarrow s'_k = -s_k$  vypočítame následovne

$$dV = -J \sum_{l \in n.n.(k)} (s'_k - s_k) s_l - H(s'_k - s_k), \quad (5.48)$$

kde  $n.n.(k)$  označuje najbližších susedov  $k$ -tého spinu. Všimnime si, že  $s'_k - s_k = 2s'_k$

## 5.6 MC simulácie Isingovho modelu

---

(2 pre  $-1 \rightarrow 1$  a  $-2$  pre  $1 \rightarrow -1$ ), takže

$$dV = -2s'_k \left( J \sum_{l \in n.n.(k)} s_l + H \right). \quad (5.49)$$

Pre dvojrozmerný Isingov model existujú len dve možné hodnoty pre  $s'_k : \pm 1$  a päť hodnôt pre  $\sum_{l \in n.n.(k)} : 0, \pm 2, \pm 4$ . Preto je výhodné zostaviť tabuľku (veľkosti  $2 \times 5$ ) pre všetky možné hodnoty  $\exp(-dV/k_B T)$ , z ktorej potom pri simulácii len vyberáme príslušajúce hodnoty namiesto pracného opakovaného výpočtu exponenciálneho výrazu. Vzhľadom na to že obyčajne potrebujeme veľa MC krokov, tým sa ušetrí obrovské množstvo počítačového času.

### Výpočet niektorých základných fyzikálnych veličín

Okrem energie, definovanej vzťahom (5.47), počítame magnetizáciu

$$M = \sum_i s_i \quad (5.50)$$

a taktiež je užitočné počítať aj rozptyl, jednak energie ako aj magnetizácie

$$Var\{X\} = \langle X^2 \rangle - \langle X \rangle^2 \quad (5.51)$$

kde  $X = V$  alebo  $M$ , ktoré sú úmerne fyzikálnym veličinám tepelná kapacita ( $C = Var\{V\}/k_B T^2$ ) a magnetická susceptibilita ( $\chi = Var\{M\}/k_B T$ ). Poznamenajme, že taktiež výpočet magnetizácie si je možné uľahčiť tým, že po každom preklopení spinu z  $s_k$  na  $s'_k$  aktualizujeme priebežnú hodnotu  $M_{tot}$  tak, že ju nahradíme  $M_{tot} + 2s'_k$ .

## Literatúra

PETROVIČ, P., NADRCHAL J., PETROVIČOVÁ J., *Programovanie a spracovanie dát I*, vyd. ES UPJŠ, Košice, 1989.

PŘIKRYL P., *Numerické metody matematické analýzy*, vyd. SNTL, Praha, 1988.

VITÁSEK, E., *Numerické metody*, vyd. SNTL, Praha, 1987.

POZRIKIDIS C., *Numerical Computation in Science and Engineering*, vyd. Oxford University Press, 1998.

CAP F.F., *Mathematical Methods in Physics and Engineering with Mathematica*, vyd. CRC Press, Boca Raton, FL, USA, 2003.

SCHEINERMAN E.R., *Invitation to Dynamical Systems*, vyd. Prentice Hall, 1995.

GIORDANO, N.J., NAKANISHI, H., *Computational Physics*, vyd. Prentice Hall, 2006.

QUARTERONI, A., SALERI, F., *Scientific Computing with MATLAB and Octave*, vyd. Springer-Verlag, Berlin Heidelberg, 2006.

BUŠA, J., PIRČ, V., SCHRÖTTER, Š., *Numerické metody, pravdepodobnosť a matematická štatistika*, vyd. TU Košice, 2006.

LANDAU, D.P., BINDER, K., *A Guide to Monte Carlo Simulations in Statistical Physics*, vyd. Cambridge University Press, 2000.

BERG, B.A., *Markov Chain Monte Carlo Simulations and Their Statistical Analysis*, vyd. World Scientific, Singapore, 2004.

NEWMAN, M.E.J., BARKEMA, G.T., *Monte Carlo Methods in Statistical Physics*, vyd. Clarendon Press, Oxford, 1999.

BINDER, K., HEERMANN, D.W., *Monte Carlo Simulations in Statistical Physics: An Introduction, 4th edition*, vyd. Springer, Berlin, 2002.

## **Počítačová fyzika I.**

Autor: doc. RNDr. Milan Žukovič PhD.,  
Vydavateľ: Univerzita Pavla Jozefa Šafárika v Košiciach  
Rozsah strán: 86  
Rozsah AH: 4,3  
Vydanie: prvé  
Umiestnenie: <http://unibook.upjs.sk/predaj-vydanych-titulov/prirodovedecka-fakulta>  
Dostupné od: 03.02.2015

ISBN 978-80-8152-228-4