# Numerické řešení rovnice vedení tepla

V této části se budeme zabývat numerickým řešením smíšené úlohy pro rovnici vedení tepla v jedné prostorové dimenzi. Hledáme funkci $u = u(x, t)$ definovanou pro $x \in [0, 1]$ a $t \geq 0$ takovou, že

$$u_t - u_{xx} = 0 \qquad \text{v } (0,1) \times \mathbb{R}^+ , \tag{1}$$

$$u(0,t) = u(1,t) = 0 \qquad \forall\, t > 0 , \tag{2}$$

$$u(x,0) = u^0(x) \qquad \forall\, x \in [0,1] . \tag{3}$$

Budeme předpokládat, že úloha (1)–(3) má klasické řešení. To mimo jiné vyžaduje, aby $u^0(0) = u^0(1) = 0$.

Úlohy tohoto typu popisují šíření tepla v tenké izolované homogenní tyči konečné délky bez přítomnosti tepelných zdrojů. Úloha typu (1)–(3) též popisuje šíření tepla napříč nekonečnou deskou, přičemž $x$ je souřadnice kolmá ke stěnám desky, na každé z nichž je předepsána konstantní teplota.

**Poznámka 1** Transformací $v(x,t) = u(x, \kappa\, t) + \alpha\,(1-x) + \beta\, x$ získáme úlohu s tepelnou difuzivitou $\kappa$ a nehomogenními Dirichletovými okrajovými podmínkami:

$$\begin{aligned}
&v_t - \kappa\, v_{xx} = 0 \qquad \text{v } (0,1) \times \mathbb{R}^+ ,\\
&v(0,t) = \alpha, \quad v(1,t) = \beta \qquad \forall\, t > 0 ,\\
&v(x,0) = v^0(x) \equiv u^0(x) + \alpha\,(1-x) + \beta\, x \qquad \forall\, x \in [0,1] .
\end{aligned}$$

Uvažování $\kappa = 1$ a homogenních okrajových podmínek v (1)–(3) tedy nepředstavuje újmu na obecnosti. Podobně lze též jednoduchou transformací proměnné $x$ přejít k úloze definované na libovolném zadaném prostorovém intervalu.

## Řešení úlohy (1)–(3) Fourierovou metodou

Hledejme řešení úlohy (1)–(3) ve tvaru $u(x,t) = X(x)\, T(t)$. Dosazením do (1) a (2) snadno zjistíme, že

$$u(x,t) = a_m\, e^{-(m\,\pi)^2\, t} \sin(m\,\pi\, x)$$

pro nějaké $m \in \mathbb{N}$ a $a_m \in \mathbb{R}$. Rovnice (1) a (2) splňuje zřejmě i libovolná lineární kombinace těchto funkcí. Nabízí se tedy hledat řešení $u$ ve tvaru

$$u(x,t) = \sum_{m=1}^{\infty} a_m\, e^{-(m\,\pi)^2\, t} \sin(m\,\pi\, x) . \tag{4}$$

Z počáteční podmínky (3) plyne

$$u^0(x) = \sum_{m=1}^{\infty} a_m \sin(m\,\pi\, x) , \tag{5}$$

a tudíž koeficienty $a_m$ jsou Fourierovy koeficienty funkce $u^0$ při rozvoji do sinové řady. Platí

$$a_m = 2 \int_0^1 u^0(x)\, \sin(m\,\pi\,x)\,\mathrm{d}x\,.$$

Je známo, že pro libovolné $u^0 \in L^2(0,1)$ řada (5) konverguje v $L^2(0,1)$. Součet řady (4) je pak nekonečně hladká funkce v $[0,1] \times \mathbb{R}^+$, která řeší parciální diferenciální rovnici (1) a splňuje okrajovou podmínku (2). Pokud řada Fourierových koeficientů $a_m$ konverguje absolutně, je součet řady (4) spojitá funkce na $[0,1] \times \mathbb{R}_0^+$, která splňuje počáteční podmínku (3). K tomu stačí, aby funkce $u^0$ byla absolutně spojitá na $[0,1]$, $(u^0)' \in L^2(0,1)$ a $u^0(0) = u^0(1) = 0$ (konzistence s okrajovou podmínkou).

V praxi výsledek (4) umožňuje získat pouze numerickou aproximaci řešení $u$, neboť koeficienty $a_m$ jsme obecně schopni určit pouze přibližně a navíc jsme schopni sečíst pouze konečně mnoho členů řady. Skutečným omezením uvedené metody však je, že ji nelze snadno zobecnit na komplikovanější úlohy. Je proto nutné hledat jiné způsoby výpočtu přibližného řešení úlohy (1)–(3) a jedním z možných postupů je aplikace metody konečných diferencí.

## Explicitní schéma pro úlohu (1)–(3)

Podobně jako dříve zavedeme rovnoměrnou síť, kterou pokryjeme výpočetní oblast. Interval $[0,1]$ nejprve rozdělíme na $J \in \mathbb{N}$ intervalů stejné délky $h = 1/J$, čímž vzniknou body $x_j = j\,h$, $j = 0,1,2,\ldots,J$. Kromě prostorového kroku sítě $h$ zavedeme též časový krok sítě $\tau > 0$ a definujeme časové hladiny $t_n = n\,\tau$, $n = 0,1,2,\ldots$. Řešení $u$ úlohy (1)–(3) aproximujeme v uzlech sítě $(x_j, t_n)$ hodnotami $U_j^n$, kde $j = 0,1,2,\ldots,J$ a $n = 0,1,2,\ldots$. Opět položíme $u_j^n = u(x_j, t_n)$.

Uvažujeme-li rovnici (1) v uzlu $(x_j, t_n)$ s $j \in \{1,\ldots,J-1\}$ a $n \in \mathbb{N}_0$, můžeme provést následující aproximace:

$$0 = (u_t - u_{xx})(x_j, t_n) \approx \frac{\Delta_{+t}\, u_j^n}{\tau} - \frac{\delta_x^2\, u_j^n}{h^2} \approx \frac{\Delta_{+t}\, U_j^n}{\tau} - \frac{\delta_x^2\, U_j^n}{h^2}\,.$$

To vede k diferenčním rovnicím

$$U_j^{n+1} = U_j^n + \mu\,(U_{j+1}^n - 2\,U_j^n + U_{j-1}^n)\,, \qquad j = 1,2,\ldots,J-1\,, \tag{6}$$

kde

$$\mu = \frac{\tau}{h^2}\,.$$

Každou hodnotu na časové hladině $t_{n+1}$ lze nezávisle spočítat z hodnot na časové hladině $t_n$, a proto se taková metoda nazývá *explicitní diferenční schéma*.

K rovnicím (6) musíme ještě přidat okrajové a počáteční podmínky

$$U_0^n = U_J^n = 0\,, \qquad n = 1,2,\ldots\,, \tag{7}$$

$$U_j^0 = u^0(x_j)\,, \qquad j = 0,1,\ldots,J\,. \tag{8}$$

Hodnoty přibližného řešení získáme tak, že nejprve definujeme hodnoty $U_j^0$ pomocí (8) a pak evolučně počítáme hodnoty přibližného řešení ná následujících časových hladinách z rovnic (6) a okrajových podmínek (7). Přibližné řešení $U_j^n$ je tedy vztahy (6)–(7) jednoznačně určeno.

**Cvičení 1** *Uvažujte schéma (6)–(7) pro úlohu (1)–(3) s počáteční podmínkou*

$$u^0(x) = \begin{cases} 2\,x & \text{pro} \quad x \in [0, \frac{1}{2}]\,, \\ 2 - 2\,x & \text{pro} \quad x \in [\frac{1}{2}, 1]\,. \end{cases}$$

*Zvolte $J = 20$ a proveďte výpočet pro $\tau = 0.0012$ a $\tau = 0.0013$.*

Provedeme-li výpočty z cvičení 1, zjistíme, že pro $\tau = 0.0012$ získáme dobrou aproximaci řešení úlohy (1)–(3), zatímco pro $\tau = 0.0013$ se v přibližném řešení objeví oscilace, které se zvětšujícím se $n$ rychle rostou. Jedná se o typický příklad stability či nestability numerického schématu. Jak uvidíme později, numerické výsledky zásadně závisejí na hodnotě $\mu$, tj. na vztahu mezi prostorovým a časovým krokem.

Chyba diskretizace schématu (6) je dána vztahem

$$\varepsilon_j^n = \frac{\Delta_{+t}\, u_j^n}{\tau} - \frac{\delta_x^2\, u_j^n}{h^2}$$

pro $j \in \{1, \ldots, J-1\}$ a $n \in \mathbb{N}_0$. Je-li $x \in [h, 1-h]$ a $t \geq 0$, můžeme též položit

$$\varepsilon_{h,\tau}(x,t) = \frac{\Delta_{+t}\, u(x,t)}{\tau} - \frac{\delta_x^2\, u(x,t)}{h^2}\,.$$

Pak $\varepsilon_j^n = \varepsilon_{h,\tau}(x_j, t_n)$. Použitím Taylorova vzorce získáme

$$\varepsilon_{h,\tau}(x,t) = \tfrac{1}{2}\, u_{tt}(x,\eta)\, \tau - \tfrac{1}{12}\, u_{xxxx}(\xi, t)\, h^2\,, \qquad \xi \in (x-h, x+h),\ \eta \in (t, t+\tau)\,.$$

Předpokládáme-li, že existují konstanty $M_1$ a $M_2$ takové, že $|u_{tt}| \leq M_1$, $|u_{xxxx}| \leq M_2$ na $[0,1] \times [0,T]$, kde $T > 0$ je pevně zvolený čas, pak

$$(9) \quad \varepsilon_{h,\tau}(x,t)| \leq \tfrac{1}{2}\, M_1\, \tau + \tfrac{1}{12}\, M_2\, h^2 = \tfrac{1}{2}\, \tau \left( M_1 + \tfrac{1}{6\,\mu}\, M_2 \right) \quad \forall\ x \in [h, 1-h] \times [0, T-\tau]\,.$$

Vidíme tedy, že pro pevný poměr $\mu$ se $\varepsilon_{h,\tau}$ chová jako $O(\tau)$ pro $\tau \to 0$. Říkáme, že schéma je *prvního řádu přesnosti* (lze též říci, že schéma je prvního řádu přesnosti v čase a druhého řádu přesnosti v prostoru).

**Poznámka 2** Jelikož $u_t = u_{xx}$, je $u_{tt} = u_{xxt} = (u_t)_{xx} = u_{xxxx}$, a tudíž

$$\varepsilon_{h,\tau}(x,t) = \tfrac{1}{2} \left( 1 - \tfrac{1}{6\,\mu} \right) u_{xxxx}(x,t)\, \tau + O(\tau^2)\,.$$

Pro $\mu = \frac{1}{6}$ je tedy schéma druhého řádu přesnosti. Jedná se ale o velmi speciální případ, se kterým se v obecnějších situacích nesetkáme.

## Konvergence explicitního schématu

**Věta 1** *Uvažujme posloupnost* $(h_i, \tau_i) \to (0,0)$ *pro* $i \to \infty$ *a předpokládejme, že* $\mu_i \equiv \tau_i/h_i^2 \le \frac{1}{2}$. *Nechť* $T > 0$ *a* $|u_{tt}| \le M_1$, $|u_{xxxx}| \le M_2$ *v* $[0,1] \times [0,T]$. *Pak pro libovolný bod* $(x,t) \in [0,1] \times [0,T]$ *a libovolnou posloupnost* $(j_i, n_i) \in \mathbb{N} \times \mathbb{N}$ *takovou, že* $j_i\, h_i \to x$, $n_i\, \tau_i \to t$, *konvergují aproximace* $U_{j_i}^{n_i}$ *generované explicitním schématem (6)–(7) k řešení* $u(x,t)$, *přičemž tato konvergence je stejnoměrná v* $[0,1] \times [0,T]$.

*Důkaz.* Uvažujme libovolnou dvojici $h, \tau$ ($\tau < T$) a libovolný bod $(x_j, t_n) \in (0,1) \times (0,T)$. Definujme *chybu aproximace*

$$e_j^n = U_j^n - u(x_j, t_n).$$

Pak pro $j = 1, \ldots, J-1$ platí

$$e_j^{n+1} = e_j^n + \mu\,[e_{j+1}^n - 2\,e_j^n + e_{j-1}^n] - \tau\,\varepsilon_j^n\,, \qquad \text{kde} \qquad \varepsilon_j^n = \varepsilon_{h,\tau}(x_j, t_n)\,.$$

Definujeme-li $\|e^n\|_\infty = \max_{l=1,\ldots,J-1} |e_l^n|$, pak (díky $e_0^n = e_J^n = 0$)

$$\|e^{n+1}\|_\infty \le (|1 - 2\,\mu| + 2\,\mu)\,\|e^n\|_\infty + \tau\,\|\varepsilon^n\|_\infty \le \|e^n\|_\infty + \tau\,\|\varepsilon^n\|_\infty\,.$$

Jelikož $e_l^0 = U_l^0 - u^0(x_l) = 0$, $l = 0, \ldots, J$, dostáváme

$$\|e^n\|_\infty \le \tau \sum_{k=0}^{n-1} \|\varepsilon^k\|_\infty \le t_n \max_{k=0,\ldots,n-1} \|\varepsilon^k\|_\infty\,.$$

Použitím (9) získáváme

$$|U_j^n - u(x_j, t_n)| \le T\,(\tfrac{1}{2}\,M_1\,\tau + \tfrac{1}{12}\,M_2\,h^2)\,.$$

Tvrzení nyní plyne ze spojitosti $u$ na $[0,1] \times [0,T]$. $\square$

## Fourierova analýza chyby

Víme, že řešení úlohy (1)–(3) lze zapsat ve tvaru Fourierovy řady (4). Ukážeme, že v podobném tvaru lze zapsat i přibližné řešení splňující (6)–(7). Za tím účelem zapíšeme řady (4) a (5) pomocí komplexních exponenciel:

$$(10) \qquad u(x,t) = \sum_{m=-\infty}^{\infty} A_m\, e^{\mathrm{i}\,m\,\pi\,x - (m\,\pi)^2\,t}\,,$$

$$(11) \qquad u^0(x) = \sum_{m=-\infty}^{\infty} A_m\, e^{\mathrm{i}\,m\,\pi\,x}\,.$$

Položíme-li $u^0(x) = -u^0(-x)$ pro $x \in [-1, 0)$, pak

$$A_m = \frac{1}{2} \int_{-1}^{1} u^0(x)\, e^{-\mathrm{i}\,m\,\pi\,x}\,\mathrm{d}x\,.$$

Snadno se ověří, že pro $m \in \mathbb{N}$ je $A_m = -A_{-m} = -\mathrm{i}\, a_m/2$. Budeme předpokládat, že Fourierova řada (11) je absolutně konvergentní, tj.

$$\sum_{m=-\infty}^{\infty} |A_m| < \infty\,. \tag{12}$$

Jak víme, postačující podmínkou pro to je, aby funkce $u^0$ byla absolutně spojitá na $[0,1]$, $(u^0)' \in L^2(0,1)$ a $u^0(0) = u^0(1) = 0$.

Připomeňme, že funkce $e^{\mathrm{i}\, m\,\pi\, x - (m\,\pi)^2\, t}$ jsou řešeními rovnice (1). V uzlech sítě platí

$$e^{\mathrm{i}\, m\,\pi\, x_j - (m\,\pi)^2\, t_n} = e^{\mathrm{i}\, k\, j\, h} \left[e^{-k^2\,\tau}\right]^n,$$

kde jsem zavedli *vlnové číslo* $k = m\,\pi$. V analogii k tomu se můžeme ptát, kdy

$$U_j^n = e^{\mathrm{i}\, k\, j\, h}\, \lambda^n$$

řeší diferenční rovnici (6). Dosazením do (6) získáme

$$e^{\mathrm{i}\, k\, j\, h}\, \lambda^{n+1} = e^{\mathrm{i}\, k\, j\, h}\, \lambda^n\, [1 + \mu\, (e^{\mathrm{i}\, k\, h} - 2 + e^{-\mathrm{i}\, k\, h})],$$

a tudíž

$$\lambda \equiv \lambda(k) = 1 - 2\,\mu\, [1 - \cos(k\, h)] = 1 - 4\,\mu\, \sin^2 \frac{k\, h}{2}\,. \tag{13}$$

Číslo $\lambda(k)$ se nazývá amplifikační (zesilující) faktor členu Fourierovy řady. Následující věta ukazuje, že přibližné řešení $U_j^n$ můžeme zapsat ve tvaru podobném jako $u$.

**Věta 2** *Nechť síťová funkce $U_j^n$ splňuje (6)–(8). Pak*

$$U_j^n = \sum_{m=-\infty}^{\infty} A_m\, e^{\mathrm{i}\, m\,\pi\, j\, h}\, [\lambda(m\,\pi)]^n\,, \qquad j = 0, 1, \ldots, J\,, \quad n \geq 0\,. \tag{14}$$

*Důkaz.* Jelikož amplifikační faktory jsou omezené a platí (12), konverguje řada (14) absolutně, a jelikož každý její člen řeší diferenční rovnici (6), řeší i součet řady (14) rovnici (6). Dále pro každé $m \in \mathbb{N}$ platí $A_m\, [\lambda(m\,\pi)]^n = -A_{-m}\, [\lambda(-m\,\pi)]^n$, a tudíž je pro $j = 0$ a $j = J$ součtem řady 0. Jsou tedy splněny okrajové podmínky (7). Konečně pro $n = 0$ se řada (14) redukuje na řadu (11) s $x = j\, h$, a součet řady tudíž splňuje i počáteční podmínku (8). $\square$

**Poznámka 3** Všimněme si, že z vlastností amplifikačního faktoru a koeficientů $A_m$ plyne, že řadu (14) lze též přepsat do tvaru

$$U_j^n = \sum_{m=1}^{\infty} a_m\, \sin(m\,\pi\, j\, h)\, [\lambda(m\,\pi)]^n\,, \qquad j = 0, 1, \ldots, J\,, \quad n \geq 0\,. \tag{15}$$