VŠB 1849 OSTRAVA TECHNICKÁ UNIVERZITA

ZÁPADOČESKÁ UNIVERZITA V PLZNI

NUMERICKÉ METODY PRO ŘEŠENÍ EVOLUČNÍCH PARCIÁLNÍCH DIFERENCIÁLNÍCH ROVNIC

Marek Brandner
Jiří Egermaier
Hana Kopincová

Text byl vytvořen v rámci realizace projektu *Matematika pro inženýry 21. století* (reg. č. CZ.1.07/2.2.00/07.0332), na kterém se společně podílela Vysoká škola báňská – Technická univerzita Ostrava a Západočeská univerzita v Plzni

evropský sociální fond v ČR | EVROPSKÁ UNIE | MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY | OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Marek Brandner, Jiří Egermaier, Hana Kopincová
Numerické metody pro řešení evolučních parciálních diferenciálních rovnic

ISBN

V Plzni 31. 1. 2012 Marek Brandner, Jiří Egermaier, Hana Kopincová

Obsah

Kapitola 1

Základní teoretické poznatky

Průvodce studiem

V této kapitole se budeme zabývat základními teoretickými poznatky týkajícími se evolučních parciálních diferenciálních rovnic. První část kapitoly bude zaměřena na hyperbolické parciální diferenciální rovnice a jejich soustavy. Uvedeme příklady matematických modelů, které vedou na hyberbolické rovnice a soustavy a hlavně zde rozebereme speciální úlohu s nespojitými počátečními podmínkami nazývanou klasický Riemannův problém. Tato úloha a její řešení bude podrobně popsána pro různé typy možných rovnic i soustav, tj. lineární i nelineární. Na konci této části přiblížíme tzv. zobecněný Riemannův problém, osvětlíme souvislost s klasickým Riemannovým problémem a zavedeme pojem Hugoniotových množin. Druhá část kapitoly se bude stručně věnovat parabolickým evolučním rovnicím, kde bude uvedena základní formulace problému, jeho řešení a základní teoretické poznatky, potřebné pro studium numerických metod v dalších kapitolách.

Cíle

Po prostudování této kapitoly budete schopni:

- formulovat Riemannův problém a představit si jeho řešení pro všechny typy rovnic a soustav,
- porozumět pojmům potřebným k řešení hyperbolických a parabolických rovnic, jako jsou charakteristiky, princip maxima, apod.
- porozumět základním teoretickým principům z oblasti evolučních rovnic, na jejichž základě je možné porozumět i následným numerickým metodám a postupům.

1.1 Příklady matematických modelů

Matematické modelování různých fyzikálních situací vychází z využití zákonů zachování (nejčastěji hmotnosti, momentu, energie) nebo bilančních vztahů (nenulové zdrojové členy). Jedním z nejobecnějších modelů z oblasti modelování proudění je založen na Navierových-Stokesových rovnicích, popisujících proudění vazkých stlačitelných i nestlačitelných tekutin. Často však není nutné používat takto obecný model, proto za jistých vhodných omezujících předpokladů lze získat modely mnohem jednodušší. Je důležité v úvodu zmínit, že zákony zachování či bilanční vztahy jsou integrální rovnosti, neboť uvažujeme hmotnost, moment nebo energii v určité oblasti a v určitém časovém intervalu. Častěji je však nalezneme zapsané v jednodušším diferenciálním tvaru, který se často používá pouze jako reprezentace původních obecných integrálních tvarů zákonů zachování nebo bilancí.

Příkladem, jak zjednodušit Navierovy-Stokesovy rovnice pro modelování říčního proudění korytem neměnného obdélníkového průřezu v jedné dimenzi jsou tzv. Saint-Venantovy rovnice (reprezentují zákon zachování hmotnosti a momentovou bilanci), které v diferenciálním tvaru lze zapsat následovně

$$\begin{aligned} h_t + (hv)_x &= 0, \\ (hv)_t + \left(hv^2 + \frac{1}{2}gh^2\right)_x &= -ghb_x. \end{aligned} \tag{1.1}$$

Funkce $h(x,t)$ reprezentuje hledanou hloubku, $v(x,t)$ horizontální rychlost, $q(x,t) = hv$ průtok, $b = b(x)$ je daná funkce popisující tvar dna a g je tíhová konstanta. Jak jsme již popsali ve skriptech [1], lze zvolit i jiný tvar koryta.

Dalším modelem, velmi často využívaným v modelování proudění stlačitelné nevazké tekutiny v jedné dimenzi, jsou tzv. Eulerovy rovnice (za předpokladu chemické a termodynamické rovnováhy a vnitřní energie je známá funkce tlaku a hustoty), reprezentující zákon zachování hmotnosti, hybnosti a energie, které opět v diferenciálním tvaru zapisujeme následovně

$$\begin{aligned} \rho_t + (\rho u)_x &= 0, \\ (\rho u)_t + (\rho u^2 + p)_x &= 0, \\ E_t + (Eu + up)_x &= 0, \end{aligned} \tag{1.2}$$

kde $\rho(x,t)$ reprezentuje hledanou hustotu plynu, $u(x,t)$ je hledaná rychlost proudění, $E(x,t)$ reprezentuje hledanou hustotu celkové energie a $p(x,t)$ je tlak.

Dále uveďme model dopravního proudu, též označovaný jako LWR (Lighthill, Whitham, Richards) model (více lze nalézt v [9])

$$u_t + \left[v_{max}\left(1 - \frac{u}{u_{max}}\right)u\right]_x = 0, \tag{1.3}$$

kde $u = u(x,t)$ reprezentuje hustotu aut (jde o relativní veličinu, a tudíž bezrozměrnou). Konstanta v_{max} je rychlost při nulové hustotě, tj. když je prázdná silnice. u_{max} je maximální hustota aut, tj. úplně plná silnice (obvykle klademe $u_{max} = 1$).

Posledním modelem, který zde uvedeme je jednoduchá nelineární skalární rovnice popisujícím dopravní proud, tzv. Burgersova rovnice. Zde uvedeme jednak homogenní tzv. nevazkou Burgersovu rovnici

$$u_t + \left(\frac{1}{2}u^2\right) = 0, \tag{1.4}$$

kde $u(x,t)$ reprezentuje rychlost dopravního proudu. Vazká Burgersova rovnice je nehomogenní nelineární skalární parabolicko-hyperbolická (advekčně-difúzní) rovnice následujícího tvaru

$$u_t + \left(\frac{1}{2}u^2\right) = \epsilon u_{xx}, \tag{1.5}$$

kde ϵ popisuje viskozitu. Burgersova rovnice se využívá spíše než k modelování fyzikálních situací, jako příklad jednoduché nelineární rovnice k ilustrování některých klíčových vlastností modelování proudění tekutin.

1.2 Základní definice

Na úvod pouze ve stručnosti shrneme základní formulaci problému a definice klasického, slabého a entropického řešení a dalších pojmů, které budeme potřebovat. Všechny jsou k nalezení např. v [4] nebo v [9], nebo jsou podrobně vysvětleny ve skriptech [1]. Problémy (1.1) s nulovou pravou stranou (tj. rovným dnem), (1.2), (1.3), (1.4) a (1.5) s nulovou pravou stranou (tj. nulovou viskozitou), lze formulovat jako počáteční úlohu

$$\begin{aligned} \mathbf{u}_t + [\mathbf{f}(\mathbf{u})]_x &= \mathbf{0}, \quad x \in \mathbb{R},\, t \in (0,T),\, T > 0, \\ \mathbf{u}(x,0) &= \mathbf{u}_0(x), \quad x \in \mathbb{R}, \end{aligned} \tag{1.6}$$

kde $\mathbf{u} = \mathbf{u}(x,t) : \mathbb{R} \times (0,T) \to \mathbb{R}^m$ je hledaná funkce (vektor zachovávaných veličin), $\mathbf{u}_0 = \mathbf{u}_0(x) : \mathbb{R} \to \mathbb{R}^m$ a $\mathbf{f} = \mathbf{f}(\mathbf{u}) : \mathbb{R}^m \to \mathbb{R}^m$ jsou dané funkce. Předpokládáme, že funkce $\mathbf{f} = \mathbf{f}(\mathbf{u})$ je dostatečně hladká.

Definice 1.1. *Klasickým řešením* úlohy (1.6) nazveme $\boldsymbol{u}(x,t)$ takové, že $\boldsymbol{u} \in [\mathrm{C}(\mathbb{R} \times \langle 0,T))]^m$, má všechny derivace obsažené v rovnici (1.6) spojité na $\mathbb{R} \times (0,T)$ a splňuje všechny rovnice (1.6) na $\mathbb{R} \times (0,T)$ a počáteční podmínku na $\mathbb{R}$.

Definice 1.2. *Slabým řešením* úlohy (1.6) nazveme $\boldsymbol{u}(x,t)$ takové, že $\boldsymbol{u}(x,t) \in [\mathcal{L}^\infty_{loc}(\mathbb{R} \times (0,\infty))]^m$, $\boldsymbol{u}_0(x) \in [\mathcal{L}^\infty_{loc}(\mathbb{R})]^m$, $\boldsymbol{f}(\boldsymbol{u}) \in [\mathcal{C}^1(\mathbb{R})]^m$ a pro libovolné $\boldsymbol{\varphi} \in [\mathcal{C}^\infty_0(\mathbb{R} \times \langle 0,\infty))]^m$ platí následující integrální rovnost

$$\int_0^\infty \int_{-\infty}^\infty [\boldsymbol{\varphi}_t \boldsymbol{u} + \boldsymbol{\varphi}_x \boldsymbol{f}(\boldsymbol{u})]\, \mathrm{d}x\mathrm{d}t = -\int_{-\infty}^\infty \boldsymbol{\varphi}(x,0)\boldsymbol{u}_0(x)\, \mathrm{d}x. \tag{1.7}$$

Definice 1.3. Slabé řešení $\boldsymbol{u}(x,t)$ úlohy (1.6) nazveme *entropické* (často též přípustné), pokud pro libovolné $\boldsymbol{\varphi} \in [\mathcal{C}_0^1(\mathbb{R} \times (-\infty, T))]^m$, $\boldsymbol{\varphi} \geqq \mathbf{0}$ a pro každou konvexní a spojitou entropii E s entropickým tokem F, splňuje

$$\int_0^\infty \int_{-\infty}^\infty [\boldsymbol{\varphi}_t E(\boldsymbol{u}) + \boldsymbol{\varphi}_x F(\boldsymbol{u})] \,\mathrm{d}x\,\mathrm{d}t + \int_{-\infty}^\infty \boldsymbol{\varphi}(x,0) E(\boldsymbol{u}_0(x)) \,\mathrm{d}x \geqq 0. \tag{1.8}$$

Poznámka 1.4. Prostor funkcí $\mathcal{C}(X)$ značí množinu všech spojitých funkcí na odpovídající množině X. Symbol $[\mathrm{C}(X)]^m$ značí, že všech m složek vektoru leží v $\mathcal{C}(X)$. Prostor funkcí $\mathcal{C}_0^\infty$ označuje množinu všech funkcí, které jsou nekoněčněkrát spojitě diferencovatelné a mají kompaktní nosič (tj. pro dostatečně velké $|x|$ a t je $\varphi(x,t) = 0$). Prostor funkcí $\mathcal{C}_0^1$ je množina jednou spojitě diferencovatelných funkcí s kompaktním nosičem. Prostor funkcí $\mathcal{L}_{loc}^\infty$ značí množinu všech funkcí definovaných skoro všude na Ω (tj. $\mathbb{R} \times (0,\infty)$ nebo $\mathbb{R}$), které jsou esenciálně omezené na všech kompaktních podmnožinách Ω. Více o vlastnostech a definicích prostorů lze nalézt například v [8].

Definice 1.5. Soustava nelineárních parciálních diferenciálních rovnic (1.6) se nazývá *slabě hyperbolická*, pokud Jacobiho matice $\boldsymbol{f}'(\boldsymbol{u}_0)$ má reálná vlastní čísla pro jakýkoliv fyzikálně relevantní stav $\boldsymbol{u}_0 \in \mathbb{R}^m$.

Soustava nelineárních parciálních diferenciálních rovnic (1.6) se nazývá *(silně) hyperbolická*, pokud Jacobiho matice $\boldsymbol{f}'(\boldsymbol{u}_0)$ je diagonalizovatelná a má reálná vlastní čísla pro jakýkoliv fyzikálně relevantní stav $\boldsymbol{u}_0 \in \mathbb{R}^m$.

Soustava nelineárních parciálních diferenciálních rovnic (1.6) se nazývá *ryze (striktně) hyperbolická*, pokud Jacobiho matice $\boldsymbol{f}'(\boldsymbol{u}_0)$ je diagonalizovatelná a má navzájem různá reálná vlastní čísla pro jakýkoliv fyzikálně relevantní stav $\boldsymbol{u}_0 \in \mathbb{R}^m$.

Vlastní čísla Jacobiho matice $\boldsymbol{f}'(\boldsymbol{u})$ se nazývají charakteristické rychlosti, značí se $\lambda_i(\boldsymbol{u})$ a definují tzv. charakteristická pole λ_i.

Poznámka 1.6. Gradient vlastních čísel (tj. charakteristických rychlostí) je dán

$$\nabla \lambda_i(\boldsymbol{u}) = \left[\frac{\partial \lambda_i}{\partial u_1}, \ldots, \frac{\partial \lambda_i}{\partial u_m}\right]^T.$$

Vektor $\boldsymbol{r}^i(\boldsymbol{u})$ značí vlastní vektory (přesněji pravé vlastní vektory) příslušející vlastnímu číslu $\lambda_i(\boldsymbol{u})$.

Definice 1.7. Charakteristické pole λ_i se nazývá *lineárně degenerované* pokud platí

$$\nabla \lambda_i(\boldsymbol{u}) \boldsymbol{r}^i(\boldsymbol{u}) = 0, \quad \forall \boldsymbol{u} \in \mathbb{R}^m. \tag{1.9}$$

Definice 1.8. Charakteristické pole λ_i se nazývá *ryze nelineární* pokud platí

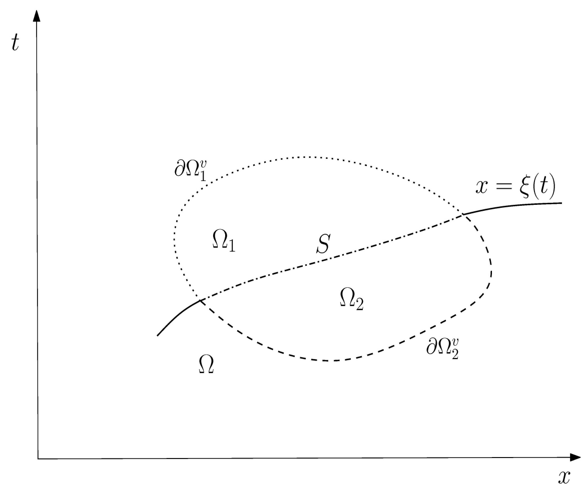
$$\nabla\lambda_i(\boldsymbol{u})\boldsymbol{r}^i(\boldsymbol{u}) \neq 0, \quad \forall \boldsymbol{u} \in \mathbb{R}^m. \tag{1.10}$$

1.2.1 Rankineův-Hugoniotův vztah

Protože slabé řešení může obsahovat nespojitosti, například tzv. rázovou vlnu, budeme se v této sekci zabývat odvozením rovnosti, která se nazývá Rankineův-Hugoniotův vztah. Tento vztah odvodíme pro následující skalární zákon zachování, ale jeho platnost lze rozšířit i na soustavy,

$$\begin{aligned} u_t + [f(u)]_x &= 0, \quad x \in \mathbb{R},\, t \in (0,T),\, T > 0, \\ u(x,0) &= u_0(x), \end{aligned} \tag{1.11}$$

Předpokládejme oblast Ω, která je rozdělena na dvě části Ω_1 a Ω_2 hladkou křivkou $x = \xi(t)$, viz Obr. 1.1. Dále předpokládejme platnost zákona zachování v diferenciálním tvaru (1.11) a dostatečně hladkou funkci u, která tuto rovnost splňuje pro $x < \xi(t)$ a $x > \xi(t)$ se skokem podél hladké křivky $x = \xi(t)$, navíc tato funkce splňuje počáteční podmínku (jde o klasické řešení ve smyslu definice 1.1), více viz [11]. Tato



Obr. 1.1 Ilustrace k odvození Rankineovy-Hugoniotovy podmínky.

funkce má pro libovolné t limity zprava a zleva, které budeme značit následujícím způsobem

$$\lim_{x\to\xi(t)^+} u(x,t) = u^+(\xi(t),t) := u_r,$$

$$\lim_{x\to\xi(t)^-} u(x,t) = u^-(\xi(t),t) := u_l,$$

analogicky

$$\lim_{x\to\xi(t)^+} f(u(x,t)) := f(u_r),$$
$$\lim_{x\to\xi(t)^-} f(u(x,t)) := f(u_l).$$

Námi předpokládaná funkce $u(x,t)$ musí v Ω_1 a v Ω_2 splňovat definici slabého řešení (viz definice 1.2), kde nebude zahrnuta počáteční podmínka, protože oblast Ω je volena mimo počátek, tj. musí platit

$$\int\limits_{\Omega_i} [\varphi_t u + \varphi_x f(u)]\,\mathrm{d}x\mathrm{d}t = 0 \quad \text{pro} \quad i = 1,2$$

pro všechny funkce φ, které nyní uvažujeme nulové na hranici $\partial\Omega = \partial\Omega_1^v \cup \partial\Omega_2^v$, tj. z podprostoru funkcí $[\mathcal{C}_0^\infty(\mathbb{R}\times\langle 0,\infty))]$. Označíme hranici oblasti Ω_1 jako $\partial\Omega_1 = \partial\Omega_1^v \cup S$ a hranici oblasti Ω_2 jako $\partial\Omega_2 = \partial\Omega_2^v \cup S$. Dále označíme složky vnějších normál k $\partial\Omega_1$ a $\partial\Omega_2$ následovně: $\boldsymbol{n}_1 = (n_1^x, n_1^t)$ a $\boldsymbol{n}_2 = (n_2^x, n_2^t)$. Použitím Greenovy věty, (1.11) v oblastech Ω_1, Ω_2 a faktu, že φ jsme speciálně zvolili nulové na hranici $\partial\Omega$, dostaneme

$$\begin{aligned}
0 &= \int\limits_{\Omega_1} [\varphi_t u + \varphi_x f(u)]\,\mathrm{d}x\mathrm{d}t + \int\limits_{\Omega_2} [\varphi_t u + \varphi_x f(u)]\,\mathrm{d}x\mathrm{d}t = \\
&= \int\limits_{\partial\Omega_1} \varphi\, u\, n_1^t\,\mathrm{d}S + \int\limits_{\partial\Omega_1} \varphi f(u)\, n_1^x\,\mathrm{d}S - \int\limits_{\Omega_1} \varphi\, u_t\,\mathrm{d}x\mathrm{d}t - \int\limits_{\Omega_1} \varphi\, f(u)_x\,\mathrm{d}x\mathrm{d}t + \\
&+ \int\limits_{\partial\Omega_2} \varphi\, u\, n_2^t\,\mathrm{d}S + \int\limits_{\partial\Omega_2} \varphi f(u)\, n_2^x\,\mathrm{d}S - \int\limits_{\Omega_2} \varphi\, u_t\,\mathrm{d}x\mathrm{d}t - \int\limits_{\Omega_2} \varphi\, f(u)_x\,\mathrm{d}x\mathrm{d}t =
\end{aligned}$$

$$\begin{aligned}
&= \int\limits_{\partial\Omega_1} (f(u)\, n_1^x + u\, n_1^t)\varphi\,\mathrm{d}S + \int\limits_{\partial\Omega_2} (f(u)\, n_2^x + u\, n_2^t)\varphi\,\mathrm{d}S - \int\limits_{\Omega_1} (u_t + f(u)_x)\,\varphi\,\mathrm{d}x\mathrm{d}t - \\
&- \int\limits_{\Omega_2} (u_t + f(u)_x)\,\varphi\,\mathrm{d}x\mathrm{d}t = \int\limits_{\partial\Omega_1} (f(u),u)\,\varphi\,\boldsymbol{n}_1\,\mathrm{d}x\mathrm{d}t + \int\limits_{\partial\Omega_2} (f(u),u)\,\varphi\,\boldsymbol{n}_2\,\mathrm{d}x\mathrm{d}t = \\
&= \int\limits_{\partial\Omega_1^v} (f(u),u)\,\varphi\,\boldsymbol{n}_1\,\mathrm{d}x\mathrm{d}t + \int\limits_{S} \lim_{x\to\xi(t)^-} (f(u),u)\,\varphi\,\boldsymbol{n}_1\,\mathrm{d}x\mathrm{d}t + \int\limits_{\partial\Omega_2^v} (f(u),u)\,\varphi\,\boldsymbol{n}_2\,\mathrm{d}x\mathrm{d}t + \\
&+ \int\limits_{S} \lim_{x\to\xi(t)^+} (f(u),u)\,\varphi\,\boldsymbol{n}_2\,\mathrm{d}x\mathrm{d}t = \int\limits_{S} (f(u_r),u_r)\,\varphi\,\boldsymbol{n}_1\,\mathrm{d}S + \int\limits_{S} (f(u_l),u_l)\,\varphi\,\boldsymbol{n}_2\,\mathrm{d}S.
\end{aligned}$$

Pro S platí, že $\boldsymbol{n} = \boldsymbol{n}_1 = -\boldsymbol{n}_2$, a tedy

$$\int\limits_{S} (f(u_r) - f(u_l), u_r - u_l)\,\varphi\,\boldsymbol{n}\,\mathrm{d}S = 0.$$

Funkce φ je libovolná, a tedy platí

$$(f(u_r) - f(u_l), u_r - u_l)\,\boldsymbol{n} = 0.$$

Pro křivku S parametrizovanou $x = \xi(t)$ je jednotkový vektor normály

$$\boldsymbol{n} = \left(\frac{1}{\sqrt{1+\dot{\xi}(t)^2}}, -\frac{\dot{\xi}(t)}{\sqrt{1+\dot{\xi}(t)^2}}\right),$$

a tedy na křivce S platí

$$f(u_r) - f(u_l) = \dot{\xi}(t)(u_r - u_l), \tag{1.12}$$

kde $\dot{\xi}(t) = s$ je rychlost šíření nespojitosti. Výše uvedená rovnost se nazývá Rankineův-Hugoniotův vztah (v této souvislosti velmi často nazývaný Rankineova-Hugoniotova podmínka nebo podmínka skoku).

Funkce $u(x,t)$ je slabé řešení, pokud je řešením v klasickém smyslu v obou regionech Ω_1 a Ω_2 a splňuje Rankineův-Hugoniotův vztah podél hladké křivky $x = \xi(t)$.

1.3 Skalární lineární rovnice

V této kapitole uvažujme nejjednodušší z modelů využívající se k popisu proudění a to počáteční úlohu pro lineární skalární rovnici

$$u_t + a u_x = 0, \quad x \in \mathbb{R},\ t \in (0,T),\ T > 0,\ a \in \mathbb{R}, \tag{1.13}$$

$$u(x,0) = u_0(x), \quad x \in \mathbb{R}. \tag{1.14}$$

Tato rovnice se nazývá též advekční. Modeluje advekci (transport) například nečistoty spolu s proudem. Koncentrace nečistoty se předpokládá dostatečně malá, tak aby její malé změny neovlivnily dynamiku proudění. Hledaná veličina $u = u(x,t)$ zde tedy může reprezentovat například hustotu (koncentraci) nečistot ve vodě a $a = $ $= konst.$ je rychlost proudu.

Řešení výše uvedené počáteční úlohy nebudeme studovat na celém časoprostorovém válci $\mathbb{R} \times (0,T)$, ale bude nás zajímat chování řešení na určitých křivkách. Tyto křivky nazýváme charakteristikami, neboli charakteristickým systémem diferenciální rovnice:

Definice 1.9. Rovnici (1.13) přiřadíme soustavu obyčejných diferenciálních rovnic (charakteristický systém) pro neznámé $x = x(s)$ a $t = t(s)$

$$\frac{\mathrm{d}x(s)}{\mathrm{d}s} = a, \quad \text{(koeficient u } u_x\text{)}, \tag{1.15}$$

$$\frac{\mathrm{d}t(s)}{\mathrm{d}s} = 1, \quad \text{(koeficient u } u_t\text{)}. \tag{1.16}$$

Každé klasické řešení $Z(s) = (x(s), t(s))$ nazveme charakteristikou.

Pro výše uvedenou advekční rovnici přidáme počáteční podmínky $x(0) = x_0$ a $t(0) = 0$ odpovídající počátečním podmínkám pro advekční rovnici (1.14) a lze již vypočítat hledané charakteristiky:

$$x(s) = as + C_1, \quad C_1 = konst. \tag{1.17}$$
$$t(s) = s + C_2, \quad C_2 = konst. \tag{1.18}$$

Dosazením počátečních podmínek určíme konstanty C_1 a C_2:

$$x(0) = C_1 = x_0, \tag{1.19}$$
$$t(0) = C_2 = 0. \tag{1.20}$$

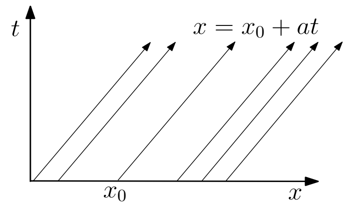
Získáváme charakteristiku $Z(s) = (as + x_0, s)$, neboli vyloučením parametru s získáme polopřímky $x = x_0 + at$. Dále lze snadno ukázat, že řešení $u(x,t)$ je podél polopřímek $Z(s)$ konstantní (využitím (1.15), (1.16) a (1.13)):

$$\frac{\mathrm{d}}{\mathrm{d}s}u(x(s),t(s)) = \frac{\partial u}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}s} + \frac{\partial u}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}s} = u_t + au_x = 0. \tag{1.21}$$

Výše uvedená advekční rovnice má tedy jednoduché řešení

$$u(x,t) = u_0(x - at). \tag{1.22}$$

Tedy podél polopřímek $x_0 + at$, viz Obr. 1.2, kde x_0 je libovolný bod na ose x, bude řešení $u(x,t)$ rovno hodnotě $u_0(x - at)$. Jinými slovy, hodnota u_0 se pouze transportuje podél polopřímek. Jednoduchým zobecněním skalární advekce je předpoklad, že



Obr. 1.2 Charakteristiky advekční rovnice pro $a > 0$.

$a = a(x)$, tedy získáváme advekční rovnici s nekonstantními koeficienty (jde o tzv. nekonzervativní tvar):

$$u_t + a(x)u_x = 0, \quad x \in \mathbb{R},\ t \in (0,T),\ T > 0, \tag{1.23}$$
$$u(x,0) = u_0(x). \tag{1.24}$$

Tento model popisuje například proudění nečistot (opět předpokládáme nízkou koncentraci) v řece, kdy v každém bodě může být rychlost proudění různá. Opět určíme charakteristiky této nové advekční rovnice řešením následující soustavy obyčejných diferenciálních rovnic se stejnými počátečními podmínkami jako v předchozím případě

$$\begin{aligned} \frac{\mathrm{d}x(s)}{\mathrm{d}s} &= a(x), \qquad (1.25)\\ \frac{\mathrm{d}t(s)}{\mathrm{d}s} &= 1. \qquad (1.26)\end{aligned}$$

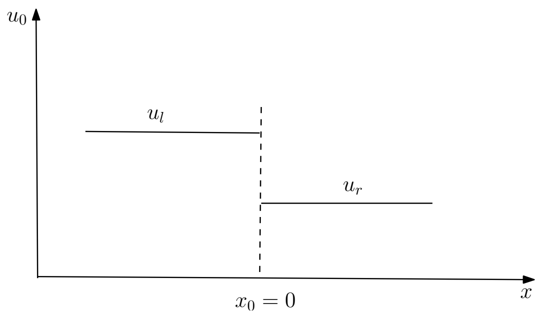
Získáme tak soustavu charakteristik $Z(s) = (x(s), s)$, které pro $a \neq konst.$ již nebudou přímkami. Podél těchto křivek však stále platí, že je řešení konstantní.

1.3.1 Riemannův problém pro skalární lineární rovnici

V této podkapitole definujeme speciální počáteční úlohu pro advekční rovnici nazývanou Riemannův problém, viz Obr. 1.3:

$$\begin{aligned} u_t + au_x &= 0, \quad x \in \mathbb{R}, t \in (0, T), T > 0, a \in \mathbb{R}, \qquad (1.27)\\ u(x,0) = u_0(x) &= \begin{cases} u_l & \text{pro} \quad x < 0, \\ u_r & \text{pro} \quad x > 0, \end{cases} \qquad (1.28)\end{aligned}$$

kde $u_l = konst.$ a $u_r = konst.$ Pokud předpokládáme nespojité počáteční podmínky, které jsou konstantní, mluvíme také o klasickém Riemannově problému. V případě, že předpokládáme po částech polynomiální počáteční podmínky, mluvíme o zobecněném Riemannově problému, viz podkapitola 1.7.2.

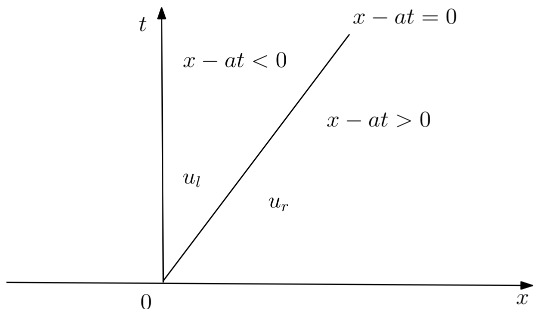


Obr. 1.3 Počáteční podmínky Riemannova problému.

Pro nespojitou počáteční podmínku získáme následující entropické řešení Riemannova problému (odvozené řešení (1.22) lze použít i pro nespojité počáteční podmínky), viz Obr. 1.4,

$$u(x,t) = u_0(x - at) = \begin{cases} u_l & \text{pro} \quad x - at < 0, \\ u_r & \text{pro} \quad x - at > 0. \end{cases} \qquad (1.29)$$

Jinými slovy, nespojitost (skok $u_r - u_l$) se šíří rychlostí a doprava nebo doleva (v závislosti na znaménku u rychlosti a) s neměnným profilem.



Obr. 1.4 Ilustrace řešení Riemannova problému pro advekční rovnici.

1.4 Soustava lineárních rovnic

V této kapitole se budeme zabývat počáteční úlohou pro soustavu lineárních rovnic

$$\begin{aligned} \mathbf{u}_t + \mathbf{A}\mathbf{u}_x &= \mathbf{0}, \quad x \in \mathbb{R},\, t \in (0,T),\, T > 0, && (1.30)\\ \mathbf{u}(x,0) &= \mathbf{u}_0(x), \quad x \in \mathbb{R}. && (1.31)\end{aligned}$$

kde $\mathbf{u} = \mathbf{u}(x,t)$ je vektor neznámých funkcí (zachovávané veličiny), soustava se předpokládá ryze hyperbolická, tj. matice $\mathbf{A} \in \mathbb{R}^{m\times m}$ je diagonalizovatelná s reálnými různými vlastními čísly, více viz například [9], nebo [1].

Jednoduchým příkladem takovéto lineární soustavy může být model lineární akustiky:

$$\begin{bmatrix} p \\ u \end{bmatrix}_t + \begin{bmatrix} 0 & K_0 \\ \frac{1}{\rho_0} & 0 \end{bmatrix} \begin{bmatrix} p \\ u \end{bmatrix}_x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{1.32}$$

kde $p = p(x,t)$ a $u = u(x,t)$ jsou perturbace tlaku a rychlosti (v akustice). Parametr K_0 reprezentuje objemový modul pružnosti a ρ_0 je hustota. Tento model popisuje například šíření zvukových vln v jednodimenzionální trubici naplněné plynem. Jde o linearizovaný (zjednodušený) model, protože obecně jsou problémy akustiky nelineárními úlohami, které obsahují takzvané rázové vlny (například tzv. sonický třesk).

Abychom mohli řešit lineární soustavu (1.30), je potřeba provést následující kroky: využijeme diagonalizovatelnosti matice $\mathbf{A}$, tedy existence regulární matice $\mathbf{R}$ (jejími sloupci jsou pravé vlastní vektory matice $\mathbf{A}$), takové, že platí

$$\mathbf{A} = \mathbf{R}\mathbf{\Lambda}\mathbf{R}^{-1}, \tag{1.33}$$

kde $\mathbf{\Lambda} = \operatorname{diag}(\lambda^1, \dots, \lambda^m)$, kde λ^p, $p = 1, \dots, m$, jsou vlastní čísla matice $\mathbf{A}$. Tedy získáme

$$\mathbf{u}_t + \mathbf{R}\mathbf{\Lambda}\mathbf{R}^{-1}\mathbf{u}_x = \mathbf{0}. \tag{1.34}$$

Dalším krokem je vynásobení soustavy maticí $\mathbf{R}^{-1}$ zleva:

$$\mathbf{R}^{-1}\mathbf{u}_t + \mathbf{R}^{-1}\mathbf{R}\mathbf{\Lambda}\mathbf{R}^{-1}\mathbf{u}_x = \mathbf{0}. \tag{1.35}$$

Posledním krokem je zavedení nové neznámé a transformace počáteční podmínky

$$\mathbf{v}(x,t) = \mathbf{R}^{-1}\mathbf{u}(x,t), \tag{1.36}$$
$$\mathbf{v}(x,0) = \mathbf{R}^{-1}\mathbf{u}(x,0), \tag{1.37}$$

tedy získáme následující počáteční úlohu

$$\mathbf{v}_t + \mathbf{\Lambda}\mathbf{v}_x = \mathbf{0}, \quad x \in \mathbb{R},\, t \in (0,T),\, T > 0, \tag{1.38}$$
$$\mathbf{v}(x,0) = \mathbf{v}_0(x), \quad x \in \mathbb{R}. \tag{1.39}$$

Protože matice $\mathbf{\Lambda}$ je diagonální, lze počáteční úlohu (1.38), (1.39) rozdělit na m nezávislých lineárních advekčních počátečních úloh, tedy

$$v_t^p + \lambda_p v_x^p = 0, \quad x \in \mathbb{R},\, t \in (0,T),\, T > 0,\, p = 1,\dots,m, \tag{1.40}$$
$$v^p(x,0) = v_0^p(x), \quad x \in \mathbb{R}. \tag{1.41}$$

Protože jsme předpokládali, že matice $\mathbf{A}$ je ryze hyperbolická, tedy vlastní čísla jsou reálná, mají tyto rovnice stejný fyzikální smysl, jako v kapitole 1.3. Jak bylo ukázáno, informace (počáteční podmínka) se bude šířit po charakteristických křivkách (polopřímkách) $x(s) = x_0 + \lambda^p t$ rychlostí (charakteristickou rychlostí) λ^p, neboli řešení advekčních rovnic je následující (viz (1.22))

$$v^p(x,t) = v_0^p(x - \lambda^p t), \quad p = 1,\dots,m. \tag{1.42}$$

Návratem k původní proměnné $\mathbf{u}$, viz (1.36), získáme následující

$$\mathbf{u}(x,t) = \mathbf{R}\mathbf{v}(x,t) = \sum_{p=1}^{m} v^p(x,t)\mathbf{r}^p = \sum_{p=1}^{m} v_0^p(x - \lambda^p t)\mathbf{r}^p, \tag{1.43}$$

kde $\mathbf{r}^p$, $p = 1,\dots,m$, jsou pravé vlastní vektory matice $\mathbf{A}$ (sloupce matice $\mathbf{R}$). Řešení je tedy lineární kombinace pravých vlastních vektorů matice $\mathbf{A}$, neboli je superpozicí vln (šířená informace se často nazývá vlna), které se šíří rozdílnými rychlostmi λ^p a svůj tvar, tj. $v_0^p(x - \lambda^p t)\mathbf{r}^p$, nemění. Pokud zavedeme levé vlastní vektory $\mathbf{l}^p$, tedy sloupce matice $\mathbf{R}^{-1}$, získáme řešení závislé přímo na počáteční podmínce $\mathbf{u}_0$, tedy

$$\mathbf{u}(x,t) = \sum_{p=1}^{m} \mathbf{l}^p\mathbf{u}_0(x - \lambda^p t)\mathbf{r}^p. \tag{1.44}$$

K určení řešení soustavy lineárních rovnic je tedy potřeba znát vlastní strukturu matice $\mathbf{A}$, tj. vlastní čísla a příslušné levé a pravé vlastní vektory.

1.4.1 Oblast závislosti

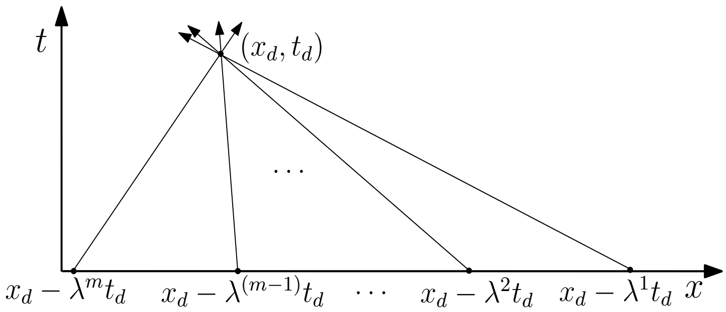
V této podkapitole rozebereme pojem oblast (též obor) závislosti. Ve stručnosti byl popsán v [1]. Předpokládejme soustavu rovnic s vlastními čísly λ^p, $p = 1,\dots,m$.

Hledáme oblast závislosti v nějakém bodě (x_d, t_d), v tomto bodě je řešení $\mathbf{u}(x_d, t_d)$ ovlivněno počáteční podmínkou pouze v bodech $x_d - \lambda^p t_d$, viz (1.44). Definujeme tedy oblast závislosti následovně:

Definice 1.10 (Oblast závislosti). Oblastí (oborem) závislosti bodu (x_d, t_d) hyperbolické soustavy rovnic (1.30) nazveme následující množinu

$$D(x_d, t_d) = \{x_d - \lambda^p t_d,\, p = 1, \ldots, m\}, \tag{1.45}$$

kde λ^p, $p = 1, \ldots, m$, jsou vlastní čísla matice $\mathbf{A}$.



Obr. 1.5 Oblast závislosti.

Situaci ilustruje Obr. 1.5, řešení v bodě (x_d, t_d) ovlivňuje počáteční podmínka v m vyznačených bodech. Pro hyperbolické rovnice je oblast závislosti vždy omezená množina (viz poznámka 1.12), protože se informace šíří vždy konečnou rychlostí (na rozdíl od parabolických rovnic), viz model advekční rovnice, kapitola 1.3. Oblast závislosti úzce souvisí s numerickou oblastí závislosti, tzv. CFL (Courant, Friedrichs, Levy) podmínkou. Více viz například [9].

1.4.2 Riemannův problém pro soustavu lineárních rovnic

V této podkapitole opět definujeme speciální počáteční úlohu pro lineární soustavy s po částech konstantní počáteční podmínkou (klasický Riemannův problém)

$$\mathbf{u}_t + \mathbf{A}\mathbf{u}_x = 0, \quad x \in \mathbb{R},\, t \in (0, T),\, T > 0,\, A \in \mathbb{R}^{m \times m}, \tag{1.46}$$

$$\mathbf{u}(x, 0) = \mathbf{u}_0(x) = \begin{cases} \mathbf{u}_l & \text{pro} \quad x < 0, \\ \mathbf{u}_r & \text{pro} \quad x > 0, \end{cases} \tag{1.47}$$

kde $\mathbf{u}_l \in \mathbb{R}^m$ a $\mathbf{u}_r \in \mathbb{R}^m$ (vektory konstant).

Jak již bylo řečeno v podkapitole 1.3.1, je řešení klasického Riemannova problému pro jednotlivé advekční rovnice následující

$$v^p(x, t) = v_0^p(x - \lambda^p t) = \begin{cases} v_l^p & \text{pro} \quad x - \lambda^p t < 0, \\ v_r^p & \text{pro} \quad x - \lambda^p t > 0 \end{cases} \tag{1.48}$$

a pro počáteční podmínky platí

$$\mathbf{u}_l = \mathbf{R}\mathbf{v}_l = \sum_{p=1}^{m} v_l^p \mathbf{r}^p, \tag{1.49}$$

$$\mathbf{u}_r = \mathbf{R}\mathbf{v}_r = \sum_{p=1}^{m} v_r^p \mathbf{r}^p. \tag{1.50}$$

Tedy hledané řešení $\mathbf{u}(x,t)$ má následující tvar

$$\begin{aligned} \mathbf{u}(x,t) &= \sum_{p:\lambda^p<x/t} v_r^p \mathbf{r}^p + \sum_{p:\lambda^p>x/t} v_l^p \mathbf{r}^p = &(1.51)\\ &= \sum_{p:\lambda^p<x/t} \mathbf{l}^p \mathbf{u}_r \mathbf{r}^p + \sum_{p:\lambda^p>x/t} \mathbf{l}^p \mathbf{u}_l \mathbf{r}^p. &(1.52)\end{aligned}$$

Rankineova-Hugoniotova podmínka, viz podkapitola 1.2.1, pro soustavu lineárních rovnic je následující

$$\mathbf{A}(\mathbf{u}_r - \mathbf{u}_l) = \lambda(\mathbf{u}_r - \mathbf{u}_l), \tag{1.53}$$

tedy skok ve vektoru $\mathbf{u}_r - \mathbf{u}_l$ je vlastním vektorem matice $\mathbf{A}$ a λ, tj. λ^p, $p = 1, \ldots, m$, jsou rychlosti šíření nespojitostí.

Jak bylo vysvětleno v podkapitole 1.3.1, jednotlivé nespojitosti (skoky) $u_r^p - u_l^p$ se šíří po odpovídající p-té charakteristice rychlostí λ^p doleva nebo doprava (v závislosti na znaménku p-tého vlastního čísla). Lze tedy hledané řešení napsat v následujícím tvaru, viz (1.44),

$$\begin{aligned} \mathbf{u}(x,t) &= \mathbf{u}_l + \sum_{p:\lambda^p<x/t} \mathbf{l}^p(\mathbf{u}_r - \mathbf{u}_l)\mathbf{r}^p = &(1.54)\\ &= \mathbf{u}_r - \sum_{p:\lambda^p>x/t} \mathbf{l}^p(\mathbf{u}_r - \mathbf{u}_l)\mathbf{r}^p. &(1.55)\end{aligned}$$

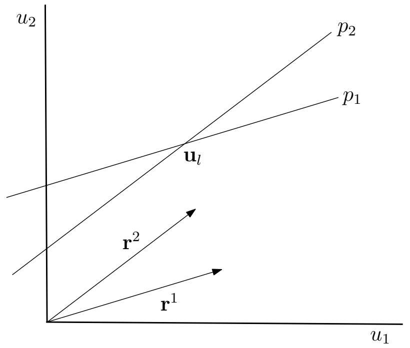
1.5 Hugoniotovy množiny

V této kapitole rozebereme pojem Hugoniotových množin, podrobněji lze toto téma nalézt například v [9], nebo v [22]. Předpokládejme soustavu lineárních rovnic pro dvě neznámé, viz například model (1.32). Zopakujme Riemannův problém pro soustavu lineárních rovnic, podrobněji viz podkapitola 1.4.2,

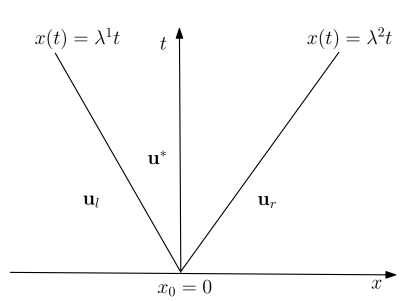
$$\begin{aligned} \mathbf{u}_t + \mathbf{A}\mathbf{u}_x &= 0, \quad x \in \mathbb{R},\ t \in (0,T),\ T > 0,\ A \in \mathbb{R}^{m\times m}, &(1.56)\\ \mathbf{u}(x,0) = \mathbf{u}_0(x) &= \begin{cases} \mathbf{u}_l & \text{pro} \quad x < 0, \\ \mathbf{u}_r & \text{pro} \quad x > 0, \end{cases} &(1.57)\end{aligned}$$

kde $\mathbf{u}_l \in \mathbb{R}^m$ a $\mathbf{u}_r \in \mathbb{R}^m$. Připomeňme, že soustavu předpokládáme ryze hyperbolickou s vlastními čísly λ^1, λ^2 a odpovídajícími vlastními vektory $\mathbf{r}^1$, $\mathbf{r}^2$. Možné řešení výše uvedeného Riemannova problému pro soustavu lineárních rovnic je na Obr. 1.7.

Fázovou rovinou nazveme rovinu závislosti u_1 a u_2, často značeno (u_1, u_2)–rovina. Každý vektor $\mathbf{u}(x,t) = [u_1, u_2]^T$ je v této rovině reprezentován bodem. Stavy $\mathbf{u}_r$ a $\mathbf{u}_l$ jsou tedy v (u_1, u_2)–rovině reprezentovány body. Jak bylo již řečeno v 1.4.2 skok $\mathbf{u}_r - \mathbf{u}_l$ se šíří jako nespojitost pouze pokud je vlastním vektorem matice $\mathbf{A}$. Tedy úsečka $\overline{\mathbf{u}_r\mathbf{u}_l}$ musí být rovnoběžná s jedním z vlastních vektorů $\mathbf{r}^1$ nebo $\mathbf{r}^2$.



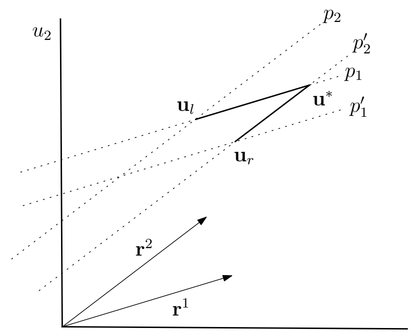
Obr. 1.6 Ilustrace k Hugoniotovým množinám.



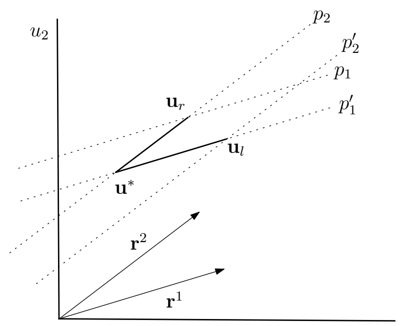
Obr. 1.7 Ilustrace řešení Riemannova problému.

Situaci ilustruje Obr. 1.6, skok mezi $\mathbf{u}_r - \mathbf{u}_l$ se může šířit jen pokud stav $\mathbf{u}_r$ bude ležet na jedné z vyznačených přímek p_1 nebo p_2, kdy platí již zmíněná rovnoběžnost s vlastními vektory, tj. $\mathbf{r}^1 \| \, p_1$ a $\mathbf{r}^2 \| \, p_2$. Tyto přímky jsou množinami všech bodů, které mohou spojovat stav $\mathbf{u}_l$ 1. nebo 2. vlnou (nespojitostí, informací). Tuto množinu bodů nazýváme Hugoniotovou množinou.

Uvažujme dvě možnosti počátečních podmínek Riemannova problému pro lineární soustavy a jeho řešení $\mathbf{u}^*$, který je spojen se stavem $\mathbf{u}_l$ první vlnou (nespojitostí) a zároveň se stavem $\mathbf{u}_r$ druhou vlnou, viz obrázky 1.8 a 1.9.



Obr. 1.8 Řešení Riemannova problému ve fázové rovině 1.



Obr. 1.9 Řešení Riemannova problému ve fázové rovině 2.

Podrobněji vysvětlíme situaci na Obr. 1.8. Tedy stav (ve fázové rovině bod) $\mathbf{u}_l$ je spojen vlnou se stavem (bodem) $\mathbf{u}^*$, protože leží na úsečce p_1 rovnoběžnou s $\mathbf{r}^1$, procházející stavem (bodem) $\mathbf{u}_l$. Stejně tak je stav (bod) $\mathbf{q}_r$ spojen vlnou se stavem (bodem) $\mathbf{q}^*$, protože leží na úsečce p_2' rovnoběžnou s $\mathbf{r}^2$ procházející stavem (bodem) $\mathbf{u}_r$. Analogické vysvětlení lze provést u Obr. 1.9.

1.6 Skalární nelineární rovnice

Věnujme se nyní následující počáteční úloze pro skalární nelineární rovnici

$$u_t + [f(u)]_x = 0, \quad x \in \mathbb{R}, t \in (0, T), T > 0, \tag{1.58}$$

$$u(x, 0) = u_0(x), \quad x \in \mathbb{R}. \tag{1.59}$$

Budeme předpokládat, že funkce $f(u)$ je konvexní, tedy $f''(u)$ je vždy kladné. Upozorněme, že pokud $f(u) = au$, kde $a = konst.$ získáváme lineární skalární rovnici, viz kapitola 1.3. Nelineární rovnice lze přepsat do tzv. kvazilineárního tvaru, tj.

$$u_t + f'(u)u_x = 0, \tag{1.60}$$

kde $f'(u) = \frac{\mathrm{d}f(u)}{\mathrm{d}u}$. Lze tedy tuto rovnici chápat jako advekční rovnici (1.13) s proměnou rychlostí proudění $f'(u)$. Charakteristiky budou, dle definice 1.9, řešení následující soustavy obyčejných diferenciálních rovnic

$$\frac{\mathrm{d}x(s)}{\mathrm{d}s} = f'(u), \tag{1.61}$$

$$\frac{\mathrm{d}t(s)}{\mathrm{d}s} = 1, \tag{1.62}$$

s počátečními podmínkami $x(0) = x_0$ a $t(0) = 0$. Řešením jsou tedy křivky o rovnici $x = x_0 + f'(u)t$. Na těchto křivkách je řešení konstantní, neboť, stejně jako v kapitole 1.6, platí:

$$\frac{\mathrm{d}}{\mathrm{d}s}u(x(s), t(s)) = \frac{\partial u}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}s} + \frac{\partial u}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}s} = u_t + f'(u)u_x = 0. \tag{1.63}$$

Protože řešení je na křivce konstantní, tj. $u = konst.$, musí platit, že i $f'(u) = konst.$ a křivky jsou tedy opět přímkami. Na rozdíl od lineární skalární rovnice, kdy se charakteristiky nikdy neprotly, může být struktura charakteristik u nelineární skalární rovnice velmi různá. Tato struktura umožňuje vzniknout rázovým vlnám, které jsou speciálním případem řešení Riemannova problému, viz podkapitola 1.6.1.

O existenci, jednoznačnosti a vlastnostech řešení nelineární skalární rovnice pojednává následující fundamentální věta a poznámka (i s důkazy k nalezení např. v [4]):

Věta 1.11 (Kružkov). *Skalární počáteční úloha (1.58), (1.59), kde* $f \in \mathcal{C}^1(\mathbb{R})$ *a* $u_0 \in \mathcal{L}^\infty(\mathbb{R})$ *má jednoznačné entropické řešení* $u(x,t) \in \mathcal{L}^\infty(\mathbb{R} \times \mathbb{R}^+)$.

Poznámka 1.12 (Vlastnosti řešení). Lze ukázat, že entropické řešení z předchozí věty splňuje následující vlastnosti (viz [4]):

1. stabilita: pro skoro všechna $t \in \mathbb{R}^+$ platí

$$\|u(x,t)\|_{\mathcal{L}^\infty(\mathbb{R}\times\mathbb{R}^+)} \leqq \|u_0\|_{\mathcal{L}^\infty(\mathbb{R})},$$

2. monotónnost řešení: pokud pro dvě počáteční podmínky platí: $u_0(x) \leqq r_0(x)$ pro skoro všechna $x \in \mathbb{R}$, potom pro odpovídající řešení platí: $u(x,t) \leqq r(x,t)$ pro skoro všechna $x \in \mathbb{R}$ a skoro všechna $t \in \mathbb{R}^+$,

3. klesající totální variace (TVD vlastnost): pokud $u_0 \in \mathrm{BV}(\mathbb{R})$, potom $u(\cdot,t) \in \mathrm{BV}(\mathbb{R})$ a

$$\mathrm{TV}(u(\cdot,t)) \leqq \mathrm{TV}(u_0(\cdot)),$$

4. konzervativita: pokud $u_0 \in \mathcal{L}^1(\mathbb{R})$, potom

$$\int_{\mathbb{R}} u(x,t)\,\mathrm{d}x = \int_{\mathbb{R}} u_0(x)\,\mathrm{d}x$$

pro skoro všechna $t \in \mathbb{R}^+$,

5. konečná oblast závislosti: pokud $u(x,t)$ a $r(x,t)$ jsou dvě entropická řešení odpovídající počátečním podmínkám $u_0(x), r_0(x) \in \mathcal{L}^\infty$ a $M = \max\left\{|f'(s)| : |s| \leqq \max\left(\|u_0(x)\|_{\mathcal{L}^\infty(\mathbb{R})}, \|r_0(x)\|_{\mathcal{L}^\infty(\mathbb{R})}\right)\right\}$, potom platí

$$\int_a^b |r(x,t) - u(x,t)|\,\mathrm{d}x \leqq \int_{a-Mt}^{b-Mt} |r_0(x) - u_0(x)|\,\mathrm{d}x,$$

jednoduše řečeno: pokud $u_0(x)$ a $r_0(x)$ bude ležet v $\{x : |x - x_0| < d\}$, potom $u(x,t)$ a $r(x,t)$ bude ležet v trojúhelníku $\{(x,t) : |x - x_0| + Mt < d\}$, viz podkapitola 1.4.1.

Poznámka 1.13. Prostor $\mathrm{BV}(\mathbb{R})$ značí prostor funkcí, které mají omezenou totální variaci na $\mathbb{R}$, symbol $\mathrm{TV}(\cdot)$ označuje totální variaci nějaké funkce, která je definována následovně: pro libovolnou funkci $u(x)$ je $\mathrm{TV}(u) = \sup \sum_{j=1}^{N} |u(\xi_j) - u(\xi_{j-1})|$, kde supremum je přes všechna dělení $-\infty = \xi_0 < \xi_1 < \cdots < \xi_N = \infty$. Definice a vlastnosti prostorů $\mathcal{L}^1$, $\mathcal{L}^\infty$ a $\mathcal{C}^1$ lze nalézt v poznámce 1.4, nebo v [8]. Výše uvedená věta nám kromě jednoznačnosti řešení také uvádí mnoho vlastností tohoto řešení. Je snaha, aby numerické řešení nějakým vhodným způsobem kopírovalo tyto vlastnosti.

1.6.1 Riemannův problém pro skalární nelineární rovnici

Definujme klasický Riemannův problém pro skalární nelineární rovnici

$$u_t + [f(u)]_x = 0, \quad x \in \mathbb{R}, t \in (0,T),\ T > 0, \tag{1.64}$$

s nespojitými počátečními podmínkami

$$u(x,0) = u_0(x) = \begin{cases} u_l & \text{pro} \quad x < 0, \\ u_r & \text{pro} \quad x > 0, \end{cases} \tag{1.65}$$

kde u_l a u_r jsou konstanty.

Jako motivační příklad zde použijeme nevazkou Burgersovu rovnici, viz (1.4), tj. řešíme následující úlohu

$$u_t + \left(\frac{1}{2}u^2\right)_x = 0, \quad x \in \mathbb{R},\ t \in (0,T),\ T > 0 \tag{1.66}$$

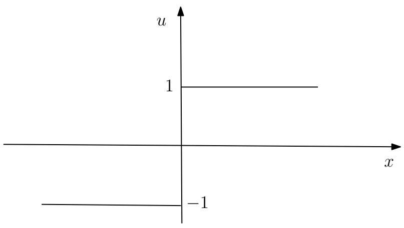
s počátečními podmínkami

$$u(x,0) = u_0(x) = \begin{cases} -1 & \text{pro} \quad x < 0, \\ 1 & \text{pro} \quad x \geqq 0. \end{cases} \tag{1.67}$$

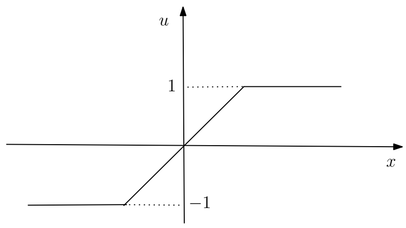
Lze ukázat, že funkce $u_1(x,t) = u_0(x)$ je slabým řešením této rovnice, viz Obr. 1.10. Také lze ukázat, že i funkce

$$u_2(x,t) = \begin{cases} -1 & \text{pro} \quad x < -t, \\ \frac{x}{t} & \text{pro} \quad -t \leqq x \leqq t, \\ 1 & \text{pro} \quad x > t \end{cases} \tag{1.68}$$

je slabým řešením, viz Obr. 1.11. Ovšem pouze řešení $u_2(x,t)$ je entropické, tedy fyzikálně správné.



Obr. 1.10 Řešení Burgersovy rovnice u_1.



Obr. 1.11 Řešení Burgersovy rovnice u_2.

Pokud $u(x,t)$ je řešením Riemannova problému, potom i $u(\alpha x, \alpha t)$, pro libovolné $\alpha > 0$, bude řešením, protože rovnice (1.64) i počáteční podmínky (1.65) jsou invariantní vzhledem ke změně souřadnic $x \to \alpha x$ a $t \to \alpha t$. Velmi stručně načrtněme myšlenku důkazu: pokud $u_0(x)$ je počáteční podmínka a $u(x,t)$ je odpovídající entropické řešení, potom pro počáteční podmínku $u_0(\alpha x)$ (což je stejná počáteční podmínka jako $u_0(x)$) existuje entropické řešení $u(\alpha x, \alpha t)$. Podle věty 1.11 je toto řešení jednoznačné, a tedy musí platit $u(\alpha x, \alpha t) \equiv u(x,t)$. Pokud zvolíme $\alpha = 1/t$, získáváme $u(x,t) = u(x/t, 1) = w(x/t)$. V případě Riemannova problému se tedy často mluví o tzv. podobnostním řešení („self-similar solution“) a počáteční úloha (1.64) a (1.65) se redukuje na hledání funkce $w(\eta) \in \mathcal{L}^\infty$, kde jsme označili $\eta = x/t$, pro kterou platí následující rovnice ve smyslu distribucí

$$-\eta w'_\eta(\eta) + [f(w(\eta))]'_\eta = 0, \quad \eta \in \mathbb{R} \tag{1.69}$$

s počátečními podmínkami

$$\begin{aligned} w(-\infty) &= u_l, \\ w(\infty) &= u_r. \end{aligned} \tag{1.70}$$

Více lze nalézt v [4].

Budeme předpokládat, že řešení Riemannova problému (pokud existuje) lze zkonstruovat kombinací z následujících tří elementárních vln (řešení):

1. Konstantní stav, tj. $u(x,t) = konst$. Jde o klasické řešení.
2. Rázová vlna.
3. Vlna zředění.

Viz například [6], [5], [4] a další.

Rázová vlna

Rázová vlna je řešení Riemannova problému (1.64), (1.65) ve tvaru

$$u(x,t) = \begin{cases} u_l & \text{pro} \quad x < st, \\ u_r & \text{pro} \quad x > st, \end{cases} \tag{1.71}$$

kdy musí platit Rankineova-Hugoniotova podmínka

$$f(u_r) - f(u_l) = s(u_r - u_l), \tag{1.72}$$

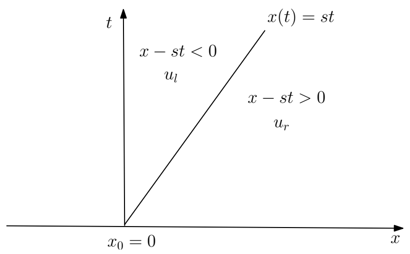
kde s je rychlost šíření nespojitosti, a navíc podmínka entropie $f'(u_l) > s > f'(u_r)$, viz Obr. 1.12. Odpovídající struktura charakteristik je na Obr. 1.13.

Vlna zředění

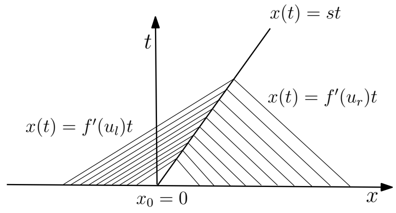
Vlna zředění, viz Obr. 1.14, je spojité řešení Riemannova problému (1.64) a (1.65), resp. (1.69) a (1.70). Platí (budeme používat značení $u(x,t) = u(\eta) = w(\eta)$, protože, jak bylo naznačeno v podkapitole 1.6, jde o jedno a totéž řešení, a proto není nutné zavádět jiné symboly)

$$[-\eta + f'(u(\eta))]u'_\eta(\eta) = 0. \tag{1.73}$$

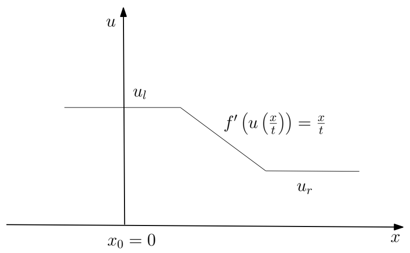
Pro $u'_\eta(\eta) = 0$ je řešení konstantní, a tedy klasické. Pro $u'_\eta(\eta) \neq 0$ bude mít rovnice $\eta = f'(u(\eta))$ jednoznačné řešení podle věty o řešitelnosti rovnic. Struktura charakteristik v případě vlny zředění je na Obr. 1.15, kdy platí $f'(u_l) < f'(u_r)$.



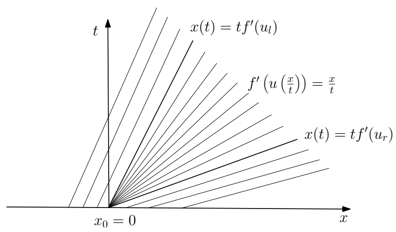
Obr. 1.12 Řešení Riemannova problému – rázová vlna.



Obr. 1.13 Struktura charakteristik – rázová vlna.



Obr. 1.14 Řešení Riemannova problému – vlna zředění.



Obr. 1.15 Struktura charakteristik – vlna zředění.

1.7 Soustavy nelineárních rovnic

Uvažujme nyní soustavu nelineárních hyperbolických rovnic

$$\mathbf{u}_t + [\mathbf{f}(\mathbf{u})]_x = \mathbf{0}, \quad x \in \mathbb{R},\, t \in (0,T),\, T > 0, \tag{1.74}$$

$$\mathbf{u}(x,0) = \mathbf{u}_0(x), \quad x \in \mathbb{R}. \tag{1.75}$$

Soustavu (1.74) přepíšeme do kvazilineárního tvaru

$$\mathbf{u}_t + \mathbf{f}'(\mathbf{u})\mathbf{u}_x = \mathbf{0}, \tag{1.76}$$

kde $\mathbf{f}'(\mathbf{u})$ je Jacobiho matice funkce $\mathbf{f} = \mathbf{f}(\mathbf{u})$. Soustava (1.74) se, stejně jako v případě lineárních soustav, předpokládá ryze hyperbolická. Vlastní čísla Jacobiho matice $\lambda^p(\mathbf{u})$, $p = 1, \ldots, m$, se opět nazývají charakteristické rychlosti šíření vln (nespojitostí, informací).

1.7.1 Riemannův problém pro soustavy nelineárních rovnic

Stejně jako v předchozích případech definujme klasický Riemannův problém pro soustavu nelineárních rovnic

$$\mathbf{u}_t + [\mathbf{f}(\mathbf{u})]_x = \mathbf{0}, \quad x \in \mathbb{R},\, t \in (0,T),\, T > 0 \tag{1.77}$$

s nespojitými počátečními podmínkami

$$\mathbf{u}(x,0) = \mathbf{u}_0(x) = \left\{ \begin{array}{lll} \mathbf{u}_l & \text{pro} & x < 0, \\ \mathbf{u}_r & \text{pro} & x > 0, \end{array} \right. \tag{1.78}$$

kde $\mathbf{u}_l \in \mathbb{R}^m$ a $\mathbf{u}_r \in \mathbb{R}^m$ jsou vektory konstant.

Stejně jako ve skalárním případě platí, že pokud $\mathbf{u}$ je slabým řešením (1.77) a (1.78), potom i $\mathbf{u}(\alpha x, \alpha t)$, pro libovolné $\alpha > 0$, bude řešením. Riemannův problém lze přepsat pomocí nové proměnné $\eta = \frac{x}{t}$ na soustavu obyčejných diferenciálních rovnic

$$\begin{aligned} -\eta \mathbf{u}'_\eta(\eta) + \mathbf{f}'(\mathbf{u}(\eta))\mathbf{u}'_\eta &= \mathbf{0}, \quad \eta \in \mathbb{R}, & (1.79)\\ \mathbf{u}(-\infty) &= \mathbf{u}_l, & (1.80)\\ \mathbf{u}(\infty) &= \mathbf{u}_r. & (1.81) \end{aligned}$$

Existenci řešení výše uvedeného klasického Riemannova problému pro nelineární soustavy řeší následující věta, která je i s důkazem k nalezení například v [4], nebo v [21]:

Věta 1.14 (Lax). *Za předpokladu, že $|\mathbf{u}_r - \mathbf{u}_l|$ je dostatečně malé, bude řešení Riemannova problému ve tvaru $\mathbf{u} = \mathbf{u}(\eta)$, $\eta = \frac{x}{t}$. Toto řešení se bude skládat z m vln oddělujících $m+1$ stavů $\mathbf{u}_l = \mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_m = \mathbf{u}_r$. Pro libovolné $k = 1, \ldots, m$ je k-tá vlna buď rázová vlna, vlna zředění, popř. kontaktní nespojitost, pokud odpovídající charakteristické pole je ryze nelineární, popř. lineárně degenerované. Pokud amplituda všech m vln je dostatečně malá, potom řešení je jednoznačné.*

Výše uvedené poznatky lze shrnout do následujících tří případů:

1. Dva konstantní stavy $\mathbf{u}_r$ a $\mathbf{u}_l$ jsou spojeny rázovou vlnou, viz Obr. 1.16, v ryze nelineárním charakteristickém poli λ^i a platí Rankineův-Hugoniotův vztah

$$\mathbf{f}(\mathbf{u}_r) - \mathbf{f}(\mathbf{u}_l) = s^i(\mathbf{u}_r - \mathbf{u}_l) \tag{1.82}$$

   a podmínka entropie

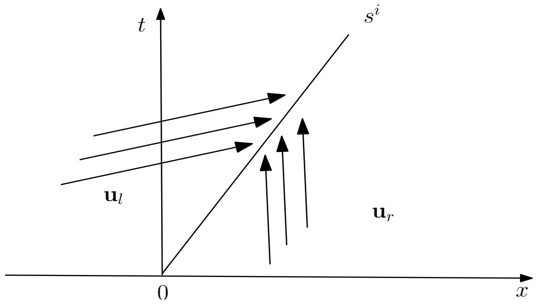
$$\lambda^i(\mathbf{u}_l) > s^i > \lambda^i(\mathbf{u}_r). \tag{1.83}$$

2. Dva konstantní stavy $\mathbf{u}_r$ a $\mathbf{u}_l$ jsou spojeny kontaktní nespojitostí , viz Obr. 1.17, v lineárně degenerovaném charakteristickém poli λ^i a platí Rankineův-Hugoniotův vztah

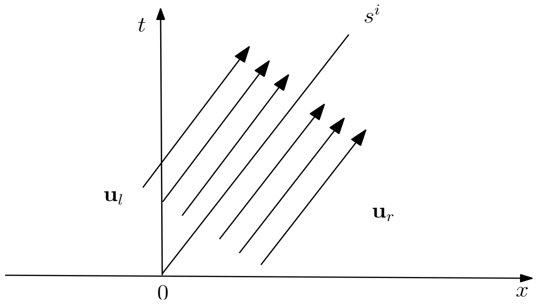
$$\mathbf{f}(\mathbf{u}_r) - \mathbf{f}(\mathbf{u}_l) = s^i(\mathbf{u}_r - \mathbf{u}_l) \tag{1.84}$$

   a podmínka

$$\lambda^i(\mathbf{u}_l) = s^i = \lambda^i(\mathbf{u}_r). \tag{1.85}$$



Obr. 1.16 Řešení Riemannova problému – rázová vlna.



Obr. 1.17 Řešení Riemannova problému – kontaktní nespojitost.

3. Dva konstantní stavy $\mathbf{u}_r$ a $\mathbf{u}_l$ jsou spojeny vlnou zředění, viz Obr. 1.18, v ryze nelineárním charakteristickém poli λ^i a platí

$$\lambda^i(\mathbf{u}_l) < \lambda^i(\mathbf{u}_r). \tag{1.86}$$

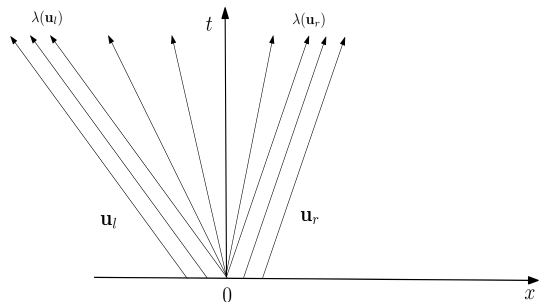
Příklad 1.15. Určete řešení Riemannova problému pro Saint-Venantovy rovnice s neměnným obdélníkovým průřezem a rovným dnem, tj. (1.1) s nulovou pravou stranou, tedy pro připomenutí jde o následující model:

$$\begin{aligned} h_t + q_x &= 0, \\ q_t + \left(\frac{q^2}{h} + \frac{1}{2}gh^2\right)_x &= 0 \end{aligned} \tag{1.87}$$

s nespojitými počátečními podmínkami

$$\mathbf{u}(x,0) = \begin{cases} \begin{bmatrix} h_l \\ q_l \end{bmatrix} & \text{pro} \quad x < 0, \\ \begin{bmatrix} h_r \\ q_r \end{bmatrix} & \text{pro} \quad x > 0, \end{cases} \tag{1.88}$$

kde h_l, h_r, q_l, $q_r \in \mathbb{R}$ a $q = hv$.



Obr. 1.18 Řešení Riemannova problému – vlna zředění.

Řešení. Velmi podrobně lze řešení uvedené úlohy nalézt v [9]. Pro řešení Riemannova problému pro (1.87) budeme potřebovat vlastní čísla a vlastní vektory Jacobiho matice soustavy. Uvedeme zde pouze výsledky, protože podrobný výpočet lze nalézt v [1] nebo v [9]. Tedy Jacobiho matice soustavy je

$$\boldsymbol{f}'(\boldsymbol{u}) = \begin{bmatrix} 0 & 1 \\ -\frac{q^2}{h^2} + gh & 2\frac{q}{h} \end{bmatrix} \tag{1.89}$$

s vlastními čísly

$$\lambda^{1,2} = \frac{q}{h} \mp \sqrt{gh} \tag{1.90}$$

a odpovídajícími vlastními vektory

$$\boldsymbol{r}^1 = \begin{bmatrix} 1 \\ \frac{q}{h} - \sqrt{gh} \end{bmatrix}, \qquad \boldsymbol{r}^2 = \begin{bmatrix} 1 \\ \frac{q}{h} + \sqrt{gh} \end{bmatrix}. \tag{1.91}$$

Dále připomeňme, že charakteristická pole definovaná vlastními čísly $\lambda^{1,2}$ jsou obě ryze nelineární, odvození opět v [1] nebo v [9].

Jako první rozebereme rázovou vlnu, tedy nespojité řešení Riemannova problému. V úvodu stručně načrtněme způsob řešení: víme že rázová vlna je takové řešení Riemannova problému, které musí splňovat Rankineovu-Hugoniotovu podmínku a také podmínku entropie. Navíc budeme hledat řešení v takovém tvaru, abychom ho mohli jednoduše zobrazit ve fázové rovině. Tedy podrobněji: hledáme všechny stavy $\mathbf{u}$, které mohou být spojeny s pevně zvoleným stavem $\mathbf{u}_c$, reprezentující $\mathbf{u}_l$ nebo $\mathbf{u}_r$. Jak bylo již řečeno, podél rázové vlny musí být splněna Rankineova-Hugoniotova podmínka, tedy

$$s(\mathbf{u}_c - \mathbf{u}) = \mathbf{f}(\mathbf{u}_c) - \mathbf{f}(\mathbf{u}), \tag{1.92}$$

kde s je rychlost šíření rázové vlny. Pro uvedený model proudění obdélníkovým korytem je Rankineova-Hugoniotova podmínka následující

$$\begin{aligned} s(h_c - h) &= q_c - q, \\ s(q_c - q) &= \frac{q_c^2}{h_c} - \frac{q^2}{h} + \frac{g}{2}\left(h_c^2 - h^2\right). \end{aligned} \tag{1.93}$$

Získáváme dvě rovnice pro tři neznámé h, q a s. Nejjednodušší způsob je získat řešení závislé na jednom parametru, vyberme jako parametr h a eliminujme rychlost s. Důvodem je, že když získáme závislost $q(h)$, lze ji pak ihned zobrazit ve fázové rovině, a my tak získáme rychlým způsobem Hugoniotovy množiny.

Z první rovnice soustavy (1.93) určíme rychlost šíření rázové vlny

$$s = \frac{q_c - q}{h_c - h} \tag{1.94}$$

a dosazením do druhé rovnice soustavy (1.93) získáme

$$\frac{(q_c - q)^2}{h_c - h} = \frac{q_c^2}{h_c} - \frac{q^2}{h} + \frac{g}{2}\left(h_c^2 - h^2\right). \tag{1.95}$$

Po úpravách získáme následující kvadratickou rovnici pro q

$$q^2 - \left(2q_c\frac{h}{h_c}\right) q + \left(q_c^2\frac{h}{h_c} - \frac{g}{2}\frac{h}{h_c}(h_c + h)(h_c - h)^2\right) = 0, \tag{1.96}$$

jejíž řešením je

$$q_{1,2} = q_c\frac{h}{h_c} \pm \sqrt{q_c^2\left(\frac{h^2}{h_c^2} - \frac{h}{h_c}\right) + \frac{g}{2}\frac{h}{h_c}(h_c + h)(h_c - h)^2}. \tag{1.97}$$

V případě, že zvolíme $h = h_c$, potom dosazením získáme $q = q_c$, neboť požadujeme, aby křivka procházela bodem (h_c, q_c), neboli body (h_l, q_l) nebo (h_r, q_r). Na obrázku 1.19 je zobrazena Hugoniotova množina bodů $\mathbf{u}$, které mohou být spojeny se stavem $\mathbf{u}_l$ 1. „rázovou vlnou“. Uvozovky jsou použity z toho důvodu, že jde o vlnu, která splňuje Rankineovu-Hugoniotovu podmínku. Aby šlo o rázovou vlnu (tedy entropické řešení) musí být navíc splněna podmínka entropie, viz dále. Tedy Hugoniotovou množinou jsou pouze některé body vyznačené (čárkované) křivky. Na Obr. 1.20 je Hugoniotova množina bodů, které mohou být spojeny 2. „rázovou vlnou“. Je potřeba zde uvést, že 1. nebo 2. rázová vlna není určena znaménky + nebo − v rovnici (1.97). Volbou znaménka získáme pouze část množiny, druhá část patří ke druhé množině.

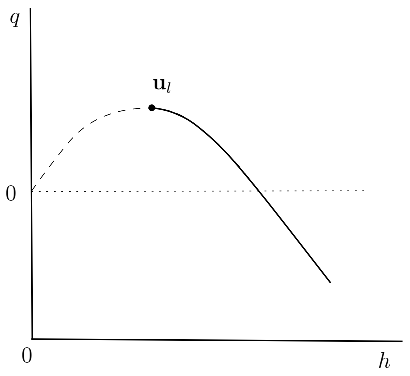
Abychom získali entropické řešení, je potřeba splnit tzv. podmínku entropie (konkrétně používáme Laxovu podmínku entropie), tedy musí platit pro 1. rázovou vlnu spojující stav $\mathbf{u}_l$ a $\mathbf{u}^*$ (ze všech výše určených $\mathbf{u}$ vybereme ty, co splňují Laxovu podmínku entropie a označíme je $\mathbf{u}^*$)

$$\lambda^1(\mathbf{u}_l) > s^1 > \lambda^1(\mathbf{u}^*) \tag{1.98}$$

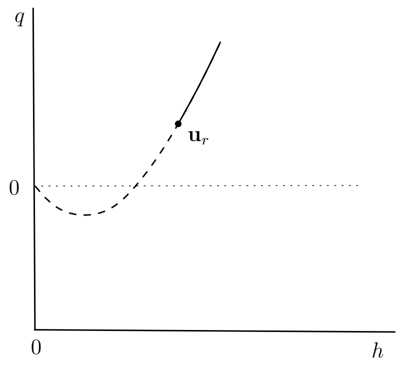
a pro 2. rázovou vlnu spojující stav $\mathbf{u}^*$ a $\mathbf{u}_r$

$$\lambda^2(\mathbf{u}^*) > s^2 > \lambda^2(\mathbf{u}_r). \tag{1.99}$$

V případě našeho modelu proudění neměnným obdélníkovým korytem lze odvodit jednoduché kritérium, abychom určili, které body na křivce na Obr. 1.19 jsou Hugoniotovou množinou (tj. splňující podmínku entropie), a které ne. V případě 1. rázové



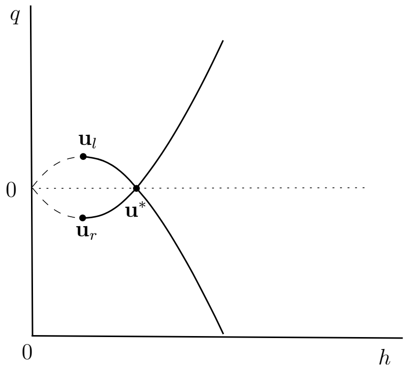
Obr. 1.19 Hugoniotova množina bodů spojující stav $\mathbf{q}_l$ 1. rázovou vlnou.



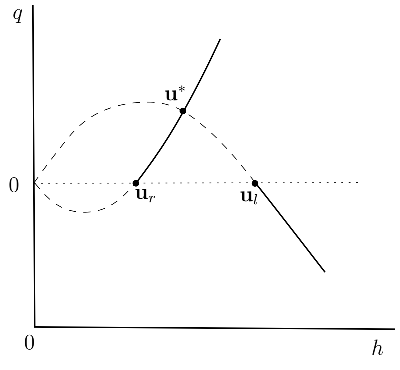
Obr. 1.20 Hugoniotova množina bodů spojující stav $\mathbf{q}_r$ 2. rázovou vlnou.

vlny, tj. vlny spojující stav $\mathbf{u}_l$ a $\mathbf{u}^*$, musí charakteristická rychlost $\lambda^1 = v - \sqrt{gh}$ klesat (jak plyne z podmínky (1.98)). Rankineova-Hugoniotova podmínka nám tedy říká, že h musí růst, a tedy musí platit $h^* > h_l$. Stejnou úvahu lze provést pro 2. rázovou vlnu, kdy získáme podmínku $h^* > h_r$. Hugoniotovy množiny jsou pro jednotlivé rázové vlny vyznačeny na Obr. 1.19 a 1.20 plnou čarou.

Nyní uvedeme dvě možné kombinace vln, viz Obr. 1.21 a 1.22. Rozeberme nejprve první situaci. Na Obr. 1.21 jsou dvě množiny Hugoniotových bodů pro stavy $\mathbf{u}_l$ a $\mathbf{u}_r$ zobrazeny plnou čarou. Jak bylo řečeno v kapitole 1.5, řešení Riemannova problému $\mathbf{u}^*$ je na průsečíku uvedených dvou křivek. Protože pro obě vlny, splňující Rankineovu-Hugoniotovu podmínku (leží na křivce vyznačené čárkovaně) splňují i entropickou podmínku (leží zároveň na plně vyznačené části křivek), tedy platí výše uvedené nerovnosti $h^* > h_l$ a $h^* > h_r$, jsou obě vlny rázové, tj. stav $\mathbf{u}_l$ je spojen se stavem $\mathbf{u}^*$ 1. rázovou vlnou a stav $\mathbf{u}_r$ je spojen se stavem $\mathbf{u}^*$ 2. rázovou vlnou. Jiná



Obr. 1.21 Hugoniotova množina bodů pro dvě rázové vlny.



Obr. 1.22 Hugoniotova množina bodů pro 2. rázovou vlnu.

je situace na Obr. 1.22. Jak je vidět, opět je řešení průsečíkem obou čárkovaných křivek, tj. opět splňuje Rankineovu-Hugoniotovu podmínku, ovšem pouze 2. vlna je vlnou rázovou, tj. splňující zároveň entropickou podmínku (leží na plné čáře), tj.

platí $h^* > h_r$. První vlna, spojující stav $\mathbf{u}_l$ se stavem $\mathbf{u}^*$ není rázovou vlnou, neboť nesplňuje podmínku entropie, tj. neleží na plné čáře, tj. neplatí nerovnost $h^* > h_l$. Protože charakteristické pole definované vlastním číslem λ^1 je ryze nelineární, může tyto stavy spojovat vlna zředění.

Nyní se tedy zabývejme vlnou zředění, což je spojité řešení Riemannova problému. Opět v úvodu hledání řešení stručně nastíníme jednotlivé kroky. Místo hledání Hugoniotových množin budeme hledat integrální křivky a opět pouze tu část, která splňuje podmínku entropie. Na takovýchto křivkách určíme chování vlny zředění. Jak již bylo řečeno v odstavci 1.6.1 a v 1.7.1, lze řešení $\mathbf{u}(x,t)$ hledat ve tvaru $\mathbf{u}(\eta)$, kde $\eta = x/t$, která splňuje následující rovnici

$$\mathbf{f}'(\mathbf{u}(\eta))\mathbf{u}'(\eta) = \eta\mathbf{u}'(\eta). \tag{1.100}$$

V případě skalární nelineární rovnice lze eliminovat $u'(\eta)$, jak bylo ukázáno v odstavci 1.6.1, ovšem v případě nelineárních soustav je $\mathbf{u}'(\eta)$ vektor. Podle (1.100) je nutné, aby vektor $\mathbf{u}'(\eta)$ byl vlastním vektorem Jacobiho matice $\mathbf{f}'(\mathbf{u}(\eta))$ pro každou hodnotu $\eta = x/t$. Neboli hledáme integrální křivku vektorového pole $\mathbf{r}^{1,2}$.

Integrální křivka $\mathbf{u}(\eta)$ parametrizovaná skalárním parametrem η je definována jako křivka, v jejímž každém bodě je vektor $\mathbf{r}^{1,2}$ ve směru tečny, tj. ve směru vektoru $\mathbf{u}'(\eta)$. Pokud máme množinu vlastních vektorů, potom skalární násobek je také vlastním vektorem, tedy platí

$$\mathbf{u}'(\eta) = \alpha(\eta)\mathbf{r}^{1,2}(\mathbf{u}(\eta)), \tag{1.101}$$

kde $\alpha(\eta)$ závisí na zvolené parametrizaci η a zvolené normalizaci vektorů $\mathbf{r}^{1,2}$.

Pro náš model zvolíme pro jednoduchost $\alpha(\eta) = 1$ a tedy musí platit:

$$\mathbf{u}'(\eta) = \mathbf{r}^{1,2}(\mathbf{u}(\eta)) = \begin{bmatrix} 1 \\ \frac{q}{h} \mp \sqrt{gh} \end{bmatrix}. \tag{1.102}$$

Tím získáme soustavu dvou obyčejných diferenciálních rovnic

$$h'(\eta) = 1 \tag{1.103}$$
$$q'(\eta) = \frac{q}{h} \mp \sqrt{gh}. \tag{1.104}$$

Budeme hledat řešení ve tvaru $q(h)$, tj. parametrizované podle h, abychom mohli řešení jednoduše zobrazit ve fázové rovině. První diferenciální rovnice má jednoduché řešení:

$$\frac{dh(\eta)}{d\eta} = 1 \quad \Rightarrow \quad h(\eta) = \eta. \tag{1.105}$$

Dosadíme do druhé rovnice a získáme následující diferenciální rovnici s pravou stranou, ke které připojíme počáteční podmínku $q(h_c) = q_c$, protože požadujeme, aby integrální křivka procházela ve fázové rovině bodem (h_c, q_c), kde $q_c = h_c v_c$,

$$\frac{dq(\eta)}{d\eta} = \frac{q}{\eta} \mp \sqrt{g\eta}. \tag{1.106}$$

Protože jde o rovnici s pravou stranou, musíme nejprve nalézt homogenní řešení (nulová pravá strana), poté partikulární řešení, například metodou variace konstant. Součtem homogenního a partikulárního řešení a zohledněním počátečních podmínek získáme obecné řešení diferenciální rovnice (1.106). Tedy nejprve řešme homogenní rovnici:

$$\frac{dq_H(\eta)}{d\eta} = \frac{q_H}{\eta} \quad \Rightarrow \quad q_H(\eta) = C\eta. \tag{1.107}$$

Dále hledáme partikulární řešení metodou variace konstant, tedy předpokládáme partikulární řešení v následujícím tvaru

$$q_p(\eta) = C(\eta)\eta, \tag{1.108}$$

určíme derivaci

$$\frac{dq_p(\eta)}{d\eta} = C'(\eta)\eta + C(\eta) \tag{1.109}$$

a dosazením do diferenciální rovnice (1.106) získáme

$$C'(\eta) = \mp\sqrt{\frac{g}{\eta}}, \tag{1.110}$$

tedy přímou integrací získáme

$$C(\eta) = \mp 2\sqrt{g\eta}. \tag{1.111}$$

Partikulární řešení má následující tvar

$$q_P(\eta) = \mp 2\eta\sqrt{g\eta}. \tag{1.112}$$

Obecné řešení po návratu proměnné h je

$$q(h) = Kh \mp 2h\sqrt{gh}, \tag{1.113}$$

zohledněním počátečních podmínek získáme

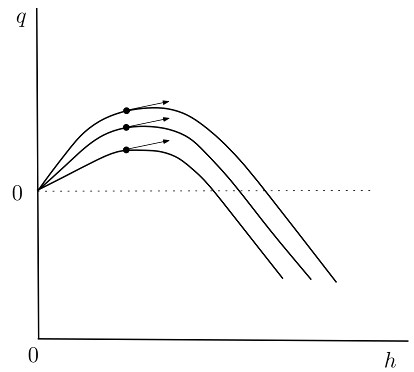
$$q(h) = hv_c \pm 2h(\sqrt{gh_c} \mp \sqrt{gh}). \tag{1.114}$$

Vypočtené integrální křivky jsou zobrazeny na Obr. 1.23 a 1.24. Upozorňujeme, že nejde o stejné křivky jako v případě Hugoniotových množin, ačkoliv jsou velmi podobné.

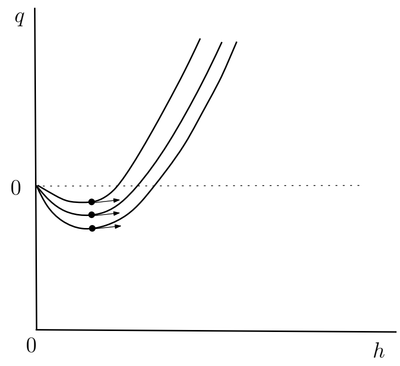
Pro zájemce:

Pouze pro zajímavost zde uvedeme pojem Riemannových invariantů. Vyjděme ze vztahu 1.114, máme tedy následující dva vztahy

$$hv = hv_c + 2h(\sqrt{gh_c} - \sqrt{gh}), \quad \text{resp.} \quad hv = hv_c - 2h(\sqrt{gh_c} + \sqrt{gh}). \tag{1.115}$$



Obr. 1.23 Integrální křivky pro vektorové pole $\mathbf{r}^1$.



Obr. 1.24 Integrální křivky pro vektorové pole $\mathbf{r}^2$.

po zkrácení h a drobné úpravě, získáváme

$$v + 2\sqrt{gh} = v_c + 2\sqrt{gh_c}, \quad \text{resp.} \quad v - 2\sqrt{gh} = v_c - 2\sqrt{gh_c}. \tag{1.116}$$

Uvědomme si, že body (h_c, v_c) a (h, v) jsou dva body na dané integrální křivce pro vektorové pole $\mathbf{r}^1$ resp. $\mathbf{r}^2$, a tedy funkce

$$w^1(\mathbf{u}) = v + 2\sqrt{gh}, \quad \text{resp.} \quad w^2(\mathbf{u}) = v - 2\sqrt{gh} \tag{1.117}$$

mají stále stejnou hodnotu podél odpovídající integrální křivky. Tyto funkce se nazývají 1. resp. 2. Riemannův invariant.

Vraťme se k vlně zředění (konkrétně středová vlna zředění), jde o řešení v následujícím tvaru

$$\mathbf{u}(x,t) = \begin{cases} \mathbf{u}_l & \text{pro} \quad x/t \leqq \eta_1, \\ \mathbf{u}(\eta) & \text{pro} \quad \eta_1 \leqq x/t \leqq \eta_2, \\ \mathbf{u}_r & \text{pro} \quad \eta_2 \leqq x/t, \end{cases} \tag{1.118}$$

kde $\mathbf{u}_l$ a $\mathbf{u}_r$ jsou dva body na stejné integrální křivce, pro které platí $\lambda^p(\mathbf{u}_l) < \lambda^p(\mathbf{u}_r)$ (abychom získali fyzikálně správné, tj. entropické řešení), viz podkapitola 1.7.1. Tedy hledáme řešení $\mathbf{u}(x,t) = \mathbf{u}(\eta)$, které splňuje rovnici (1.100) a leží na integrální křivce příslušející jednomu z vlastních vektorů $\mathbf{r}^{1,2}$. Také bylo řečeno, že $\mathbf{u}'(\eta)$ je vlastním vektorem Jacobiho matice $\mathbf{f}'(\mathbf{u}(\eta))$, tedy platí následující

$$[\mathbf{f}(\mathbf{u}(\eta)) - \eta]\, \mathbf{u}'(\eta) = \mathbf{0} \quad \Rightarrow \quad [\lambda^p(\mathbf{u}(\eta)) - \eta]\, \mathbf{u}'(\eta) = \mathbf{0}, \tag{1.119}$$

neboli

$$\lambda^p(\mathbf{u}(\eta)) = \eta. \tag{1.120}$$

Výše uvedené lze psát přímo z rovnice (1.100), neboť podle definice vlastních čísel a vlastních vektorů musí platit, že pokud $\mathbf{u}'(\eta)$ je vlastním vektorem, pak $\eta = \frac{x}{t}$ je vlastním číslem.

Zderivujeme (1.120) podle η:

$$1 = \nabla\lambda^p(\mathbf{u}(\eta))\mathbf{u}'(\eta), \tag{1.121}$$

kde $\nabla\lambda^p(\mathbf{u}(\eta))$ je gradient p-tého vlastního čísla. Využitím (1.101) a drobnou úpravou můžeme formulovat soustavu obyčejných diferenciálních rovnic

$$\mathbf{u}'(\eta) = \frac{\mathbf{r}^p(\mathbf{u}(\eta))}{\nabla\lambda^p(\mathbf{u}(\eta))\mathbf{r}^p(\mathbf{u}(\eta))}. \tag{1.122}$$

Řešením v intervalu $\eta_1 \leqq \eta \leqq \eta_2$ s počátečními podmínkami:

$$\mathbf{u}(\eta_1) = \mathbf{u}_l, \quad \mathbf{u}(\eta_2) = \mathbf{u}_r \tag{1.123}$$

získáme vlnu zředění uvnitř tzv. vějíře zředění, viz Obr. 1.18. Popsaný postup je obecný, nyní určíme soustavu obyčejných diferenciálních rovnic pro náš model (1.87). Vlastní čísla a vlastní vektory lze nalézt v (1.90) a (1.91), proto zde uveďme pouze následující výpočty:

$$\nabla\lambda^1(\mathbf{u}(\eta)) = \begin{bmatrix} -\frac{q}{h^2} - \frac{1}{2}\sqrt{\frac{g}{h}} \\ \frac{1}{h} \end{bmatrix}, \quad \nabla\lambda^2(\mathbf{u}(\eta)) = \begin{bmatrix} -\frac{q}{h^2} + \frac{1}{2}\sqrt{\frac{g}{h}} \\ \frac{1}{h} \end{bmatrix}. \tag{1.124}$$

$$\nabla\lambda^1(\mathbf{u}(\eta))\mathbf{r}^1(\mathbf{u}(\eta)) = -\frac{3}{2}\sqrt{\frac{g}{h}}, \quad \nabla\lambda^2(\mathbf{u}(\eta))\mathbf{r}^2(\mathbf{u}(\eta)) = \frac{3}{2}\sqrt{\frac{g}{h}}. \tag{1.125}$$

Dosazením do (1.122), získáme následující obyčejné diferenciální rovnice

$$\mathbf{u}'(\eta) = -\frac{2}{3}\sqrt{\frac{h}{g}}\begin{bmatrix} 1 \\ \frac{q}{h} - \sqrt{gh} \end{bmatrix}, \quad \mathbf{u}'(\eta) = \frac{2}{3}\sqrt{\frac{h}{g}}\begin{bmatrix} 1 \\ \frac{q}{h} + \sqrt{gh} \end{bmatrix} \tag{1.126}$$

s počátečními podmínkami (1.123). Podrobně vyřešíme pouze první ze dvou soustav, máme tedy následující soustavu

$$h'(\eta) = -\frac{2}{3}\sqrt{\frac{h(\eta)}{g}}, \tag{1.127}$$

$$q'(\eta) = -\frac{2}{3}\frac{q}{\sqrt{gh}} + \frac{2}{3}h. \tag{1.128}$$

Separací proměnných a poté integrací (1.127) získáme

$$\frac{dh(\eta)}{\sqrt{h}} = -\frac{2}{3}\frac{1}{\sqrt{g}}d\eta \quad \Rightarrow \quad h(\eta) = \frac{1}{9g}(K-\eta)^2. \tag{1.129}$$

Dále musíme identifikovat konstantu K zohledněním počátečních podmínek (1.123) a též rovnice (1.120), neboli

$$h(\eta_1) = h_l, \qquad \eta_1 = \lambda^1(\mathbf{u}(\eta_1)) = \frac{q_l}{h_l} - \sqrt{gh_l} = v_l - \sqrt{gh_l}, \tag{1.130}$$

$$h(\eta_2) = h_r, \qquad \eta_2 = \lambda^1(\mathbf{u}(\eta_2)) = \frac{q_r}{h_r} - \sqrt{gh_r} = v_r - \sqrt{gh_r}. \tag{1.131}$$

Po užití obou podmínek získáme hodnotu konstanty K:

$$K = v_l + 2\sqrt{gh_l} = v_r + 2\sqrt{gh_r}. \tag{1.132}$$

Pro zájemce:

Opět zde pro zájemce nastíníme souvislost konstanty K a Riemannova invariantu. Protože K je 1. Riemannovým invariantem, viz (1.117), tj. funkce, která je podél odpovídající integrální křivky konstantní, není tedy nutné dosazovat druhou podmínku a rovnítko v (1.132) je možno psát ihned.

Řešme nyní druhou diferenciální rovnici, tj. (1.128):

$$\frac{dq(\eta)}{d\eta} = -\frac{2}{3}\frac{q}{\sqrt{gh}} + \frac{2}{3}h. \tag{1.133}$$

Protože řešení $h(\eta)$ již známe, dosadíme (1.129) do (1.133) a získáme následující diferenciální rovnici s pravou stranou

$$\frac{dq(\eta)}{d\eta} = -\frac{2q}{K-\eta} + \frac{2}{27g}(K-\eta)^2, \tag{1.134}$$

kde konstanta K je definována v (1.132). Opět budeme hledat homogenní a partikulární řešení jejichž součtem získáme obecné řešení nehomogenní rovnice. Tedy nejprve hledejme homogenní řešení, tj. řešíme rovnici (1.134) s nulovou pravou stranou:

$$\frac{dq_H(\eta)}{d\eta} = -\frac{2q_H}{K-\eta}, \tag{1.135}$$

separací proměnných a následnou přímou integrací získáme

$$q_H(\eta) = C_1(K-\eta)^2, \tag{1.136}$$

kde C_1 je libovolná konstanta. Tuto konstantu upravíme ještě před dalším výpočtem následovně (z důvodu jednoduššího vyjádření řešení): $C_1 = \frac{C}{27g}$, kde C je stále libovolná konstanta. Získáme následující homogenní řešení

$$q_H(\eta) = \frac{C}{27g}(K-\eta)^2. \tag{1.137}$$

Partikulární řešení nalezneme metodou variace konstant, tedy hledáme jej ve tvaru

$$q_P(\eta) = \frac{C(\eta)}{27g}(K-\eta)^2. \tag{1.138}$$

Zderivujeme partikulární řešení podle η:

$$\frac{dq_P(\eta)}{d\eta} = \frac{C'(\eta)}{27g}(K-\eta)^2 - 2\frac{C(\eta)}{27g}(K-\eta). \tag{1.139}$$

Dosazením (1.138) na (1.139) do diferenciální rovnice (1.133) několika drobnými úpravami a poté přímou integrací získáme

$$C'(\eta) = 2 \Rightarrow C(\eta) = 2\eta. \tag{1.140}$$

Partikulární řešení má tedy tvar

$$q_p(\eta) = \frac{2}{27g}\eta(K-\eta)^2. \tag{1.141}$$

Obecné řešení získáme součtem homogenního a partikulárního řešení

$$q(\eta) = \frac{C}{27g}(K-\eta)^2 + \frac{2}{27g}\eta(K-\eta)^2. \tag{1.142}$$

Posledním krokem je určení konstanty C zohledněním počátečních podmínek a též dosazení konstanty K, tj.

$$q(\eta_1) = q_l, \qquad \eta_1 = \lambda^1(\mathbf{u}(\eta_1)) = v_l - \sqrt{gh_l}, \quad K = v_l + 2\sqrt{gh_l} \tag{1.143}$$
$$q(\eta_2) = q_r, \qquad \eta_2 = \lambda^1(\mathbf{u}(\eta_2)) = v_r - \sqrt{gh_r}, \quad K = v_r + 2\sqrt{gh_r}. \tag{1.144}$$

Opět získáme

$$C = v_l + 2\sqrt{gh_l} = v_r + 2\sqrt{gh_r} = K. \tag{1.145}$$

Je vidět, že konstanta je stejná jako předchozí konstanta K (z tohoto důvodu jsme mírně upravovali tvar konstanty C_1). Pokud bychom chtěli vyjádřit závislost $q(h)$, využijeme řešení (1.129), a získáme

$$q(h) = Kh - 2h\sqrt{gh}, \tag{1.146}$$

neboli

$$v(h) = K - 2\sqrt{gh}, \tag{1.147}$$

Nyní tedy známe chování $\mathbf{u}$ přes vlnu zředění.

Pro zájemce:

Stejně jako v předchozím případě je konstanta C 1. Riemannovým invariantem (tj. funkce, která je konstantní na odpovídající integrální křivce), tedy rovnítko mezi $v_l + 2\sqrt{gh_l}$ a $v_r + 2\sqrt{gh_r}$ je oprávněné.

Dále zde pro zájemce načrtneme jiný, elegantnější, způsob řešení druhé rovnice. Protože známe chování $h(\eta)$ a známe i chování integrální křivky (stavy popisující vlnu zředění musí ležet na odpovídající integrální křivce). Lze tedy pomocí řešení (1.129) s konstantou definovanou v (1.132) a odpovídající charakteristikou (po vydělení h), viz (1.114), a využitím vlastnosti 1. Riemannového invariantu, získáme

$$\begin{aligned} v(h) &= v_l + 2\sqrt{gh_l} - 2\sqrt{gh} = && (1.148)\\ &= v_r + 2\sqrt{gh_r} - 2\sqrt{gh} = && (1.149)\\ &= K - 2\sqrt{gh}. && (1.150) \end{aligned}$$

Získali jsme tedy stejné řešení druhé rovnice, viz (1.147), a řešili jsme pouze jednodušší diferenciální rovnici.

Řešením druhé soustavy (1.126), získáme následující řešení (detailní řešení necháváme již na čtenáři, postup je identický jako v předchozím případě):

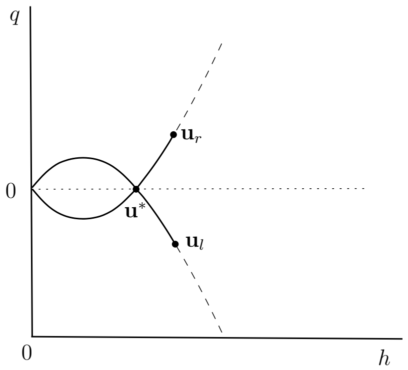
$$h(\eta) = \frac{1}{9g}(\eta - K)^2, \tag{1.151}$$

kde $K = u_l - 2\sqrt{gh_l} = u_r - 2\sqrt{gh_r}$. A řešení druhé rovnice je

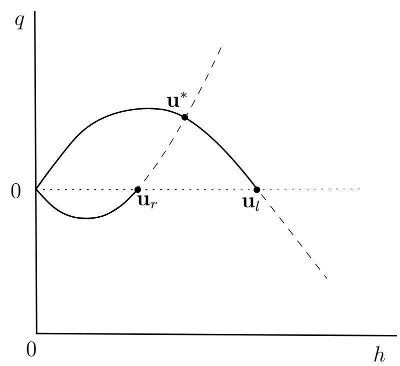
$$q(\eta) = \frac{C}{27g}(\eta - K)^2 + \frac{2}{27g}\eta(\eta - K), \tag{1.152}$$

kde $C = u_l - 2\sqrt{gh_l} = u_r - 2\sqrt{gh_r} = C$.

Pro představu uveďme dvě možnosti, jak může řešení vlny zředění vypadat. Na Obr. 1.25 a 1.26 jsou dvě možné situace. Řešení na Obr. 1.25 zobrazuje dvě vlny



Obr. 1.25 Integrální křivky pro dvě vlny zředění.



Obr. 1.26 Integrální křivka pro 1. vlnu zředění.

zředění, obě splňují podmínku (1.86), neboť například u druhé vlny spojující stav $\mathbf{u}^*$ a $\mathbf{u}_r$ platí:

$$h^* < h_r \quad \wedge \quad v^* < v_r \quad \Rightarrow \quad v^* + \sqrt{gh^*} < v_r + \sqrt{gh_r}. \tag{1.153}$$

Na Obr. 1.26 je stav $\mathbf{u}_l$ se stavem $\mathbf{u}^*$ vlnou zředění, kdežto druhá vlna nemůže být vlnou zředění, neboť nesplňuje podmínku (1.86). ▲

1.7.2 Zobecněný Riemannův problém, věta o řešení zobecněného Riemannova problému

Zobecnění Riemannova problému (1.77) a (1.78) lze provést dvěma způsoby. První možností je zobecnění samotné soustavy rovnic (přidáním zdrojového členu), lze nalézt například v [5], nebo zobecnění počáteční podmínky, kdy se nebudeme omezovat pouze po částech konstantní funkcí, viz například v [18], [19]. Poslední možností je samozřejmě kombinace obou předchozích.

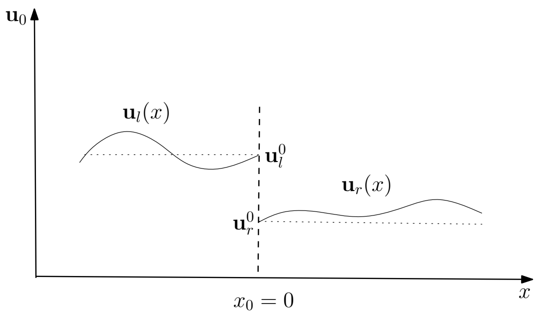
Uvažujme opět soustavu nelineárních hyperbolických rovnic

$$\mathbf{u}_t + [\mathbf{f}(\mathbf{u})]_x = \mathbf{0}, \quad x \in \mathbb{R}, t \in (0,T),\, T > 0, \tag{1.154}$$

s počátečními podmínkami

$$\mathbf{u}(x,0) = \mathbf{u}_0(x) = \begin{cases} \mathbf{u}_l(x) & \text{pro} \quad x < 0, \\ \mathbf{u}_r(x) & \text{pro} \quad x > 0, \end{cases} \tag{1.155}$$

kde $\mathbf{u}_l(x) : \mathbb{R}^- \to \mathbb{R}^m$ a $\mathbf{u}_r(x) : \mathbb{R}^+ \to \mathbb{R}^m$ jsou dané omezené, po částech hladké funkce (obvykle po částech polynomiální funkce) a $\mathbf{u}_l(0) \neq \mathbf{u}_r(0)$, viz Obr. 1.27.



Obr. 1.27Počáteční podmínky zobecněného Riemannova problému.

Označme hodnoty počátečních podmínek zobecněného Riemannova problému v bodě $x_0 = 0$

$$\mathbf{u}_l^0 = \mathbf{u}_l(0), \qquad \mathbf{u}_r^0 = \mathbf{u}_r(0) \tag{1.156}$$

a definujme klasický Riemannův problém, viz podkapitola 1.7.1,

$$\mathbf{u}_t^0 + [\mathbf{f}(\mathbf{u}^0)]_x = \mathbf{0}, \quad x \in \mathbb{R}, t \in (0,T),\, T > 0, \tag{1.157}$$

s počátečními podmínkami

$$\mathbf{u}^0(x,0) = \begin{cases} \mathbf{u}_l^0 & \text{pro} \quad x < 0, \\ \mathbf{u}_r^0 & \text{pro} \quad x > 0. \end{cases} \tag{1.158}$$

Takto definovaný Riemannův problém je častěji označován jako přidružený. Jak bylo již popsáno v podkapitole 1.7.1, má klasický Riemannův problém slabé entropické řešení (pro $|\mathbf{u}_r^0 - \mathbf{u}_l^0|$ dostatečně malé), které můžeme zapsat ve tvaru

$$\mathbf{u}^0(x,t) = \mathbf{u}^0\left(\frac{x}{t}\right) \tag{1.159}$$

a skládá se z $m + 1$ stavů spojených vlnou zředění, kontaktní nespojitostí nebo rázovou vlnou.

Následující věta dává do souvislosti strukturu řešení zobecněného Riemannova problému a tzv. přidruženého (tedy klasického) Riemannova problému a je důležitým předpokladem pro odvození numerických schémat vyššího řádu. Podrobnosti lze nalézt v [18] a v [20].

Věta 1.16 (Li, Yu). *Předpokládejme že skok $|\mathbf{u}_r^0 - \mathbf{u}_l^0|$ je dostatečně malý, potom v $R(\delta) = \{(x,t)|0 \leqq t \leqq \delta, -\infty < x < \infty\}$ má zobecněný Riemannův problém (1.154), (1.155) jednoznačné řešení, které má v okolí počátku stejnou strukturu jako řešení přidruženého Riemannova problému (1.157), (1.158).*

1.8 Parabolické rovnice

V předchozích kapitolách jsme se věnovali ryze hyperbolickým evolučním úlohám, které se typicky používají v modelování proudění. V této kapitole se budeme zabývat parabolickými rovnicemi, které jsou naopak typické v modelování vedení tepla nebo difúze. Formulujme tedy následující skalární lineární počátečně-okrajovou úlohu

$$u_t = Du_{xx}, \quad x \in (a,b),\ t \in (0,T),\ T > 0,\ D \in \mathbb{R}^+, \tag{1.160}$$
$$u(x,0) = u_0(x), \quad x \in (a,b), \tag{1.161}$$
$$u(a,t) = g_0(t), \tag{1.162}$$
$$u(b,t) = g_1(t), \tag{1.163}$$

kde $u(x,t)$ je hledaná zachovávaná veličina a D je koeficient, jehož konkrétní tvar závisí na modelovaném problému, viz dále, $g_0(t) : (0,T) \to \mathbb{R}$ a $g_1(t) : (0,T) \to \mathbb{R}$ jsou Dirichletovy okrajové podmínky. Budeme předpokládat, že jsou splněny tzv. podmínky kompatibility, tj. je splněno následující:

$$g_0(0) = u_0(a), \quad g_1(0) = u_0(b). \tag{1.164}$$

Zjednodušeně řečeno počáteční podmínka musí splňovat zadané okrajové podmínky (okrajové podmínky na sebe musí spojitě navazovat).

Jako konkrétní modely můžeme uvést například rovnici vedení tepla uvnitř tyče, tedy

$$\rho c u_t = \lambda u_{xx}, \tag{1.165}$$

kde $u = u(x,t)$ reprezentuje časový vývoj teploty, λ je koeficient pro vedení tepla, ρ je hustota materiálu a c je měrná tepelná kapacita materiálu.

Dalším příkladem může být zmíněný model difúze, tedy

$$u_t = Du_{xx}, \tag{1.166}$$

kde $u(x,t)$ reprezentuje koncentraci částic a D je difúzní koeficient.

Nyní uvedeme definice klasického a princip maxima, detaily lze nalézt například v [25].

Definice 1.17 (Klasické řešení). *Klasickým řešením* úlohy (1.160), (1.161), (1.162), (1.163) nazveme $u(x,t)$ takové, že $u \in \mathrm{C}(\langle a,b\rangle \times (0,T))$, má všechny derivace obsažené v rovnici (1.160) spojité na $(a,b)\times(0,T)$, tj. $\frac{\partial^2 u}{\partial x^2} \in \mathrm{C}((a,b)\times(0,T))$, $\frac{\partial u}{\partial t} \in \mathrm{C}((a,b)\times(0,T))$, a splňuje rovnici (1.160) na $(a,b)\times(0,T)$, počáteční podmínku na (a,b) a okrajovou podmínku v bodech a a b a podmínky kompatibility.

Věta 1.18 (Slabý princip maxima). *Nechť $u(x,t)$ je klasické řešení úlohy (1.160), (1.161), (1.162), (1.163), potom $u(x,t)$ nabývá svého maxima a minima na hranici, tj. v bodech a, b nebo v počátku, tj. v čase $t = 0$.*

Princip maxima nám v podstatě říká, že například rozložení teploty v tyči, viz model (1.165), kde není přítomen žádný zdroj tepla (tyč nic neohřívá), musí být maximální teplota buď v počátečním okamžiku (počáteční podmínka) nebo na krajích tyče. Dalším velkým rozdílem oproti evolučním hyperbolickým rovnicím (které mohou obsahovat nespojité řešení, tvz. rázovou vlnu) mají hyperbolické rovnice zhlazovací efekt. Tedy pro $t > 0$ zhlazují nespojitosti v počáteční podmínce (porovnej s řešením advekční rovnice, viz kapitola 1.3, která nespojitost pouze transportuje).

Pro úlohu (1.160), (1.161), (1.162), (1.163) lze ukázat, že existuje právě jedno klasické řešení, viz například [25].

Σ Pojmy k zapamatování

— hyperbolické evoluční rovnice
— Saint-Venantovy rovnice, Eulerovy rovnice, vazká a nevazká Burgersova rovnice
— kinematická vlnová aproximace (model KWA)
— klasické, slabé a entropické řešení,
— slabě hyperbolická, hyperbolická a ryze hyperbolická soustava
— Rankineův-Hugoniotův vztah
— skalární lineární rovnice (advekční rovnice) a její řešení
— Riemannův problém a jeho řešení pro skalární lineární rovnici
— charakteristiky (charakteristický systém)
— soustava lineárních rovnic a její řešení
— oblast závislosti
— Riemannův problém a jeho řešení pro soustavu lineárních rovnic
— skalární nelineární rovnice a vlastnosti jejího řešení
— Riemannův problém a jeho řešení pro skalární nelineární rovnici
— soustavy nelineárních rovnic
— Riemannův problém a jeho řešení pro soustavu nelineárních rovnic

— Zobecněný Riemannův problém a souvislost jeho řešení s řešením klasického Riemannova problému

— Hugoniotovy množiny

— parabolické evoluční rovnice

— rovnice vedení tepla, rovnice difúze

— klasické řešení, princip maxima, zhlazovací efekt parabolických rovnic

Kapitola 2

Diferenční metody

Průvodce studiem

V této kapitole se čtenář seznámí s řadou numerických metod, které lze použít na řešení úloh založených na evolučních rovnicích hyperbolického typu. Půjde o skalární problémy i o problémy popsané soustavami rovnic. Dozví se také, jaké vlastnosti rozhodují o kvalitě metod a přesnosti získaného řešení.

Cíle

Po prostudování této kapitoly budete schopni:

- sestavit a aplikovat několik numerických metod pro řešení příslušných parciálních diferenciálních rovnic
- rozumět vlastnostem metod nutných pro získání korektních řešení
- rozumět rozdílům mezi jednotlivými metodami a posoudit vhodnost jejich použití na konkrétní problémy
- definovat důležité pojmy jako konzistence, konzervativita a stabilita

2.1 Skalární úlohy

Uvažujme nyní skalární případ. Stejně jako v případě vektorových funkcí 1.6, můžeme i zde sestavit počáteční nelineární úlohu

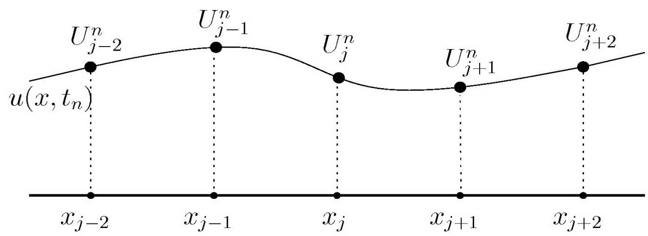
$$\begin{aligned} u_t + [f(u)]_x &= 0, \qquad t \in (0,T), x \in \mathbb{R}, T > 0, \\ u(x,0) &= u_0(x), \qquad x \in \mathbb{R}. \end{aligned} \tag{2.1}$$

nebo úlohu lineární

$$\begin{aligned} u_t + au_x &= 0, \qquad t \in (0,T), x \in \mathbb{R}, a \in \mathbb{R}, T > 0, \\ u(x,0) &= u_0(x), \qquad x \in \mathbb{R}, \end{aligned} \tag{2.2}$$

kde $u = u(x,t)$ a $f \in C^2(\mathbb{R})$. Předpokládejme, že funkce u_0 má omezený nosič, tedy $\text{supp}\ u_0 = \overline{\{x \in \mathbb{R} : u_0(x) \neq 0\}}$ je omezená množina.

Diferenční metody jsou založeny na tom, že si daný interval, na kterém hledáme řešení, rozdělíme na několik částí, které nazýváme subintervaly. Jednotlivé subintervaly jsou odděleny body x_j. Jsou-li všechny subintervaly stejně veliké, hovoříme o konstantním diskretizačním kroku $\Delta x = x_{j+1} - x_j$ a ekvidistantní síti. Dále si pro jednoduchost podobným způsobem zavedeme konstantní dělení časového intervalu s časovým krokem $\Delta t = t_{n+1} - t_n, n \in \mathbb{N}_0, T \geqq n\Delta t$. Budeme předpokládat, že $\frac{\Delta t}{\Delta x}$ = konst. Hodnotu přesného řešení v bodě x_j a čase t_n značíme $u_j^n = u(x_j, t_n)$, její aproximaci U_j^n. Dále definujme $\boldsymbol{U}^{(n)}$ jako vektor se složkami $U_j^n, j = \dots, -2, -1, 0, 1, 2, \dots$.



Obr. 2.1 Diskretizace neznámé funkce

2.1.1 Laxova-Friedrichsova metoda

Laxova-Friedrichsova metoda je jedna z centrálních metod. Tyto metody se tak nazývají proto, že k aproximacím prostorových derivací využívají centrálních diferencí. Konkrétně u této metody aproximujeme derivaci neznámé funkce v bodě x_j v čase t_n výrazem

$$u_x(x_j, t_n) \approx \frac{1}{2\Delta x}(U_{j+1}^n - U_{j-1}^n). \tag{2.3}$$

Ve smyslu této aproximace nahrazujeme i hodnotu neznámé funkce v bodě x_j a v čase t_n výrazem

$$u(x_j, t_n) \approx \frac{1}{2}(U_{j-1}^n + U_{j+1}^n). \tag{2.4}$$

Této hodnoty využijeme při stanovení aproximace časové derivace neznámé funkce

$$u_t(x_j, t_n) \approx \frac{1}{\Delta t}\left(u(x_j, t_{n+1}) - u(x_j, t_n)\right) \approx \frac{1}{\Delta t}\left(U_j^{n+1} - \frac{1}{2}(U_{j-1}^n + U_{j+1}^n)\right). \tag{2.5}$$

Dosazením aproximací (2.3) a (2.4) do diferenciální rovnice (2.2) získáme diskrétní analogii lineárního problému v následujícím tvaru

$$\frac{1}{\Delta t}\left(U_j^{n+1} - \frac{1}{2}(U_{j-1}^n + U_{j+1}^n)\right) + a\frac{1}{2\Delta x}(U_{j+1}^n - U_{j-1}^n) = 0, \tag{2.6}$$

ze kterého je možné jednoduchými úpravami získat hodnotu neznámé funkce na nové časové vrstvě, a tedy tvar Laxovy-Friedrichsovy metody pro lineární úlohu je

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{\Delta t}{2\Delta x} a(U_{j+1}^n - U_{j-1}^n). \tag{2.7}$$

V nelineárním případě je třeba obdobným způsobem zavést aproximaci derivace tokové funkce v daném bodě a to výrazem

$$[f(u)(x_j, t_n)]_x \approx \frac{1}{2\Delta x}(f(U_{j+1}^n) - f(U_{j-1}^n)). \tag{2.8}$$

Dosazení této aproximace a aproximací (2.3) a (2.5) do lineární rovnice (2.1) získáme Laxovu-Friedrichsovu metodu pro řešení nelineární rovnice

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{\Delta t}{2\Delta x}(f(U_{j+1}^n) - f(U_{j-1}^n)). \tag{2.9}$$

2.1.2 Laxova-Wendroffova metoda

Ukažme si nyní také odvození Laxovy-Wendroffovy metody pro lineární a nelineární případ. Při její konstrukci využijeme více členů Taylorova rozvoje než tomu bylo u Laxovy-Friedrichsovy metody a tím docílíme vyšší přesnosti aproximace. Předpokládáme, že existují veškeré derivace, se kterými pracujeme. V lineárním případě vyjdeme z diferenciálního zákona zachování ve tvaru

$$u_t = -au_x. \tag{2.10}$$

Na základě Taylorova rozvoje aproximujeme hodnotu $u(x_j, t_n)$ s přesností $O(\Delta t^3)$ vztahem

$$u(x_j, t_{n+1}) = u(x_j, t_n) + \frac{\partial u}{\partial t}(x_j, t_n)\Delta t + \frac{\partial^2 u}{\partial t^2}(x_j, t_n)\frac{\Delta t^2}{2}. \tag{2.11}$$

Zderivováním vztahu (2.10) získáme následující rovnosti

$$u_{tt} = -au_{xt} = -[au_t]_x = a^2 u_{xx}. \tag{2.12}$$

Dosazením rovností (2.12) a (2.10) do vztahu (2.11) získáme

$$u(x_j, t_{n+1}) = u(x_j, t_n) - a\frac{\partial u}{\partial x}(x_j, t_n)\Delta t + a^2 \frac{\partial^2 u(x_j, t_n)}{\partial x^2}\frac{\Delta t^2}{2}. \tag{2.13}$$

S využitím aproximací neznámé funkce a centrálních poměrných diferencí získáme Laxovu-Wendroffovu metodu ve tvaru

$$U_j^{n+1} = U_j^n - a\frac{\Delta t}{2\Delta x}(U_{j+1}^n - U_{j-1}^n) + \frac{a^2 \Delta t^2}{2\Delta x^2}(U_{j+1}^n - 2U_j^n + U_{j-1}^n). \tag{2.14}$$

V nelineárním případě má diferenciální zákon zachování tvar

$$u_t = -[f(u)]_x. \tag{2.15}$$

Hodnotu $u(x_j, t_n)$ opět aproximujeme vztahem (2.11). Zderivováním vztahu (2.15) získáme následující rovnosti

$$u_{tt} = -[f(u)]_{xt} = -[a(u)u_t]_x = [a(u)f(u)_x]_x, \tag{2.16}$$

kde $a(u) = f'(u)$. Dosazením rovností (2.16) a (2.15) do vztahu (2.11) získáme

$$u(x_j, t_{n+1}) = u(x_j, t_n) - \frac{\partial f(u)}{\partial x}(x_j, t_n)\Delta t + \frac{\partial}{\partial x}\left(a(u(x_j, t_n))\frac{\partial f(u(x_j, t_n))}{\partial x}\right)\frac{\Delta t^2}{2}. \tag{2.17}$$

S využitím aproximací neznámé funkce a jejích poměrných diferencí získáme Laxovu--Wendroffovu metodu ve tvaru

$$\begin{aligned} U_j^{n+1} &= U_j^n - \frac{\Delta t}{2\Delta x}(f(U_{j+1}^n) - f(U_{j-1}^n)) + \\ &+\frac{\Delta t^2}{2\Delta x^2}(a_{j+1/2}^n(f(U_{j+1}^n) - f(U_j^n)) - a_{j-1/2}^n(f(U_j^n) - f(U_{j-1}^n)), \end{aligned} \tag{2.18}$$

kde $a_{j+1/2}^n = \frac{f(U_{j+1}^n) - f(U_j^n)}{U_{j+1}^n - U_j^n}$.

Lze však získat i jiný tvar Laxovy-Wendroffovy metody. Při něm se částečně využije přístupu dvoukrokové metody, kdy nestanovujeme hodnoty na časové vrstvě t_{n+1} přímo z hodnot v čase t_n, ale nejprve určíme hodnoty v mezičase $t_{n+1/2}$. V tomto mezičase navíc změníme dělení ze subintervalů $\langle x_{j-1/2}, x_{j+1/2}\rangle$ na $\langle x_j, x_{j+1}\rangle$ (Obr. 2.2). Pro získání hodnot na v čase $t_{n+1/2}$ využijeme opět centrálních diferencí přesně tak, jako v případě Laxovy-Friedrichsovy metody

$$U_{j+1/2}^{n+1/2} = \frac{1}{2}(U_j^n + U_{j+1}^n) - \frac{\Delta t}{2\Delta x}(f(U_{j+1}^n) - f(U_j^n)). \tag{2.19}$$

Hodnoty $U_{j+1/2}^{n+1/2}$ využijeme pro aproximace tokových funkcí v diferenciálním zákoně zachování na původních intervalech $\langle x_{j-1/2}, x_{j+1/2}\rangle \times \langle t_n, t_{n+1}\rangle$ a získáme tak metodu druhého řádu přesnosti, která má tvar

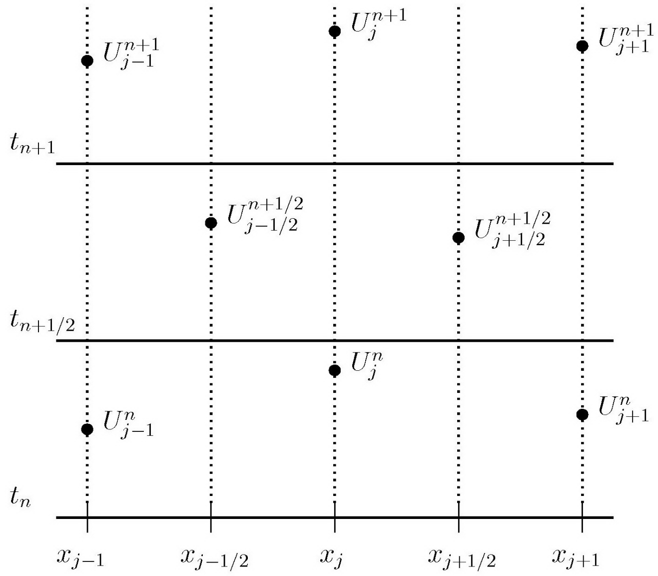
$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}(f(U_{j+1/2}^{n+1/2}) - f(U_{j-1/2}^{n+1/2})), \tag{2.20}$$

kde hodnoty $U_{j+1/2}^{n+1/2}$ jsou definovány vztahem (2.19).

2.1.3 MacCormackova metoda

MacCormackova metoda je dvoukroková metoda. Využívá pomocné hodnoty mezi časovými vrstvami t_n a t_{n+1} podobně jako tomu bylo u Laxovy-Wendroffovy metody. Nejprve získáme pomocnou hodnotu $\tilde{U}_j^{n+1}$ využitím vztahu

$$\tilde{U}_j^{n+1} = U_j^n - a\frac{\Delta t}{\Delta x}(U_{j+1}^n - U_j^n). \tag{2.21}$$



Obr. 2.2 Poloviční krok Laxovy-Wendroffovy metody

Tuto hodnotu následně využijeme k získání nové hodnoty neznámé funkce na časové vrstvě t_{n+1} vztahem.

$$U_j^{n+1} = \frac{U_j^n + \tilde{U}_j^{n+1}}{2} - a\frac{\Delta t}{2\Delta x}(\tilde{U}_j^{n+1} - \tilde{U}_{j-1}^{n+1}). \tag{2.22}$$

V nelineárním případě pak obdobným způsobem získáme metodu tvaru

$$U_j^{n+1} = \frac{U_j^n + \tilde{U}_j^{n+1}}{2} - \frac{\Delta t}{2\Delta x}(f(\tilde{U}_j^{n+1}) - f(\tilde{U}_{j-1}^{n+1})), \tag{2.23}$$

kde

$$\tilde{U}_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}(f(U_{j+1}^n) - f(U_j^n)). \tag{2.24}$$

2.1.4 Metoda typu upwind

Při řešení lineární skalární úlohy (2.2) touto metodou je důležité znaménko konstanty a, které určuje směrnici charakteristik. Jak již bylo zmíněno v kapitole 1, charakteristiky jsou v tomto případě přímky, na kterých je řešení dané úlohy konstantní. V případě, že $a < 0$, směrnice je záporná, charakteristiky směřují doleva a proto zavedeme diskretizaci úlohy pomocí dopředných diferencí

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + a\frac{U_{j+1}^n - U_j^n}{\Delta x} = 0. \tag{2.25}$$

Je to z toho důvodu, že se informace potřebné pro získání řešení na nové časové vrstvě díky směru charakteristik šíří zprava doleva. Pokud $a > 0$, situace je opačná, charakteristiky směřují doprava a my sestavíme metodu pomocí zpětné prostorové diference

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + a\frac{U_j^n - U_{j-1}^n}{\Delta x} = 0. \tag{2.26}$$

Abychom mohli zapsat metodu v kompaktnější podobě, která zahrnuje obě předchozí varianty, zavedeme hodnoty

$$a^+ = \max\{a, 0\}, \qquad a^- = \min\{a, 0\}. \tag{2.27}$$

Metodu typu upwind pak můžeme napsat následovně

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}\left[a^+(U_j^n - U_{j-1}^n) + a^-(U_{j+1}^n - U_j^n)\right]. \tag{2.28}$$

Metodu typu upwind pro nelineární skalární úlohu (2.1) můžeme psát v tzv. *divergentním tvaru*, který má následující podobu

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}(F_{j+1/2}^n - F_{j-1/2}^n). \tag{2.29}$$

funkce $F_{j+1/2}^n$ budeme nazývat numerické toky. K tomu, abychom určili hodnoty těchto numerických toků využijeme kvazilineárního tvaru naší úlohy. Ten získáme analogicky na základě tvaru lineární úlohy (2.2) zavedením

$$a(u) = \frac{\partial f(u)}{\partial u}. \tag{2.30}$$

Potom lze diferenciální rovnici v úloze (2.1) přepsat do kvazilineárního tvaru

$$u_t + f'(u)u_x = 0. \tag{2.31}$$

Diskrétní podobu vztahu (2.30) aproximujeme na časové vrstvě t_n mezi buňkami j a $j+1$ výrazem $a_{j+1/2}^n$

$$a_{j+1/2}^n = \frac{f(U_{j+1}^n) - f(U_j^n)}{U_{j+1}^n - U_j^n}, \tag{2.32}$$

kterým aproximujeme hodnotu směrnice charakteristiky procházející bodem $x_{j+1/2}^n$. Numerické toky pak volíme podobně jako v případě metody typu upwind pro lineární úlohu na základě znaménka $a_{j+1/2}^n$

$$F_{j+1/2}^n = \begin{cases} f(U_j^n), & \text{když } a_{j+1/2}^n \geqq 0, \\ f(U_{j+1}^n), & \text{když } a_{j+1/2}^n < 0. \end{cases} \tag{2.33}$$

Takto zvolený numerický má potom následující podobu

$$F_{j+1/2}^n = \frac{1}{2}\left[f(U_j^n) + f(U_{j+1}^n)\right] - \frac{1}{2}\left|a_{j+1/2}^n\right|\left(U_{j+1}^n - U_j^n\right). \tag{2.34}$$

Dosazením těchto hodnot numerických toků (2.34) do metody (2.29) získáme metodu typu upwind pro nelineární skalární úlohy.

Lze sestavit i jiný tvar metody, který je obdobný tvaru (2.28) použitého pro lineární úlohu. Zavedeme-li na rozhraní všech buněk

$$a^{n,+}_{j+1/2} = \max\{a^n_{j+1/2}, 0\}, \qquad a^{n,-}_{j+1/2} = \min\{a^n_{j+1/2}, 0\}, \tag{2.35}$$

získáme metodu ve tvaru

$$U^{n+1}_j = U^n_j - \frac{\Delta t}{\Delta x}\left[a^{n,+}_{j-1/2}(U^n_j - U^n_{j-1}) + a^{n,-}_{j+1/2}(U^n_{j+1} - U^n_j)\right]. \tag{2.36}$$

Je vhodné poznamenat, že rovnost (2.31 lze například aproximovat také takto

$$U^{n+1}_j = U^n_j - \frac{\Delta t}{\Delta x} f'(U^n_j)(U^n_{j+1} - U^n_j). \tag{2.37}$$

Dále si však ukážeme, že aproximace pomocí numerických toků (2.34) je mnohem vhodnější.

2.1.5 Metody pro soustavy rovnic

Metody pro řešení skalárních úloh, které jsme si odvodili na začátku této kapitoly, lze s jistými úpravami použít i na řešení soustav hyperbolických rovnic, které jsme popsali v části 1.7. Tvary některých metod pro lineární úlohu (2.47) si zde uvedeme bez odvození, která jsou analogická ke skalárnímu případu.

- Laxova-Friedrichsova metoda

$$\boldsymbol{U}^{n+1}_j = \frac{1}{2}(\boldsymbol{U}^n_{j-1} + \boldsymbol{U}^n_{j+1}) - \frac{\Delta t}{2\Delta x}\boldsymbol{A}(\boldsymbol{U}^n_{j+1} - \boldsymbol{U}^n_{j-1}). \tag{2.38}$$

- Laxova-Wendroffova metoda

$$\boldsymbol{U}^{n+1}_j = \boldsymbol{U}^n_j - \boldsymbol{A}\frac{\Delta t}{2\Delta x}(\boldsymbol{U}^n_{j+1} - \boldsymbol{U}^n_{j-1}) + \boldsymbol{A}^2\frac{\Delta t^2}{2\Delta x^2}(\boldsymbol{U}^n_{j+1} - 2\boldsymbol{U}^n_j + \boldsymbol{U}^n_{j-1}). \tag{2.39}$$

- McCormackova metoda

$$\boldsymbol{U}^{n+1}_j = \frac{\boldsymbol{U}^n_j + \tilde{\boldsymbol{U}}^{n+1}_j}{2} - \boldsymbol{A}\frac{\Delta t}{2\Delta x}(\tilde{\boldsymbol{U}}^{n+1}_j - \tilde{\boldsymbol{U}}^{n+1}_{j-1}), \tag{2.40}$$

  kde

$$\tilde{\boldsymbol{U}}^{n+1}_j = \boldsymbol{U}^n_j - \boldsymbol{A}\frac{\Delta t}{\Delta x}(\boldsymbol{U}^n_{j+1} - \boldsymbol{U}^n_j). \tag{2.41}$$

V nelineárním případě mají pak tyto metody tvary

- Laxova-Friedrichsova metoda

$$\boldsymbol{U}_j^{n+1} = \frac{1}{2}(\boldsymbol{U}_{j-1}^n + \boldsymbol{U}_{j+1}^n) - \frac{\Delta t}{2\Delta x}(\boldsymbol{f}(\boldsymbol{U}_{j+1}^n) - \boldsymbol{f}(\boldsymbol{U}_{j-1}^n)). \tag{2.42}$$

- Laxova-Wendroffova metoda

$$\boldsymbol{U}_j^{n+1} = \boldsymbol{U}_j^n - \frac{\Delta t}{\Delta x}(\boldsymbol{f}(\boldsymbol{U}_{j+1/2}^{n+1/2}) - \boldsymbol{f}(\boldsymbol{U}_{j-1/2}^{n+1/2})), \tag{2.43}$$

kde

$$\boldsymbol{U}_{j+1/2}^{n+1/2} = \frac{1}{2}(\boldsymbol{U}_j^n + \boldsymbol{U}_{j+1}^n) - \frac{\Delta t}{2\Delta x}(\boldsymbol{f}(\boldsymbol{U}_{j+1}^n) - \boldsymbol{f}(\boldsymbol{U}_j^n)). \tag{2.44}$$

- McCormackova metoda

$$\boldsymbol{U}_j^{n+1} = \frac{\boldsymbol{U}_j^n + \tilde{\boldsymbol{U}}_j^{n+1}}{2} - \frac{\Delta t}{2\Delta x}(\boldsymbol{f}(\tilde{\boldsymbol{U}}_j^{n+1}) - \boldsymbol{f}(\tilde{\boldsymbol{U}}_{j-1}^{n+1})), \tag{2.45}$$

kde

$$\tilde{\boldsymbol{U}}_j^{n+1} = \boldsymbol{U}_j^n - \frac{\Delta t}{\Delta x}(\boldsymbol{f}(\boldsymbol{U}_{j+1}^n) - \boldsymbol{f}(\boldsymbol{U}_j^n)). \tag{2.46}$$

Metoda typu upwind

Východiskem pro sestavení metody typu upwind pro řešení soustavy lineárních rovnic je rozklad na soustavu m nezávislých rovnic pro m nezávislých proměnných. Mějme lineární soustavu

$$\begin{array}{rcll} \boldsymbol{U}_t + \boldsymbol{A}\boldsymbol{U}_x & = & \boldsymbol{0}, & x \in \mathbb{R}, t \in (0,T), \\ \boldsymbol{U}(x,0) & = & \boldsymbol{u_0}(x), & x \in \mathbb{R}, \end{array} \tag{2.47}$$

s reálnou čtvercovou maticí $\boldsymbol{A}$ a hladkou počáteční funkcí $\boldsymbol{u_0} = \boldsymbol{u_0}(x) : \mathbb{R} \to \mathbb{R}^m$. Předpokládáme, že matice $\boldsymbol{A}$ je diagonalizovatelná, a lze ji tedy napsat jako

$$\boldsymbol{A} = \boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{R}^{-1}, \tag{2.48}$$

kde $\boldsymbol{R}$ je regulární matice, jejíž sloupce jsou tvořeny vlastními vektory $\boldsymbol{R}^m$ matice $\boldsymbol{A}$ a $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda^1, \ldots, \lambda^m)$ diagonální matice tvořena vlastními čísly matice $\boldsymbol{A}$. Soustavy tohoto typu lze rozložit na soustavy m rovnic pro m nezávislých proměnných

$$v_t^p + \lambda^p v_x^p = 0 \qquad p = 1, \ldots, m, \tag{2.49}$$

kde $\boldsymbol{U} = \boldsymbol{R}\boldsymbol{v}$. Každou z těchto rovnic lze řešit metodou typu upwind odvozenou v části 2.1.4. Ta má pro p-tou rovnici v bodě x_j tvar

$$V_j^{p,n+1} = V_j^{p,n} - \frac{\Delta t}{\Delta x}\left[\lambda^{p,+}(V_j^{p,n} - V_{j-1}^{p,n}) + \lambda^{p,-}(V_{j+1}^{p,n} - V_j^{p,n})\right], \tag{2.50}$$

kde

$$\lambda^{p,+} = \max\{\lambda^p, 0\}, \qquad \lambda^{p,-} = \min\{\lambda^p, 0\}. \tag{2.51}$$

Zavedeme-li vektor $\boldsymbol{V}_j^n$ o m složkách, které tvoří funkce $V_j^{p,n}$, můžeme tyto rovnice zapsat v maticové podobě

$$\boldsymbol{V}_j^{n+1} = \boldsymbol{V}_j^n - \frac{\Delta t}{\Delta x}\left[\boldsymbol{\Lambda}^+(\boldsymbol{V}_j^n - \boldsymbol{V}_{j-1}^n) + \boldsymbol{\Lambda}^-(\boldsymbol{V}_{j+1}^n - \boldsymbol{V}_j^n)\right], \tag{2.52}$$

kde matice $\boldsymbol{\Lambda}^+ = \operatorname{diag}(\lambda^{1,+}, \ldots, \lambda^{m,+})$ a $\boldsymbol{\Lambda}^- = \operatorname{diag}(\lambda^{1,-}, \ldots, \lambda^{m,-})$. Jelikož $\boldsymbol{u} = \boldsymbol{R}\boldsymbol{v}$, pak aproximace $\boldsymbol{V}_j^n = \boldsymbol{R}^{-1}\boldsymbol{U}_j^n$. Soustavu lze tedy přepsat do tvaru

$$\boldsymbol{R}^{-1}\boldsymbol{U}_j^{n+1} = \boldsymbol{R}^{-1}\boldsymbol{U}_j^n - \frac{\Delta t}{\Delta x}\left[\boldsymbol{\Lambda}^+(\boldsymbol{R}^{-1}\boldsymbol{U}_j^n - \boldsymbol{R}^{-1}\boldsymbol{U}_{j-1}^n) + \boldsymbol{\Lambda}^-(\boldsymbol{R}^{-1}\boldsymbol{U}_{j+1}^n - \boldsymbol{R}^{-1}\boldsymbol{U}_j^n)\right]. \tag{2.53}$$

Vynásobíme ji zleva maticí $\boldsymbol{R}$

$$\boldsymbol{R}\boldsymbol{R}^{-1}\boldsymbol{U}_j^{n+1} = \boldsymbol{R}\boldsymbol{R}^{-1}\boldsymbol{U}_j^n - \frac{\Delta t}{\Delta x}\left[\boldsymbol{R}\boldsymbol{\Lambda}^+\boldsymbol{R}^{-1}(\boldsymbol{U}_j^n - \boldsymbol{U}_{j-1}^n) + \boldsymbol{R}\boldsymbol{\Lambda}^-\boldsymbol{R}^{-1}(\boldsymbol{U}_{j+1}^n - \boldsymbol{U}_j^n)\right] \tag{2.54}$$

a tím získáme metodu typu upwind pro řešení lineární soustavy rovnic tvaru

$$\boldsymbol{U}_j^{n+1} = \boldsymbol{U}_j^n - \frac{\Delta t}{\Delta x}\left[\boldsymbol{A}^+(\boldsymbol{U}_j^n - \boldsymbol{U}_{j-1}^n) + \boldsymbol{A}^-(\boldsymbol{U}_{j+1}^n - \boldsymbol{U}_j^n)\right], \tag{2.55}$$

kde $\boldsymbol{A}^+ = \boldsymbol{R}\boldsymbol{\Lambda}^+\boldsymbol{R}^{-1}$ a $\boldsymbol{A}^- = \boldsymbol{R}\boldsymbol{\Lambda}^-\boldsymbol{R}^{-1}$.

Metodu typu upwind lze pochopitelně zkonstruovat i pro řešení nelineárních soustav rovnic. Tvar je opět maticovou obdobou metody pro řešení nelineární skalární rovnice sestavené v části 2.1.4. Nelineární úlohy mají tvar

$$\begin{aligned} \boldsymbol{U}_t + [\boldsymbol{F}(\boldsymbol{U})]_x &= \boldsymbol{0}, & x \in \mathbb{R}, t \in (0,T) \\ \boldsymbol{U}(x,0) &= \boldsymbol{u_0}(x), & x \in \mathbb{R}, \end{aligned} \tag{2.56}$$

který lze upravit na

$$\begin{aligned} \boldsymbol{U}_t + \boldsymbol{A}(\boldsymbol{U})\boldsymbol{U}_x &= \boldsymbol{0}, & x \in \mathbb{R}, t \in (0,T) \\ \boldsymbol{U}(x,0) &= \boldsymbol{u_0}(x), & x \in \mathbb{R}, \end{aligned} \tag{2.57}$$

kde $\boldsymbol{A}(\boldsymbol{U}) = \boldsymbol{F}'(\boldsymbol{U})$ je Jacobiho matice vektorové funkce $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{U})$. Tvar metody typu upwind je potom

$$\boldsymbol{U}_j^{n+1} = \boldsymbol{U}_j^n - \frac{\Delta t}{\Delta x}(\boldsymbol{F}_{j+1/2}^n - \boldsymbol{F}_{j-1/2}^n). \tag{2.58}$$

Numerické toky jsou definovány

$$\boldsymbol{F}_{j+1/2}^n = \frac{1}{2}\left[\boldsymbol{f}(\boldsymbol{U}_j^n) + \boldsymbol{f}(\boldsymbol{U}_{j+1}^n)\right] - \frac{1}{2}\left|\boldsymbol{A}_{j+1/2}^n\right|\left(\boldsymbol{U}_{j+1}^n - \boldsymbol{U}_j^n\right), \tag{2.59}$$

kde matice $\boldsymbol{A}_{j+1/2}^n$ je nějaká vhodná aproximace matice $\boldsymbol{A}(\boldsymbol{u})$ v čase t_n mezi buňkami j a $j+1$. Některé volby těchto matic budou uvedeny později.

2.1.6 Vlastnosti metod

Dosud jsme konstruovali nejrůznější numerické metody, aniž bychom se příliš zajímali o to, zda jsou vhodné pro řešení našich úloh a zda vůbec „fungují". V této části si popíšeme několik základních vlastností, které by „fungující" metody měly splňovat.

Budeme zde pracovat s pojmy konvergence, konzistence a stabilita, které si nyní přiblížíme. Pod pojmem konvergence přibližného řešení k přesnému řešení úlohy míníme takovou vlastnost, kdy získané řešení danou metodou lze libovolně zpřesňovat (tedy přibližovat k přesnému řešení) tím, že zmenšujeme velikost časového kroku. V limitním případě $\Delta t \to 0$ (a tedy i $\Delta x \to 0$, neboť stále předpokládáme, že $\frac{\Delta t}{\Delta x} = \text{konst.}$) pak získáme přesné řešení. Konzistencí metody s danou diferenciální rovnicí míníme vlastnost, kdy příslušné diferenční schéma je v limitním případě $\Delta t \to 0$ shodné s danou diferenciální rovnicí. Stabilita metody pak zajišťuje, že chyby, kterými jsou zatížena vstupní data metody, zůstávají během výpočtu omezené.

V této části uvedeme definice některých základních pojmů z teorie diferenčních metod pro řešení skalárních úloh. Připomeňme, že uvažujeme $\frac{\Delta t}{\Delta x} = \text{konst.}$ Explicitní diferenční formule metod lze potom zapsat ve tvaru

$$\boldsymbol{U}^{(n+1)} = \boldsymbol{\mathcal{H}}(\boldsymbol{U}^{(n)}), \tag{2.60}$$

kde $\boldsymbol{\mathcal{H}}$ je operátor přechodu z časové vrstvy n na vrstvu $n+1$. Lineární diferenční schéma pak lze psát ve tvaru

$$\boldsymbol{U}^{(n+1)} = \boldsymbol{H}\boldsymbol{U}^{(n)}, \tag{2.61}$$

kde $\boldsymbol{H}$ je lineární operátor . Nyní uvažujme případ hladkých řešení.

Definice 2.1. *Lokální diskretizační chybu* $\boldsymbol{L}^{(n)}$ metody (4.8) definujeme vztahem

$$\boldsymbol{L}^{(n)} = \frac{1}{\Delta t}\left[\boldsymbol{U}^{(n+1)} - \mathcal{H}\boldsymbol{U}^{(n)}\right], \tag{2.62}$$

kde $\boldsymbol{U}^{(n)}$ je posloupnost o členech $u_j^n = u(x_j, t_n)$.

Definice 2.2. *Globální diskretizační chybu* $E(x_j, t_n)$ (budeme ji značit E_j^n) metody (4.8) v bodě x_j a čase t_n definujeme vztahem

$$E_j^n = U_j^n - u(x_j, t_n). \tag{2.63}$$

Definice 2.3. Diferenční formule se nazývá *konzistentní* s příslušnou aproximovanou rovnicí, jestliže pro všechna j a n taková, že $T \geqq n\Delta t$ platí

$$L_j^n \to 0 \text{ pro } \Delta t \to 0. \tag{2.64}$$

Pro příklad dosaďme do Laxovy-Friedrichsovy metody (2.7) hodnoty přesného řešení $u(x,t)$. Lokální diskretizační chyba je pak podle definice 4.1 určena vztahem

$$L_j^n = \frac{1}{\Delta t}\left[u_j^{n+1} - \frac{1}{2}\left(u_{j-1}^n + u_{j+1}^n\right) + \frac{\Delta t}{2\Delta x} a\left(u_{j+1}^n - u_{j-1}^n\right)\right]. \tag{2.65}$$

Zajímá nás limitní případ, kdy $\Delta t \to 0$. Potom i $\Delta x \to 0$. V případě klasického řešení platí

$$\frac{1}{2}\left(u_{j-1}^n + u_{j+1}^n\right) \to u_j^n \tag{2.66}$$

a lokální diskretizační chyba je tedy

$$\lim_{\Delta t\to 0} L_j^n = \lim_{\Delta t\to 0} \frac{u_j^{n+1} - u_j^n}{\Delta t} + a \lim_{\Delta x \to 0} \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} = (u_j)_t + a(u_j)_x = 0. \tag{2.67}$$

Metoda je tedy konzistentní s rovnicí (2.2).

Definice 2.4. Největší přirozené číslo p, pro které existuje konstanta $C_L > 0$ taková, že pro všechna j a n taková, že $T \geqq n\Delta t$ platí

$$L_j^n \leqq C_L(\Delta t)^p \text{ pro } \Delta t \to 0, \tag{2.68}$$

se nazývá *řád konzistence.*

Mějme opět Laxovu-Friedrichsovu metodu (2.7). Využijeme Taylorova rozvoje druhého řádu členů u_j^{n+1}, u_{j-1}^n a u_{j+1}^n

$$\begin{aligned} u_j^{n+1} &\approx u_j^n + \Delta t u_t + \frac{(\Delta t)^2}{2} u_{tt}(\xi_1), \\ u_{j-1}^n &\approx u_j^n - \Delta x u_x + \frac{(\Delta x)^2}{2} u_{xx}(\xi_2), \\ u_{j+1}^n &\approx u_j^n + \Delta x u_x + \frac{(\Delta x)^2}{2} u_{xx}(\xi_3), \end{aligned}$$

kde ξ_1, ξ_2 a ξ_3 jsou body z příslušných okolí x_j, x_{j-1} a x_{j+1}. Tyto rozvoje dosadíme do vztahu pro lokální diskretizační chybu (2.65). Dále využijeme toho, že $u_t + au_x = 0$, a tedy $u_{tt} = a^2 u_{xx}(viz(2.12))$. Lokální diskretizační chyba má potom tvar

$$L_j^n = \frac{1}{\Delta t}\left[\frac{a^2(\Delta t)^2}{2} u_{xx}(\xi_1) - \frac{(\Delta x)^2}{4}(u_{xx}(\xi_2) + u_{xx}(\xi_3)) + \frac{(\Delta x)^2}{2}(u_{xx}(\xi_2) - u_{xx}(\xi_3))\right]. \tag{2.69}$$

Metoda je tedy prvního řádu konzistence ($p = 1$) neboť platí

$$\frac{a^2\Delta t}{2} u_{xx}(\xi_1) - \frac{(\Delta x)^2}{4\Delta t}(u_{xx}(\xi_2) + u_{xx}(\xi_3)) + \frac{(\Delta x)^2}{2\Delta t}(u_{xx}(\xi_2) - u_{xx}(\xi_3)) \leqq C_L \Delta t \tag{2.70}$$

tedy

$$\frac{a^2}{2}u_{xx}(\xi_1) - \frac{(\Delta x)^2}{4(\Delta t)^2}(u_{xx}(\xi_2) + u_{xx}(\xi_3)) + \frac{(\Delta x)^2}{2(\Delta t)^2}(u_{xx}(\xi_2) - u_{xx}(\xi_3)) \leqq C_L \quad (2.71)$$

za předpokladu, že $\frac{(\Delta x)^2}{(\Delta t)^2} = \text{konst.}$ a u_{xx} je ve všech bodech omezená.

Definice 2.5. Metoda určená formulí (4.8) je *konvergentní,* jestliže pro všechna j a n taková, že $T \geqq n\Delta t$ platí

$$E_j^n \to 0 \text{ pro } \Delta t \to 0. \quad (2.72)$$

Definice 2.6. Řekneme, že lineární metoda (4.9) je *lineárně stabilní,* je-li posloupnost $\{\|\boldsymbol{H}^n\|\}$ omezená, tzn. pro každé T existuje číslo $C_S > 0$ takové, že platí

$$\|\boldsymbol{H}^n\| \leqq C_S \quad (2.73)$$

pro všechna n taková, že $n\Delta t \leqq T$.

Poznámka 2.7. Dalo by se říci, že stabilní metoda je taková, která zabraňuje tomu, aby se chyby ve vstupních datech příliš zvětšovaly. Máme-li například přesná data $\boldsymbol{V}$ a data zatížená chybou $\tilde{\boldsymbol{V}}$, pak velikost této chyby (označme ji e) definujeme například

$$\mathrm{e} = \|\boldsymbol{V} - \tilde{\boldsymbol{V}}\|. \quad (2.74)$$

Chceme-li, aby tato chyba byla během výpočtu omezená, musí existovat $C > 0$ takové, že

$$\|\boldsymbol{H}(\boldsymbol{V}) - \boldsymbol{H}(\tilde{\boldsymbol{V}})\| \leqq C\|\boldsymbol{V} - \tilde{\boldsymbol{V}}\|. \quad (2.75)$$

Tento požadavek lze použít i obecně, tedy i pro nelineární případy. Často pokládáme $C = 1$, neboť v tomto případě se chyba nezvětšuje. V lineárním případě navíc platí

$$\|\boldsymbol{H}(\boldsymbol{V}) - \boldsymbol{H}(\tilde{\boldsymbol{V}})\| = \|\boldsymbol{H}(\boldsymbol{V} - \tilde{\boldsymbol{V}})\| \leqq \|\boldsymbol{H}\|\|\boldsymbol{V} - \tilde{\boldsymbol{V}}\|. \quad (2.76)$$

Velikost chyby tedy zůstane omezená, pokud $\|\boldsymbol{H}\| \leqq C$. Jelikož úlohu řešíme po časových krocích Δt, pro hodnotu řešení v čase $t_n = n\Delta t$ platí

$$\|\boldsymbol{V}^n - \tilde{\boldsymbol{V}}^n\| = \|\boldsymbol{H}^n(\boldsymbol{V}^n - \tilde{\boldsymbol{V}}^n)\| = \|\boldsymbol{H}\|^n\|\boldsymbol{V} - \tilde{\boldsymbol{V}}\| \leqq C^n\|\boldsymbol{V} - \tilde{\boldsymbol{V}}\|. \quad (2.77)$$

Pokud tedy $C = 1$, pak $\|\boldsymbol{H}^n\| \leqq \|\boldsymbol{H}\|^n \leqq 1$ a chyba ve vstupních datech se nebude časem zvětšovat. Lze ukázat, že je možné volit $C = 1 + \alpha\Delta t$, pro nějaké $\alpha > 0$. Při této volbě má pak konstanta C_S uvedená v definici 4.4 tvar

$$C_S = \mathrm{e}^{\alpha T}. \quad (2.78)$$

Poznámka 2.8. Použitá maticová norma je indukovaná zvolenou vektorovou normou a to tak, že pro matici $\boldsymbol{A} \in \mathbb{R}^{m,m}$ je definovaná vztahem

$$\|\boldsymbol{A}\| = \sup_{\boldsymbol{v}\neq 0} \frac{\|\boldsymbol{A}\boldsymbol{v}\|}{\|\boldsymbol{v}\|}, \qquad \forall \boldsymbol{v} \in \mathbb{R}^m. \tag{2.79}$$

Je možné použít například obvyklou Eukleidovskou normu $||\cdot||_2$ nebo normu $||\cdot||_{2,\Delta x}$ definovanou

$$||\boldsymbol{U}||_{2,\Delta x} = \sqrt{\sum_{j=-\infty}^{\infty} (u_j)^2 \Delta x}. \tag{2.80}$$

Pracovat s vektorovými a maticovými normami je možné, neboť funkce $u_0(x)$ má omezený nosič a vzhledem k tomu, že se hodnoty funkce $u(x,t)$ šíří po charakteristikách, kterými jsou přímky s kladnou směrnicí, existují $x_{\min}, x_{\max} \in \mathbb{R}$ takové, že hodnoty $u(x,t)$ jsou nulové pro $x \notin \langle x_{\min}, x_{\max}\rangle$.

O konvergenci lineární metody hovoří následující věta, jejíž důkaz je možné nalézt např. v [9].

Věta 2.9 (Laxova). Nechť lineární metoda (4.9) je konzistentní. Tato metoda je konvergentní právě tehdy, když je stabilní.

Jako příklad si vezměme metodu typu upwind

$$U_j^{n+1} = U_j^n - a\frac{\Delta t}{\Delta x}(U_{j+1}^n - U_j^n) \tag{2.81}$$

pro lineární rovnici (2.2), kde je konstanta $a < 0$. Ukažme si pro tuto metodu odvození podmínky stability. Z (2.81) plyne

$$\sum_{j=-\infty}^{\infty} |U_j^{n+1}|^2 = \sum_{j=-\infty}^{\infty} |(1 + a\frac{\Delta t}{\Delta x})U_j^n - a\frac{\Delta t}{\Delta x}U_{j+1}^n|^2 \leqq$$

$$\leqq \sum_{j=-\infty}^{\infty} \left[(1 + a\frac{\Delta t}{\Delta x})^2|U_j^n|^2 + 2|1 + a\frac{\Delta t}{\Delta x}||a\frac{\Delta t}{\Delta x}||U_j^n||U_{j+1}^n| + |a\frac{\Delta t}{\Delta x}|^2|U_{j+1}^n|^2\right] \leqq$$

$$\leqq \sum_{j=-\infty}^{\infty} \left[(1 + a\frac{\Delta t}{\Delta x})^2|U_j^n|^2 + |1 + a\frac{\Delta t}{\Delta x}||a\frac{\Delta t}{\Delta x}|\left(|U_j^n|^2 + |U_{j+1}^n|^2\right) + |a\frac{\Delta t}{\Delta x}|^2|U_{j+1}^n|^2\right] \leqq$$

$$\leqq \sum_{j=-\infty}^{\infty} \left[(1 + a\frac{\Delta t}{\Delta x})^2 + 2|1 + a\frac{\Delta t}{\Delta x}||a\frac{\Delta t}{\Delta x}| + |a\frac{\Delta t}{\Delta x}|^2\right] |U_j^n|^2.$$

Po odmocnění a přenásobení diskretizačním krokem Δx tedy platí

$$||U^{n+1}||_{2,\Delta x} \leqq \left[|1 + a\frac{\Delta t}{\Delta x}| + |a\frac{\Delta t}{\Delta x}|\right] ||U^n||_{2,\Delta x},$$

$$||U^{n+1}||_{2,\Delta x} \leqq \left[|1 + a\frac{\Delta t}{\Delta x}| + |a\frac{\Delta t}{\Delta x}|\right]^{n+1} ||U^0||_{2,\Delta x}.$$

Pro

$$-1 \leqq a\frac{\Delta t}{\Delta x} \leqq 0, \qquad a < 0 \tag{2.82}$$

platí

$$|1 + a\frac{\Delta t}{\Delta x}| + |a\frac{\Delta t}{\Delta x}| = 1 + a\frac{\Delta t}{\Delta x} - a\frac{\Delta t}{\Delta x} = 1.$$

Metoda je tedy stabilní za předpokladů (2.82).

Obdobnou podmínku stability získáme i pro Laxovu-Friedrichsovu metodu, která má pro lineární skalární úlohu (2.2) tvar

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{\Delta t}{2\Delta x}a(U_{j+1}^n - U_{j-1}^n). \tag{2.83}$$

Pro názornost použijeme jiného postupu než v případě metody typu upwind. Laxovu-Friedrichsovu metodu pro řešení na celém intervalu přepíšeme do tvaru $\boldsymbol{U}^{n+1} = \boldsymbol{HU}^n$

$$\begin{bmatrix} U_1^{n+1} \\ U_2^{n+1} \\ U_3^{n+1} \\ \vdots \\ U_{N-1}^{n+1} \\ U_N^{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ c_L & 0 & c_R & 0 & 0 & \dots & 0 \\ 0 & c_L & 0 & c_R & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \dots & 0 & 0 & c_L & 0 & c_R \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^n \\ U_2^n \\ U_3^n \\ \vdots \\ U_{N-1}^n \\ U_N^n \end{bmatrix}, \tag{2.84}$$

kde $c_L = \frac{1}{2} + \frac{a\Delta t}{2\Delta x}$ a $c_R = \frac{1}{2} - \frac{a\Delta t}{2\Delta x}$. Na krajích intervalu jsou hodnoty položeny rovny nule. Zvolená maticová norma $\|.\|_1$, která je rovna maximální hodnotě sloupcového součtu $\|\boldsymbol{H}\|_1 = \max_j \sum_{i=1}^{N} |h_{ij}|$, je

$$\|\boldsymbol{H}\|_1 = \left|\frac{1}{2} + \frac{a\Delta t}{2\Delta x}\right| + \left|\frac{1}{2} - \frac{a\Delta t}{2\Delta x}\right| \leqq 1 \Leftrightarrow \left|\frac{a\Delta t}{2\Delta x}\right| \leqq \frac{1}{2}. \tag{2.85}$$

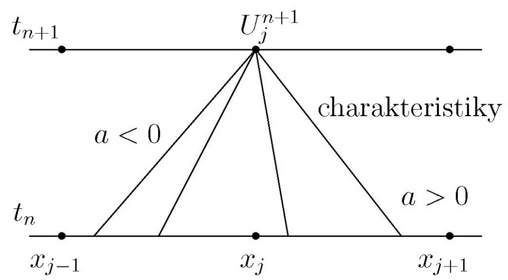
Metoda je tedy stabilní za podmínky

$$|a|\frac{\Delta t}{\Delta x} \leqq 1. \tag{2.86}$$

Poznámka 2.10. Pro konkrétní metodu a úlohu (2.2) nebo (2.1) často vyjadřujeme důsledek podmínky stability (4.16) ve tvaru

$$\frac{\Delta t}{\Delta x} \leqq \text{konst.} \tag{2.87}$$

Hovoříme pak o *Courantově-Friedrichsově-Lewyho (CFL) podmínce.* Pro názornost si ukažme geometrickou interpretaci CFL podmínky na lineární rovnici $u_t + au_x = 0$.



Obr. 2.3 Geometrická interpretace CFL podmínky

Ta má tvar (2.86). Řešení je v tomto případě konstantní na charakteristikách, které mají směrnici $\frac{1}{a}$. CFL Podmínka pak zaručuje přenos řešení z místa mezi „správnými“ uzly (viz Obr.2.3).

V tabulce 2.1 jsou uvedena schémata některých metod a jim odpovídající podmínky stability.

Název metody	Diferenční schéma	pro	Podmínka stability
jednostranná L	$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} a(U_j^n - U_{j-1}^n)$	$a>0$	$a\frac{\Delta t}{\Delta x} \leqq 1$
jednostranná R	$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} a(U_{j+1}^n - U_{j-1}^n)$	$a<0$	$-a\frac{\Delta t}{\Delta x} \leqq 1$
Laxova-Friedrichsova	$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{\Delta t}{2\Delta x} a(U_{j+1}^n - U_{j-1}^n)$	$a \in \mathbb{R}$	$\|a\|\frac{\Delta t}{\Delta x} \leqq 1$
Laxova-Wendroffova	$U_j^{n+1} = U_j^n - \frac{\Delta t}{2\Delta x} a(U_{j+1}^n - U_{j-1}^n) + \frac{\Delta t^2}{2\Delta x^2} a^2(U_{j+1}^n - 2U_j^n + U_{j-1}^n)$	$a \in \mathbb{R}$	$\|a\|\frac{\Delta t}{\Delta x} \leqq 1$

Tab. 2.1 Podmínky stability pro rovnici $u_t + au_x = 0$

V případě nelineární úlohy (2.1) jsou charakteristiky přímky, pro něž platí $\frac{\mathrm{d}x}{\mathrm{d}t} = f'(u)$, nejsou tedy rovnoběžné. Hodnota konkrétní směrnice je závislá na funkci u. Délku časového kroku je tedy třeba volit s ohledem na hodnotu řešení na celém intervalu, proto ji vyjadřujeme ve tvaru

$$\max_j |f'(U_j^n)| \frac{\Delta t}{\Delta x} \leqq 1. \tag{2.88}$$

Vezměme si opět Laxovu-Friedrichsovu metodu a ukažme, že při splnění podmínky (2.88) je stabilní. Mějme dvě řešení nelineární úlohy (2.1) V a $\tilde{V}$, tedy platí

$$V_j^{n+1} = \frac{1}{2}(V_{j-1}^n + V_{j+1}^n) - \frac{\Delta t}{2\Delta x}(f(V_{j+1}^n) - f(V_{j-1}^n)) \tag{2.89}$$

a

$$\tilde{V}_j^{n+1} = \frac{1}{2}(\tilde{V}_{j-1}^n + \tilde{V}_{j+1}^n) - \frac{\Delta t}{2\Delta x}(f(\tilde{V}_{j+1}^n) - f(\tilde{V}_{j-1}^n)). \tag{2.90}$$

Zaveďme si $W_j^n = V_j^n - \tilde{V}_j^n$ pro rozdíl obou řešení. Odečtením vztahu (2.90) od vztahu (2.89) získáme

$$W_j^{n+1} = \frac{1}{2}(W_{j-1}^n + W_{j+1}^n) - \frac{\Delta t}{2\Delta x}\left[\left(f(V_{j+1}^n) - f(\tilde{V}_{j+1}^n)\right) - \left(f(V_{j-1}^n) - f(\tilde{V}_{j-1}^n)\right)\right]. \tag{2.91}$$

Je-li f hladká konvexní funkce, potom pro nějaké $\xi_j^n \in \langle V_j^n, \tilde{V}_j^n\rangle$ platí

$$f(V_j^n) - f(\tilde{V}_j^n) = f'(\xi_j^n)(V_j^n - \tilde{V}_j^n) = f'(\xi_j^n)W_j^n. \tag{2.92}$$

Využitím tohoto vztahu můžeme rovnost (2.91) přepsat do tvaru

$$W_j^{n+1} = \left(\frac{1}{2} - \frac{\Delta t}{2\Delta x}f'(\xi_{j+1}^n)\right)W_{j+1}^n + \left(\frac{1}{2} + \frac{\Delta t}{2\Delta x}f'(\xi_{j-1}^n)\right)W_{j-1}^n. \tag{2.93}$$

Při předpokládaném splnění podmínky (2.88) (připomeňme, že funkce $f(u)$ je konvexní, tedy funkce $f'(u)$ je monotónní, a proto splnění podmínky (2.88) v uzlových bodech x_j zaručuje, že $\frac{\Delta t}{2\Delta x}f'(\xi_{j+1}^n) \leqq \frac{1}{2}$) jsou obě závorky v (2.93) nezáporné. Proto platí

$$|W_j^{n+1}| \leqq \left(\frac{1}{2} - \frac{\Delta t}{2\Delta x}f'(\xi_{j+1}^n)\right)|W_{j+1}^n| + \left(\frac{1}{2} + \frac{\Delta t}{2\Delta x}f'(\xi_{j-1}^n)\right)|W_{j-1}^n|. \tag{2.94}$$

Sečteme-li tento vztah přes všechny uzly j, dostáváme

$$\begin{aligned}\sum_{j=-\infty}^{\infty}|W_j^{n+1}| \leqq \frac{1}{2}\sum_{j=-\infty}^{\infty}|W_{j+1}^n| + \frac{1}{2}\sum_{j=-\infty}^{\infty}|W_{j-1}^n| - \frac{\Delta t}{2\Delta x}\sum_{j=-\infty}^{\infty}f'(\xi_{j+1}^n)|W_{j+1}^n| \quad + \\ + \frac{\Delta t}{2\Delta x}\sum_{j=-\infty}^{\infty}f'(\xi_{j-1}^n)|W_{j-1}^n|.\end{aligned} \tag{2.95}$$

Vzhledem k tomu, že v krajních uzlech předpokládáme nulové hodnoty neznámých funkcí, poslední dva součty vztahu (2.95) se vzájemně vynulují a docházíme ke vztahu

$$\sum_{j=-\infty}^{\infty}|W_j^{n+1}| \leqq \sum_{j=-\infty}^{\infty}|W_j^n| \qquad \text{neboli} \qquad \|\boldsymbol{V}^{n+1} - \tilde{\boldsymbol{V}}^{n+1}\|_1 \leqq \|\boldsymbol{V}^n - \tilde{\boldsymbol{V}}^n\|_1. \tag{2.96}$$

Laxova-Friedrichsova metoda je tedy stabilní.

Při řešení soustav lineárních rovnic vychází odvození podmínky stability z rozkladu na soustavu nezávislých rovnic (2.49). Řešení jednotlivých rovnic se šíří po charakteristikách o směrnicích $\frac{1}{\lambda^p}$. Pro stabilitu celé soustavy je tedy nutné zvolit časový krok, který je stabilní pro každou z rovnic. Docházíme tak k podmínce

$$\max_p |\lambda^p| \frac{\Delta t}{\Delta x} \leqq 1. \tag{2.97}$$

V nelineárním případě vyházíme ze soustavy (2.57) a na základě analogie k (2.97 vyjadřujeme tuto podmínku tvarem

$$\max_p |\lambda^p(u)| \frac{\Delta t}{\Delta x} \leqq 1, \tag{2.98}$$

kde $\lambda^p(u)$ jsou vlastní čísla matice $\boldsymbol{A}(\boldsymbol{u})$.

Dalším pojmem, který si zde vysvětlíme je konzervativita. Ta znamená, že veškeré změny celkové bilance neznámé funkce jsou v dané oblasti způsobeny pouze rozdílem vstupního a výstupního toku. V dané oblasti se tedy nic neztrácí ani nic nevzniká (stejně jako je tomu u popisovaného fyzikálního děje). Řešení našich úloh (2.1) a (2.2) tuto vlastnost splňují. Proto je vhodné na řešení těchto úloh použít konzervativní metody.

Definice 2.11. Řekneme, že metoda je *konzervativní*, jestliže pro aproximaci neznámé funkce platí

$$\sum_{j=-\infty}^{\infty} U_j^{n+1} = \sum_{j=-\infty}^{\infty} U_j^n. \tag{2.99}$$

Předpokládáme, že $\lim_{j \to -\infty} U_j = 0$ a $\lim_{j \to \infty} U_j = 0$.

Ukažme si nyní, že metody, které lze zapsat v divergentním tvaru jsou konzervativní. Pro řešení v bodě x_j platí

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}(F_{j+1/2}^n - F_{j-1/2}^n). \tag{2.100}$$

Sečtením těchto rovností přes všechny uzly s indexy j získáme

$$\sum_{j=-\infty}^{\infty} U_j^{n+1} = \sum_{j=-\infty}^{\infty} U_j^n - \frac{\Delta t}{\Delta x} \sum_{j=-\infty}^{\infty} (F_{j+1/2}^n - F_{j-1/2}^n). \tag{2.101}$$

Protože $F_{j+1/2}^n = F_{(j+1)-1/2}^n$, toky uvnitř intervalu se vzájemně odečtou a pro dostatečně velké indexy $|j|$ předpokládáme nulové hodnoty neznámé funkce a tedy i toků. Celková bilance neznámé funkce tak zůstane zachována

$$\sum_{j=-\infty}^{\infty} U_j^{n+1} = \sum_{j=-\infty}^{\infty} U_j^n. \tag{2.102}$$

Metody v divergentním tvaru se proto často označují jako metody v konzervativním tvaru.

Pro řešení úloh, ve kterých se vyskytují veličiny, které z fyzikálního hlediska nemohou nabývat záporných hodnot (např. hloubka vody), je vhodné použití metod, které zaručí nezápornost daných aproximací. Takové vlastnosti se říká pozitivní semidefinitnost.

Definice 2.12. Řekneme, že daná metoda je *pozitivně semidefinitní*, jestliže platí

$$(\boldsymbol{U}^n \geqq \mathbf{0}) \Longrightarrow (\boldsymbol{U}^{n+1} = \boldsymbol{H}(\boldsymbol{U}^n) \geqq \mathbf{0}). \tag{2.103}$$

Jako příklad si vezměme Laxovu-Friedrichsovu metodu a ukažme, že je konzervativní a pozitivně semidefinitní. Řešení v čase t_{n+1} v bodě x_j má hodnotu

$$U_j^{n+1} = \frac{1}{2}\left(U_{j-1}^n + U_{j+1}^n\right) - a\frac{\Delta t}{2\Delta x}\left(U_{j+1}^n - U_{j-1}^n\right). \tag{2.104}$$

Konzervativitu ukážeme tak, že rovnosti (2.104) sečteme přes všechny uzly

$$\sum_{j=-\infty}^{\infty} U_j^{n+1} = \frac{1}{2}\sum_{j=-\infty}^{\infty}\left(U_{j-1}^n + U_{j+1}^n\right) - a\frac{\Delta t}{2\Delta x}\sum_{j=-\infty}^{\infty}\left(U_{j+1}^n - U_{j-1}^n\right). \tag{2.105}$$

Opět je zřejmé, že za předpokladu $\lim\limits_{j\to-\infty} U_j = 0$ a $\lim\limits_{j\to\infty} U_j = 0$ platí

$$\frac{1}{2}\sum_{j=-\infty}^{\infty}\left(U_{j-1}^n + U_{j+1}^n\right) = \sum_{j=-\infty}^{\infty} U_j^n, \qquad \sum_{j=-\infty}^{\infty}\left(U_{j+1}^n - U_{j-1}^n\right) = 0 \tag{2.106}$$

a podmínka konzervativity $\sum\limits_{j=-\infty}^{\infty} U_j^{n+1} = \sum\limits_{j=-\infty}^{\infty} U_j^n$ je splněna.

Pro prokázání pozitivní semidefinitnosti předpokládáme, že $U_{j-1}^n \geqq 0$, $U_{j+1}^n \geqq 0$ a platí CFL podmínka $|a|\frac{\Delta t}{\Delta x} \leqq 1$. Metodu lze snadno přepsat do tvaru

$$U_j^{n+1} = U_{j-1}^n\left(\frac{1}{2} + a\frac{\Delta t}{\Delta x}\right) + U_{j+1}^n\left(\frac{1}{2} - a\frac{\Delta t}{\Delta x}\right). \tag{2.107}$$

Pokud je $a > 0$ potom

$$\begin{aligned} \frac{1}{2} + a\frac{\Delta t}{\Delta x} &> \frac{1}{2}, \\ \frac{1}{2} - a\frac{\Delta t}{\Delta x} &\geqq \frac{1}{2} - \frac{1}{2} = 0. \end{aligned} \tag{2.108}$$

Pokud je $a < 0$ potom

$$\begin{aligned} \frac{1}{2} + a\frac{\Delta t}{\Delta x} &\geqq \frac{1}{2} - \frac{1}{2} = 0, \\ \frac{1}{2} - a\frac{\Delta t}{\Delta x} &> \frac{1}{2}. \end{aligned} \tag{2.109}$$

V obou případech jsou obě závorky ve vztahu (2.107) nezáporné, a proto platí, že $U_j^{n+1} \geqq 0$. Laxova-Friedrichsova metoda je tedy pozitivně semidefinitní.

2.1.7 Lokální Laxova-Friedrichsova metoda

Centrální metody jsou, jak uvidíme, obecně zatíženy velkou numerickou difúzí. Difúze je jev, který se projevuje například při šíření tepla, kdy se teplo šíří z míst s větší teplotou do míst s menší teplotou. Obecně dochází k vyrovnávání rozdílů v hodnotách sledovaných veličin. Přesné řešení našich úloh však žádnou difúzi neobsahuje, a je tedy vhodné tento nežádoucí jev potlačit. Jako příklad si opět vezměme Laxovu-Friedrichsovu metodu pro nelineární skalární rovnici (2.9), kterou lze zapsat také v divergentním tvaru

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}\left[F_{j+1/2}^n - F_{j-1/2}^n\right], \tag{2.110}$$

kde

$$F_{j+1/2}^n = \frac{1}{2}\left[f(U_{j+1}^n) + f(U_j^n)\right] - a\frac{1}{2}(U_{j+1}^n - U_j^n), \qquad a = \frac{\Delta x}{\Delta t}. \tag{2.111}$$

Přestože je tato metoda konstruována pro řešení zákona zachování $u_t + [f(u)]_x = 0$, lze ukázat, že přesněji aproximuje bilanční vztah

$$u_t + [f(u)]_x = Du_{xx}, \tag{2.112}$$

kde D je difúzní koeficient. Přepíšeme-li Laxovu-Friedrichsovu metodu s numerickým tokem (2.111), získáme

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + \frac{f(U_{j+1}^n) - f(U_{j-1}^n)}{2\Delta x} = a\frac{\Delta x}{2}\frac{U_{j+1}^n - 2U_j^n + U_{j+1}^n}{(\Delta x)^2}. \tag{2.113}$$

V této aproximaci tak difúznímu koeficientu odpovídá člen $D \approx \frac{a}{2}\Delta x = \frac{(\Delta x)^2}{2\Delta t}$. Abychom numerickou difúzi snížili, zavedeme lokálně závislý parametr

$$a_{j+1/2}^n = \max(|f'(u)|), \qquad \forall u \in \langle U_j^n, U_{j+1}^n\rangle. \tag{2.114}$$

V případě konvexní funkce f je

$$a_{j+1/2}^n = \max\{|f'(U_j^n)|, |f'(U_{j+1}^n)|\}. \tag{2.115}$$

Takto zavedený parametr skutečně snižuje numerickou difúzi metody, neboť při splnění nutné podmínky stability $|f'(u)|\frac{\Delta t}{\Delta x} \leqq 1$ pro všechny hodnoty funkce u platí

$$|f'(u)| \leqq \frac{\Delta x}{\Delta t} = a. \tag{2.116}$$

Pro lokální Laxovu-Friedrichsovu metodu v divergentním tvaru se tedy používají numerické toky ve tvaru

$$F_{j+1/2}^n = \frac{1}{2}\left[f(U_{j+1}^n) - f(U_{j-1}^n) - a_{j+1/2}^n(U_{j+1}^n - U_j^n)\right]. \tag{2.117}$$

Poznámka 2.13. Všimněme si ještě, že takto zvolený numerický tok je jakýmsi přechodem mezi numerickými toky centrální metody a metodou typu upwind popsané v části 2.1.4, u které je vliv numerické difúze nejmenší. To je zřejmé z tvaru tokové funkce, který u metody typu upwind je

$$F^n_{j+1/2} = \frac{1}{2}\left[f(U^n_j) + f(U^n_{j+1})\right] - \frac{1}{2}\left|a^n_{j+1/2}\right|\left(U^n_{j+1} - U^n_j\right), \tag{2.118}$$

kde

$$a^n_{j+1/2} = \frac{f(U^n_{j+1}) - f(U^n_j)}{U^n_{j+1} - U^n_j}. \tag{2.119}$$

Numerická difúze zde nemůže být větší než u lokální Laxovy-Friedrichsovy metody, kde je parametr $a^n_{j+1/2}$ volen jako maximální hodnota podílu (2.119) na daném intervalu.

Poznámka 2.14. Abychom se zcela zbavili vlivu numerické difúze, použili bychom numerický tok (2.111) s volbou $a = 0$. Taková metoda je však nestabilní a ukazuje se, že určité množství numerické difúze je potřebné k zajištění stability metody.

V případě nelineární soustavy diferenciálních rovnic o m rovnicích je třeba podobně jako u metody typu upwind využít ke konstrukci lokální Laxovy-Friedrichsovy metody aproximace Jacobiho matice soustavy a jejích vlastních čísel a vlastních vektorů. Numerický tok je definován

$$\boldsymbol{F}^n_{j+1/2} = \frac{1}{2}\left[\boldsymbol{f}(\boldsymbol{U}^n_j) + \boldsymbol{f}(\boldsymbol{U}^n_{j+1})\right] - \frac{1}{2}\sum_{p=1}^{m} a^{n,p}_{j+1/2}\alpha^p_{j+1/2}\boldsymbol{r}^p_{j+1/2}, \tag{2.120}$$

kde je využito rozkladu skoku v neznámé funkci

$$\boldsymbol{U}^n_{j+1} - \boldsymbol{U}^n_j = \sum_{r=1}^{m} \alpha^p_{j+1/2}\boldsymbol{r}^p_{j+1/2} \tag{2.121}$$

a

$$a^{n,p}_{j+1/2} = \max\{|\lambda^{n,p}_j|, |\lambda^{n,p}_{j+1}|\}. \tag{2.122}$$

Další variantou metody je použití numerického toku (2.118) s volbou

$$a^n_{j+1/2} = \max\{\max_p\{|\lambda^{n,p}_j|, |\lambda^{n,p}_{j+1}|\}\}, \tag{2.123}$$

jejíž obdoba bude podrobněji popsána při konstrukci metody typu central-upwind. Podrobnosti o těchto rozkladech je možné nalézt například v [9].

2.2 Úlohy s pravou stranou

Dosud jsme se v této kapitole věnovali homogenním úlohám, tedy takovým, které mají nulovou pravou stranou. Nyní si stručně popišme způsoby, jakými lze modifikovat použité metody na nehomogenní úlohy. Jako příklad si vezměme lineární skalární úlohu se zdrojovou funkcí ψ

$$\begin{aligned} u_t + au_x &= \psi(u,x), \qquad & t \in (0,T), x \in \mathbb{R}, a \in \mathbb{R}, T > 0, \\ u(x,0) &= u_0(x), & x \in \mathbb{R}, \end{aligned} \tag{2.124}$$

kde $u = u(x,t)$, $f \in C^2(\mathbb{R})$ a $\psi = \psi(u,x)$. Na tuto úlohu lze aplikovat všechny popsané metody, do kterých je pouze nutno přidat nějakou vhodnou aproximaci zdrojového členu Ψ. Pokud bychom chtěli použít například Laxovu-Friedrichsovu metodu, měla by tvar

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{\Delta t}{2\Delta x} a(U_{j+1}^n - U_{j-1}^n) + \Delta t \Psi(U_j^n, x_j^n). \tag{2.125}$$

Metodu typu upwind bychom mohli zkonstruovat například takto

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}\left[a^+(U_j^n - U_{j-1}^n) + a^-(U_{j+1}^n - U_j^n)\right] + \Delta t \Psi(U_j^n, x_j^n), \tag{2.126}$$

kde $a^+ = \max\{a, 0\}, a^- = \min\{a, 0\}$. Ukazuje se, že u této metody je lepší využít upwindingu i v aproximaci pravé strany. Schéma pak může mít následující podobu

$$\begin{gathered} U_j^{n+1} = U_j^n - \tfrac{\Delta t}{\Delta x}\left[a^+(U_j^n - U_{j-1}^n) + a^-(U_{j+1}^n - U_j^n)\right] + \\ + \tfrac{\Delta t}{2}\left(\tfrac{a^+}{a}\left(\Psi(U_{j-1}^n, x_{j-1}) + \Psi(U_j^n, x_j)\right) + \tfrac{a^-}{a}\left(\Psi(U_j^n, x_j) + \Psi(U_{j+1}^n, x_{j+1})\right)\right). \end{gathered} \tag{2.127}$$

Poznámka 2.15. Někdy je vhodné převést danou nehomogenní úlohu na úlohu homogenní v kvazilineárním tvaru. K tomu je třeba rozšířit počet rovnic o další, kde se mezi neznámými objeví některé funkce obsažené ve zdrojovém členu. Tento postup má své výhody i nevýhody, je vhodný například pro úlohy, kdy hledáme ustálené stavy. Konstrukce tohoto způsobu řešení je nad rámec těchto skript, pro Saint-Venantovy rovnice jej čtenář nalezne například v [15].

Další možností řešení nehomogenních soustav je použití metody štěpení. Tato metoda má především tu výhodu, že lze beze změn využít všechny metody k vyřešení homogenního problému a pravou stranu uvažovat zvlášť. Na každé časové vrstvě se řeší dvě úlohy.

1. Homogenní rovnice $\boldsymbol{u}_t + [\boldsymbol{f}(\boldsymbol{u})]_x = \boldsymbol{0}$, jejíž řešení je použito jako počáteční podmínka v druhém kroku

2. Obyčejná diferenciální rovnice $\boldsymbol{u}_t = \boldsymbol{\psi}(\boldsymbol{u}, x)$.

Tato metoda je populární především pro svou jednoduchost a širokou možnost použití. Další výhodou je to, že každý krok lze řešit naprosto odlišnou metodou. Navíc díky tomu, že podproblémy představují jednodušší úlohy než původní problém (homogenní úloha a obyčejná diferenciální rovnice), existuje větší množství metod na jejich řešení.

Mohou zde však nastat problémy především v situacích, kdy se blížíme k ustálenému stavu, ve kterém $[\boldsymbol{f}(\boldsymbol{u})]_x = \boldsymbol{\psi}(\boldsymbol{u}, x)$. V takových případech je tato metoda nevhodná, neboť produkuje značné výchylky od ustáleného stavu a je tedy obtížné jej numericky určit. Dalším problémem jsou okrajové podmínky, které jsou dané pro původní úlohu se zdrojovými členy. Není vždy jednoduché na jejich základě stanovit, jaké okrajové podmínky zvolit pro obě nově sestavené úlohy. Problémy mohou být také se stabilitou celé metody štěpení, přestože metody použité na oba kroky budou stabilní, jejich kombinování může přinést nestabilitu. Podrobnosti o této metodě je možné nalézt například v [9].

Pojmy k zapamatování

— konzistence
— konzervativita
— stabilita
— centrální metody
— metody typu upwind
— lokální diskretizační chyba
— globální diskretizační chyba
— diskretizace úlohy
— CFL podmínka
— pozitivní semidefinitnost

Kontrolní otázky

1. Jak se nazývá nutná podmínka stability explicitních metod a jak ji lze geometricky interpretovat?
2. Stačí konzistence a konvergence metody k její stabilitě?
3. Jaký je rozdíl mezi lokální a globální diskretizační chybou?
4. Je metoda v divergentním tvaru konzervativní?
5. Co je to konzistence metody s diferenciální rovnicí?
6. Jaké jsou hlavní rozdíly mezi centrálními metodami a metodami typu upwind?
7. Co je to pozitivní semidefinitnost metody?

Příklady k procvičení

!

1. Stanovte řád konzistence pro Laxovu-Wendroffovu metodu.

2. Dokažte konzistenci Laxovy-Wendroffovy metody s lineární skalární rovnicí (2.2).

3. Ukažte, že metoda tvaru (2.28) je konzervativní.

4. Ukažte, že Laxovu-Friedrichsovu metodu tvaru (2.9) lze zapsat v divergentním tvaru (2.110).

5. Odvoďte rozklad soustavy lineárních rovnic $\boldsymbol{u}_t + \boldsymbol{A}\boldsymbol{u}_x = \boldsymbol{0}$ na soustavu nezávislých rovnic.

6. Napište tvar Laxovy-Wendroffovy metody a metody typu upwind pro řešení Burgersovy rovnice $u_t + \left(\frac{1}{2}u^2\right)_x = 0$.

7. Sestavte metodu pro řešení lineární skalární úlohy, pro níž bude mít CFL podmínka stability podobu
$$|a|\frac{\Delta t}{\Delta x} \leqq 10.$$

Klíč k příkladům k procvičení

1. Konzistence druhého řádu.

2. Postupujte analogicky jako při důkazu konzistence Laxovy-Friedrichsovy metody v části (2.1.6).

3. Pro konzervativitu je třeba splnit rovnost
$\sum\limits_{j=-\infty}^{\infty} U_j^n + \frac{\Delta t}{\Delta x} \sum\limits_{j=-\infty}^{\infty} \left[a^+(U_j^n - U_{j-1}^n) + a^-(U_{j+1}^n - U_j^n)\right] = \sum\limits_{j=-\infty}^{\infty} U_j^n.$
Druhá suma je nulová, neboť se členy vzájemně odečtou a při $j \to \pm\infty$ je $U_j^n = 0$.

4. Numerický tok je v tomto případě definován
$$F_{j+1/2}^n = \frac{1}{2}(f(U_j^n) + f(U_{j+1}^n)) - \frac{\Delta x}{2\Delta t}(U_{j+1}^n - U_j^n)$$

5. Využijte rozkladu (2.48) $\boldsymbol{A} = \boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{R}^{-1}$ a zaveďte vektor $\boldsymbol{v} = \boldsymbol{R}^{-1}\boldsymbol{u}$. S tímto přepíšete soustavu na tvar
$$(\boldsymbol{R}\boldsymbol{v})_t + \boldsymbol{R}\boldsymbol{\Lambda}\mathbb{R}^{-1}(\boldsymbol{R}\boldsymbol{v})_x = \boldsymbol{0}$$
$$\boldsymbol{R}\boldsymbol{v}_t + \boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{v}_x = \boldsymbol{0}$$
$$\boldsymbol{v}_t + \boldsymbol{\Lambda}\boldsymbol{v}_x = \boldsymbol{0}.$$
Získáte tak soustavu nezávislých rovnic $(v^p)_t + \lambda^p(v^p)_x = 0$.

6. Dosazením do tvaru metody pro řešení nelineární skalární rovnice (2.20) získáte Laxovu-Wendroffovu metodu
$U_j^{n+1} = U_j^n - \frac{\Delta t}{2\Delta x}\left((U_{j+1/2}^{n+1/2})^2 - (U_{j-1/2}^{n+1/2})^2\right),$

kde $U_{j+1/2}^{n+1/2} = \frac{1}{2}\left(U_j^n - U_{j+1}^n\right) - \frac{\Delta t}{4\Delta x}\left((U_{j+1}^n)^2 - (U_j^n)^2\right)$. Dosazením do (2.29) získáte metodu typu upwind
$U_j^{n+1} = U_j^n - \frac{\Delta t}{4\Delta x}\Big[(U_{j+1}^n)^2 - |U_{j+1}^n + U_j^n|(U_{j+1}^n - U_j^n) - (U_{j-1}^n)^2 + |U_j^n + U_{j-1}^n|(U_j^n - U_{j-1}^n)\Big]$.

7. Jedním z řešení je sestavení obdoby Laxovy-Friedrichsovy metody ve tvaru

$$U_j^{n+1} = \frac{1}{2}(U_{j-k}^n + U_{j+k}^n) - \frac{\Delta t}{2k\Delta x}a(U_{j+k}^n - U_{j-k}^n), \quad k \in \mathbb{N}.$$

Analogickým postupem jako v případě (2.84) získáte podmínku stability

$$\left|\frac{1}{2} + \frac{a\Delta t}{2k\Delta x}\right| + \left|\frac{1}{2} - \frac{a\Delta t}{2k\Delta x}\right| \leqq 1 \Leftrightarrow \left|\frac{a\Delta t}{2k\Delta x}\right| \leqq \frac{1}{2}.$$

Tato metoda je tedy stabilní za podmínky

$$|a|\frac{\Delta t}{\Delta x} \leqq k.$$

Zvolíte $k = 10$.

Kapitola 3

Metody pro úlohy s nehladkým řešením

Průvodce studiem

V této kapitole se čtenář seznámí s rozdíly mezi metodami prvního a vyššího řádu. Naučí se, jakými způsoby lze konstruovat metody vyššího řádu přesnosti řešení použitím korekčních toků nebo polynomiální rekonstrukce neznámé funkce. Pozná nepříznivé jevy, které metody vyššího řádu vnášejí do řešení a naučí se některým postupům, jak je omezit. Na základě znalosti vlastností jednotlivých metod pronikne do problematiky výběru vhodné metody pro konkrétní problém.

Cíle

Po prostudování této kapitoly budete schopni:

- sestavit metody vyššího řádu
- sestavit metody s vysokým rozlišením
- definovat důležité pojmy jako metody s vysokým rozlišením, ustálený stav, apod.

Ve druhé kapitole jsme věnovali pozornost diferenčním metodám pro parciální diferenciální rovnice hyperbolického typu. Zásadní otázkou je, zda jsou tyto metody použitelné pro úlohy s nehladkými daty. I v případě, že by tyto metody použitelné byly, zůstává otázkou, zda je možné pro jejich analýzu použít teoretické nástroje vyvinuté pro studium metody konečných diferencí. Již na první pohled je zřejmé, že nelze použít zavedené pojmy konzistence a řádu metody. Ty jsou totiž založeny na předpokladu diferencovatelnosti přesného řešení.

Dalším problémem je, že je pro potřeby analýzy nelineárních úloh a nelineárních metod rozpracovat teorii nelineární stability. Naše předchozí úvahy byly totiž založené na pojmu stability lineární.

Cílem této kapitoly je jednak odvodit metody, jednak vybudovat odpovídající teoretický aparát vhodný pro analýzu obecných nelineárních problémů.

3.1 Nehladká data

Uvažujme skalární lineární úlohu

$$\begin{aligned} u_t + au_x &= 0, \qquad t \in (0,T), x \in \mathbb{R}, a \in \mathbb{R}, T > 0, \\ u(x,0) &= u_0(x), \qquad x \in \mathbb{R}, \end{aligned} \tag{3.1}$$

a připomeňme si tvar přesného řešení této úlohy. Řešení je konstantní na charakteristikách (podrobněji popsány v části 1.2, což jsou v našem případě přímky v rovině (x,t), které splňují vztah

$$\frac{\mathrm{d}x}{\mathrm{d}t} = a. \tag{3.2}$$

Přesné řešení na těchto charakteristikách má tvar (viz část 1.2)

$$u(x,t) = u_0(x - at). \tag{3.3}$$

Řešení v bodě x_j a čase t_n je tedy jednoznačně určeno počáteční hodnotou neznámé funkce v bodě $x_j - at_n$. V případě hladké funkce u_0 se jedná o klasické řešení, lze však definovat i zobecněné řešení – nespojitost v počáteční podmínce se opět šíří po příslušné charakteristice.

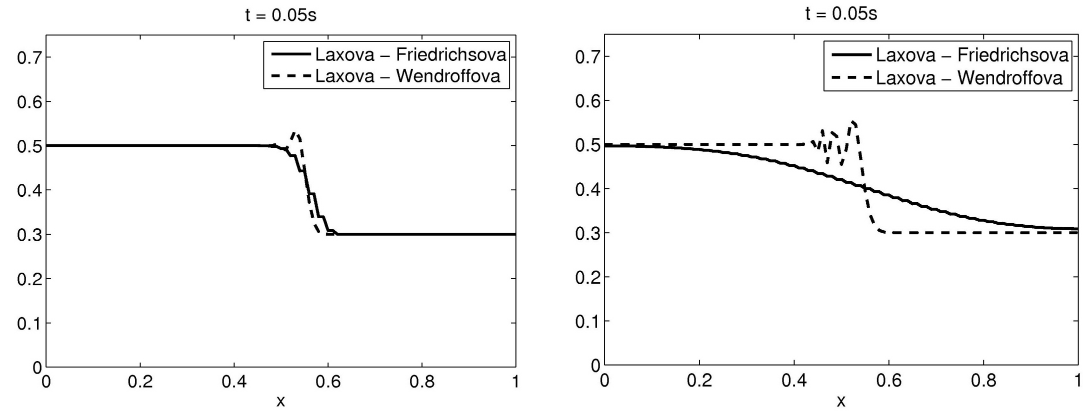
Nejdříve se podívejme na diferenční metody odvozené v předchozích kapitolách.

Ukažme si na příkladu chování metody prvního řádu (Laxova-Friedrichsova metoda) a metody druhého řádu (Laxova-Wendroffova metoda) při dvou různých volbách velikosti časové diskretizace. Uvažujme úlohu

$$\begin{aligned} u_t + u_x &= 0, \qquad t > 0, x \in \langle 0,1 \rangle, \\ u(x,0) &= \begin{cases} 0,5 & \text{pro } x \leqq 0,5, \\ 0,3 & \text{pro } x > 0,5, \end{cases} \\ u(0,t) = 0,5, &\qquad u(1,t) = 0,3. \end{aligned} \tag{3.4}$$

Obrázky 3.1 ukazují řešení v čase $T = 0,05s$ s použitím obou metod a s příslušnými časovými kroky Δt v závislosti na velikosti prostorového kroku Δx (zvoleno $\Delta x =$ $= 0,01$). Na obrázku 3.1 vidíme, že u metody druhého řádu dochází ke vzniku oscilací v místě, kde je řešení nehladké, zatímco u metody prvního řádu je zachována monotonie řešení, přibližné řešení je však příliš „rozmazáno“.

Později ukážeme, že metody prvního řádu zachovávají monotonii řešení, tedy neprodukují další lokální extrémy oproti přesnému řešení. Nevýhodou je pochopitelně samotný nízký řád přesnosti a již zmíněný efekt „rozmazávání numerického řešení v čase“. Metody druhého řádu se sice vyznačují vyšší přesností v místech, kde je řešení hladké, v případě nespojitého řešení však tyto metody produkují umělé oscilace, které přesné řešení neobsahuje.



$\Delta t = \Delta x/2$ $\Delta t = (\Delta x)^2$

Obr. 3.1 Porovnání řešení získaného metodou prvního a druhého řádu

Lze obecně říci, že diferenční metody, které jsme odvodili v předchozí kapitole nejsou vhodné pro řešení úloh, ve kterých se vyskytují nehladká data. Připomeňme ještě, že v případě nelineárních úloh můžeme obdržet nehladké řešení i pro vstupní data, která jsou hladká. V dalším textu popíšeme metody, které netrpí výše popsanými nedostatky diferenčních metod.

3.2 Modifikace diferenčních metod

3.2.1 Metoda umělé vazkosti

Jedná se o jednu z prvních nelineárních metod, která byla vyvinuta pro řešení problémů dynamiky tekutin. Byla vyvinuta J. von Neumannem a R. Richtmyerem v roce 1950. Popíšeme zde stručně zjednodušenou variantu pro advekční rovnici s konstantním koeficientem. Místo řešení parciální diferenciální rovnice

$$u_t + au_x = 0, \quad a > 0, \tag{3.5}$$

budeme hledat přibližné řešení rovnice

$$u_t + au_x = du_{xx}. \tag{3.6}$$

Přesné řešení této rovnice je pro $d > 0$ hladké. Proto lze použít diferenční metodu (např. druhého řádu). Otázkou zůstává, jak zvolit parametr d. Pokud bychom provedli volbu d =konst. a chtěli použít explicitní metodu, dostali bychom jako podmínku stability místo CFL nerovnosti nerovnost obvyklou u explicitních metod pro parciální diferenciální rovnice parabolického typu. Tak bychom ale obdrželi velmi přísný požadavek na omezení časového kroku. Proto je nutné d volit jako funkci Δx

a Δt. Fakticky se výše popsaný postup realizuje pomocí metody štěpení: nejdříve se použije metoda druhého řádu pro rovnici (3.5) v konzervativním tvaru, např.

$$\begin{aligned} \tilde{U}_j^{n+1} &= U_j^n - \tfrac{\Delta t}{\Delta x}(F_{j+1/2}^n - F_{j-1/2}^n), \\ F_{j+1/2}^n &= \tfrac{1}{2}\left[f(U_j^n) + f(U_{j+1}^n)\right] = \tfrac{1}{2}a\left(U_j^n + U_{j+1}^n\right). \end{aligned} \tag{3.7}$$

Poté se realizuje korekční krok, např. ve tvaru

$$U_j^{n+1} = \tilde{U}_j^n + a\frac{\Delta t}{2\Delta x}\left(1 - a\frac{\Delta t}{\Delta x}\right)\left(U_{j+1}^n - 2U_j^n + U_{j-1}^n\right). \tag{3.8}$$

Tento krok je aproximací vztahu

$$u_t = a\frac{\Delta x}{2}\left(1 - a\frac{\Delta t}{\Delta x}\right)u_{xx}. \tag{3.9}$$

Z (3.5) a (3.9) plyne, že aproximujeme úlohu

$$u_t + au_x = a\frac{\Delta x}{2}\left(1 - a\frac{\Delta t}{\Delta x}\right)u_{xx}. \tag{3.10}$$

Nevýhodou tohoto postupu je, že získaná metoda je prvního řádu. Konkrétně jde o metodu ve tvaru

$$U_j^{n+1} = U_j^n - a\frac{\Delta t}{\Delta x}\left(U_j^n - U_{j-1}^n\right). \tag{3.11}$$

Naším cílem je však navrhnout metodu, která je vyššího řádu pro problémy, jejichž řešení je hladké. V tomto případě musíme přidat vazké členy vyššího řádu. Původní vazký člen nižšího řádu musí záviset na datech U_j^n, aby bylo možné rozlišit případy, ve kterých je řešení hladké, tj., aby v bodech, ve kterých je řešení nehladké, platilo $d = a\frac{\Delta x}{2}\left(1 - a\frac{\Delta t}{\Delta x}\right)$ a vazký člen člen by byl roven nule, a v bodech, ve kterých je řešení hladké, musí být vazké členy voleny tak, abychom získali aproximaci řádu vyššího.

Popišme ještě semidiskrétní aproximaci. Ta má tvar

$$\begin{aligned} \frac{\mathrm{d}\tilde{U}_j}{\mathrm{d}t} &= -\tfrac{1}{\Delta x}(F_{j+1/2} - F_{j-1/2}) + \tfrac{|a|}{2\Delta x}\left(U_{j+1} - 2U_j + U_{j-1}\right), \\ F_{j+1/2} &= \tfrac{1}{2}\left[f(U_j) + f(U_{j+1})\right] = \tfrac{1}{2}a\left(U_j + U_{j+1}\right). \end{aligned} \tag{3.12}$$

Z (3.5) a (3.12) plyne, že aproximujeme úlohu

$$u_t + au_x = |a|\frac{\Delta x}{2}u_{xx}. \tag{3.13}$$

V nelineárním případě

$$u_t + [f(u)]_x = 0$$

má semidiskrétní aproximace tvar

$$\begin{aligned}
\frac{\mathrm{d}\tilde{U}_j}{\mathrm{d}t} &= -\frac{1}{\Delta x}\left[(F_{j+1/2} - F_{j-1/2}) + \frac{a_{j+1/2}}{2\Delta x}(U_{j+1} - U_j) - \frac{a_{j-1/2}}{2\Delta x}(U_j - U_{j-1})\right], \\
F_{j+1/2} &= \frac{1}{2}\left[f(U_j) + f(U_{j+1})\right],
\end{aligned} \tag{3.14}$$

kde

$$a_{j+1/2} = \max\{|f'(U_j)|, |f'(U_{j+1})|\}$$

Tato volba vychází z kvazilineárního tvaru nelineární rovnice, tj. tvaru

$$u_t + f'(u)u_x = 0.$$

Roli koeficientu a z rovnice $u_t + au_x = 0$ hraje tedy první derivace tokové funkce. Výhodou metod tohoto typu je relativně jednoduchá implementace. Navíc lze analogickou myšlenku použít i pro soustavy parciálních diferenciálních rovnic hyperbolického typu a úlohy ve více prostorových dimenzích. Nevýhodou je, že člen s druhou prostorovou derivací, který odpovídá difúznímu (příp. vazkému) členu obsahuje koeficient, který je často volen na základě zkušenosti. Podrobnější rozbor metod tohoto typu lze nalézt např. v publikaci [23].

3.2.2 Hartenovo TVNI schéma

V roce 1982 publikoval A. Harten metodu, která je založena na kombinaci schématu prvního a druhého řádu. Uvažujme opět rovnici

$$u_t + [f(u)]_x = 0 \tag{3.15}$$

a pro jednoduchost předpokládejme, že $f'(u) > 0$ pro všechna relevantní u. Metoda prvního řádu typu upwind má lokální diskretizační chybu ve tvaru

$$L(\Delta t, \Delta x, u) = \frac{\Delta x}{2}\left\{f'(u)\left[\frac{\Delta t}{\Delta x}f'(u) - 1\right]u_x\right\}_x + O(\Delta x^2). \tag{3.16}$$

Víme – zatím pouze na základě zkušenosti, že v bodech, ve kterých je řešení nehladké je vhodné použít metodu prvního řádu, a v bodech, kde je řešení hladké, je však možné použít metodu řádu druhého (příp. vyššího). Proto je ve druhém případě vhodné první člen chyby eliminovat. Eliminaci této chyby lze chápat jako druhý krok metody štěpení uplatněné na rovnici

$$u_t + \left[f(u) + \frac{\Delta x}{\Delta t}g\right]_x = 0, \tag{3.17}$$

kde

$$g = \Delta x \beta\left(u, \frac{\Delta t}{\Delta x}\right)u_x, \quad \beta\left(u, \frac{\Delta t}{\Delta x}\right) = \frac{1}{2}[\mu - \mu^2], \quad \mu = \frac{\Delta t}{\Delta x}f'(u). \tag{3.18}$$

Prvním krokem metody štěpení je aplikace metody prvního řádu na rovnici

$$u_t + [f(u)]_x = 0.$$

Na problém lze nahlížet i jinak: uplatníme-li na rovnici (3.17) vhodným způsobem schéma prvního řádu, dostaneme adaptivní schéma pro původní rovnici (3.15), tj.získáme postup, založený na přepínání mezi schématem prvního a druhého řádu. Ukazuje se, že takto navržená metoda má vlastnost TVNI (total variation non--increasing; nyní je tato vlastnost častěji označována jako TVD – total variation diminishing). Tato vlastnost je ve shodě s vlastností přesného řešení, to jest s tím, že v čase neroste totální variace řešení, kde

$$\mathrm{TV}_t(v) = \sup \sum_{j=1}^{M} |v(\xi_j) - v(\xi_{j-1})|$$

přes všechna dělení $-\infty < \xi_0 < \xi_1 < \cdots < \xi_M < \infty$.

Právě popsanou metodu v roce 1973 modifikovali A. Harten a G. Zwas. Nazvali ji hybridní metodou. Na podobné myšlence je založena také metoda J. P. Borise a D. L. Booka publikovaná v témže roce a označovaná jako metoda korekce transportního toku (FCT – Flux Corrected Transport). Konstrukce metody je založena na přidání tzv. korekčních toků k metodě prvního řádu. Korekční toky v bodech, kde je řešení hladké, způsobí, že se zvýší řád přesnosti, zatímco v bodech, kde je řešení nehladké, jejich vliv na řešení bude malý, a metoda tak zůstane pouze prvního řádu a nebude vytvářet nežádoucí oscilace.

Popišme nyní Hartenův a Zwasův přístup. Korekční toky označme $F^{n,H}_{j+1/2}$. Tyto korekční toky přidáme k metodě prvního řádu. Uvažujme případ skalárního nelineárního problému $u_t + [f(u)]_x = 0$. Metoda má následující tvar

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}\left[F^n_{j+1/2} - F^n_{j-1/2}\right] - \frac{\Delta t}{\Delta x}(F^{n,H}_{j+1/2} - F^{n,H}_{j-1/2}), \tag{3.19}$$

kde

$$F^n_{j+1/2} = \frac{1}{2}[f(U_j^n) + f(U_{j+1}^n)] - \frac{1}{2}|a^n_{j+1/2}|(U^n_{j+1} - U^n_j), \tag{3.20}$$

$$a^n_{j+1/2} = \frac{f(U^n_{j+1}) - f(U^n_j)}{U^n_{j+1} - U^n_j}, \tag{3.21}$$

je tok prvního řádu a

$$F^{n,H}_{j+1/2} = \frac{1}{2}|a^n_{j+1/2}|\left(1 - \frac{\Delta t}{\Delta x}|a^n_{j+1/2}|\right)(U^n_{j+1} - U^n_j)\Psi(\Theta^n_{j+1/2}), \tag{3.22}$$

$$\Theta^n_{j+1/2} = \begin{cases} \frac{\Upsilon^n_{j-1/2}}{\Upsilon^n_{j+1/2}} & \text{pro } a^n_{j+1/2} > 0, \\ \frac{\Upsilon^n_{j+3/2}}{\Upsilon^n_{j+1/2}} & \text{jinak,} \end{cases} \tag{3.23}$$

$$\Upsilon^n_{j+1/2} = U^n_{j+1} - U^n_j.$$

Funkce $\Psi = \Psi(\Theta)$ se nazývá omezovací funkce (limiter) a je mírou lokální hladkosti řešení. Je možné ji definovat několika způsoby, pro příklad si zde uveďme dva.

- $\Psi(\Theta) = \frac{\Theta+|\Theta|}{1+|\Theta|}$,
- $\Psi(\Theta) = \mathrm{minmod}(1, \Theta)$,

kde

$$\mathrm{minmod}(a, b) = \begin{cases} a & \text{pokud } |a| < |b| \wedge ab > 0, \\ b & \text{pokud } |b| < |a| \wedge ab > 0, \\ 0 \text{ pokud } ab \leqq 0. \end{cases} \tag{3.24}$$

Pokud je $\Psi(\Theta^n_{j+1/2}) = 1$, jedná se o metodu druhého řádu. Konkrétně získáme Laxovu-Wendroffovu metodu (2.18). Pokud je $\Psi(\Theta^n_{j+1/2}) < 1$, dochází ke snížení řádu přesnosti a zároveň k omezení oscilací numerického řešení. Toto nastává, pokud je v místě $x_{j+1/2}$ nespojitost ve funkčních hodnotách U^n, pak výraz $U^n_{j+1} - U^n_j$ nabývá vysokých hodnot. Tento výraz vystupuje jako jmenovatel ve výrazu (3.23). Pomocí funkce $\Psi(\Theta^n_{j+1/2})$ se tak snižuje vliv korekčních toků na řešení. Funkční hodnoty $\Psi(\Theta^n_{j+1/2}) \approx 1$ v případě, že v místě $x_{j+1/2}$ mají funkce $\boldsymbol{U}^n$ hladký charakter a metoda je tedy druhého řádu přesnosti. $\Psi(\Theta^n_{j+1/2}) = 0$ tehdy, když $\Theta^n_{j+1/2} < 0$ tj. jednostranné diference mají v bodě $x_{j+1/2}$ opačné znaménko. V bodě $x_{j+1/2}$ se v takovém případě nachází lokální extrém a metoda je prvního řádu přesnosti.

3.3 Metoda konečných objemů

Metoda konečných objemů nevychází z lokální formulace pomocí diferenciální rovnosti, ale z integrální bilance ve tvaru

$$\int_{x_1}^{x_2} u(x, t_2)\,\mathrm{d}x = \int_{x_1}^{x_2} u(x, t_1)\,\mathrm{d}x - \int_{t_1}^{t_2} f[u(x_2, t)]\,\mathrm{d}t + \int_{t_1}^{t_2} f[u(x_1, t)]\,\mathrm{d}t, \tag{3.25}$$

popř. bilance

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{x_1}^{x_2} u(x, t)\,\mathrm{d}x = -f[u(x_2, t)] + f[u(x_1, t)]. \tag{3.26}$$

Tyto rovnosti lze využít k formulaci zobecněného řešení tak, že vyžadujeme splnění rovnosti pro každý konečný interval $(x_1, x_2) \subset \mathbb{R}$ a případně současně pro každý interval $(t_1, t_2) \subset \langle 0, T\rangle$. V případě metody konečných objemů zvolíme diskretizační síť stejně jako u metody konečných diferencí, tj. $x_j = j\Delta x$ a $t_n = n\Delta t$, kde $j \in \mathbb{Z}$

a $n \in \mathbb{N}_0$. Dále zavádíme $x_{j+1/2} = x_j + \Delta x/2$. Dalším krokem při konstrukci metody konečných objemů je formulace integrální rovnosti (3.25) na intervalech daných diskretizací, tj.

$$\begin{aligned} \int\limits_{x_{j-1/2}}^{x_{j+1/2}} u(x,t_{n+1})\,\mathrm{d}x &= \int\limits_{x_{j-1/2}}^{x_{j+1/2}} u(x,t_n)\,\mathrm{d}x - \\ &- \int\limits_{t_n}^{t_{n+1}} f[u(x_{j+1/2},t)]\,\mathrm{d}t + \int\limits_{t_n}^{t_{n+1}} f[u(x_{j-1/2},t)]\,\mathrm{d}t, \end{aligned} \tag{3.27}$$

Tuto rovnost lze přepsat do následujícího tvaru

$$\bar{u}_j^{n+1} = \bar{u}_j^n - \frac{\Delta t}{\Delta x}(\bar{f}_{j+1/2}^{n+1/2} - \bar{f}_{j-1/2}^{n+1/2}), \tag{3.28}$$

kde

$$\bar{u}_j^n = \frac{1}{\Delta x}\int\limits_{x_{j-1/2}}^{x_{j+1/2}} u(x,t_n)\ \mathrm{d}x \tag{3.29}$$

a

$$\bar{f}_{j+1/2}^{n+1/2} = \frac{1}{\Delta t}\int\limits_{t_n}^{t_{n+1}} f(u(x_{j+1/2},t))\ \mathrm{d}t. \tag{3.30}$$

Dostali jsme tak vztah mezi integrálními průměry. Tento vztah lze chápat jako numerické schéma, pomocí kterého na základě znalosti integrálních průměrů hledané funkce na n-té časové vrstvě stanovíme hodnoty integrálních průměrů na na $n+1$ časové vrstvě.

Základním problémem je, že pro stanovení těchto integrálních průměrů potřebujeme znát hodnoty hledané funkce $u(x_{j+1/2}, t)$ pro $t \in (t_n, t_{n+1})$. Abychom tyto hodnoty mohli určit, musíme realizovat dva kroky. Tyto kroky již nevyužívají pouze údaje z integrální formulace, proto právě navrhovaným postupem nezískáme přesné hodnoty integrálních průměrů $\bar{u}_j^n$, ale jejich přibližné hodnoty, které označíme $\bar{U}_j^n$. Nejdříve z integrálních průměrů na n-té časové vrstvě přibližně zrekonstruujeme průběh funkce $u = u(x, t_n)$. Tuto rekonstrukci označíme $\hat{U}^n = \hat{U}^n(x)$. Nejjednodušším způsobem je realizovat po částech konstantní rekonstrukci, tj.

$$\hat{U}^n(x) = \bar{U}_j^n \text{ pro } x \in (x_{j-1/2}, x_{j+1/2}).$$

Dalším krokem je stanovení (přibližné) hodnoty integrálů (3.30). K jejich stanovení však musíme – jak jsme již zmínili – znát přibližné hodnoty funkce$u(x_{j+1/2}, t)$ v intervalu (t_n, t_{n+1}). Aproximaci této funkce označíme $\tilde{U}^n = \tilde{U}^n(x, t)$. Hledat ji budeme jako řešení Riemannova problému

$$\begin{aligned} u_t + [f(u)]_x &= 0, x \in R, t \in (t_n, t_{n+1}), \\ u(x, t_n) &= \begin{cases} \hat{U}^n(x_{j+1/2}-), & x < x_{j+1/2}, \\ \hat{U}^n(x_{j+1/2}+), & x > x_{j+1/2}. \end{cases} \end{aligned} \tag{3.31}$$

Problém (3.32) můžeme řešit přesně nebo přibližně. Řešení tohoto problému však není hladká funkce. Lze však ukázat, že na přímce $x = x_{j+1/2}$ řešení hladké je. Pro dostatečně malé Δt navíc toto řešení není ovlivněno řešením dalších Riemannových problémů v sousedních bodech $x_{j-1/2}$ a $x_{j+3/2}$. Poté již můžeme (opět přesně nebo přibližně) stanovit hodnoty integrálu (3.30), které označíme $\bar{F}^{n+1/2}_{j+1/2}$. Dostáváme tak rekurentní formuli ve tvaru

$$\bar{U}_j^{n+1} = \bar{U}_j^n - \frac{\Delta t}{\Delta x}(\bar{F}^{n+1/2}_{j+1/2} - \bar{F}^{n+1/2}_{j-1/2}). \tag{3.32}$$

Vzhledem k tomu, že nepotřebujeme znát celý průběh funkcí $\tilde{U}^n = \tilde{U}^n(x,t)$, ale pouze hodnoty v bodech $x = x_{j+1/2}$, stačí řešit Riemannův problém ve zjednodušené podobě. Na základě již dříve popsaného řešení Riemannova problému lze učinit následující závěry:

- Je-li $f'(\bar{U}_j^n) > 0$ a $f'(\bar{U}_{j+1}^n) > 0$, pak nezávisle na tom, zda je řešením rázová vlna, nebo vlna zředění dostáváme $\tilde{U}^n(x_{j+1/2}, t) = \bar{U}_j^n$, a tedy

$$\bar{F}^{n+1/2}_{j+1/2} = \frac{1}{\Delta t}\int\limits_{t_n}^{t_{n+1}} f(\bar{U}_j^n)\,\mathrm{d}t = \frac{1}{\Delta t} f(\bar{U}_j^n)\int\limits_{t_n}^{t_{n+1}} \mathrm{d}t = f(\bar{U}_j^n).$$

- Je-li $f'(\bar{U}_j^n) < 0$ a $f'(\bar{U}_{j+1}^n) < 0$, pak nezávisle na tom, zda je řešením rázová vlna, nebo vlna zředění dostáváme $\tilde{U}^n(x_{j+1/2}, t) = \bar{U}_{j+1}^n$, a tedy

$$\bar{F}^{n+1/2}_{j+1/2} = \frac{1}{\Delta t}\int\limits_{t_n}^{t_{n+1}} f(\bar{U}_{j+1}^n)\,\mathrm{d}t = \frac{1}{\Delta t} f(\bar{U}_{j+1}^n)\int\limits_{t_n}^{t_{n+1}} \mathrm{d}t = f(\bar{U}_{j+1}^n).$$

- Je-li $f'(\bar{U}_j^n) > 0$ a $f'(\bar{U}_{j+1}^n) < 0$, dostáváme rázovou vlnu, jejíž rychlost šíření je dána

$$s^n_{j+1/2} = \frac{f(\bar{U}_{j+1}^n) - f(\bar{U}_j^n)}{\bar{U}_{j+1}^n - \bar{U}_j^n}.$$

  Pokud $s^n_{j+1/2} < 0$, pak

$$\bar{F}^{n+1/2}_{j+1/2} = \frac{1}{\Delta t}\int\limits_{t_n}^{t_{n+1}} f(\bar{U}_{j+1}^n)\,\mathrm{d}t = \frac{1}{\Delta t} f(\bar{U}_{j+1}^n)\int\limits_{t_n}^{t_{n+1}} \mathrm{d}t = f(\bar{U}_{j+1}^n).$$

  Pokud $s^n_{j+1/2} > 0$, pak

$$\bar{F}^{n+1/2}_{j+1/2} = \frac{1}{\Delta t}\int\limits_{t_n}^{t_{n+1}} f(\bar{U}_j^n)\,\mathrm{d}t = \frac{1}{\Delta t} f(\bar{U}_j^n)\int\limits_{t_n}^{t_{n+1}} \mathrm{d}t = f(\bar{U}_j^n).$$

- Je-li $f'(\bar{U}_j^n) < 0$ a $f'(\bar{U}_{j+1}^n) > 0$, dostáváme vlnu zředění, pro kterou platí $\tilde{U}^n(x_{j+1/2}, t) = q(0)$, kde $q = q(x/t)$ je řešením rovnice $f'(q(x/t)) = 0$. Dostáváme tak

$$\bar{F}_{j+1/2}^{n+1/2} = \frac{1}{\Delta t} \int\limits_{t_n}^{t_{n+1}} f(q(0))\,\mathrm{d}t = \frac{1}{\Delta t} f(q(0)) \int\limits_{t_n}^{t_{n+1}} \mathrm{d}t = f(q(0)).$$

Právě vyložený postup patří do skupiny Godunovových metod, tj. metod, které využívají konečně-objemového přístupu kombinovaného s formulací Riemannových problémů. Na tomto místě výkladu je vhodné dát do souvislosti diferenční metodu typu upwind prvního řádu, kterou lze zapsat ve tvaru

$$\begin{aligned} U_j^{n+1} &= U_j^n - \tfrac{\Delta t}{\Delta x}\left(F_{j+1/2}^n - F_{j-1/2}^n\right), \\ F_{j+1/2}^n &= \tfrac{1}{2}[f(U_j^n) + f(U_{j+1}^n)] - \tfrac{1}{2}|a_{j+1/2}^n|(U_{j+1}^n - U_j^n), \\ a_{j+1/2}^n &= \tfrac{f(U_{j+1}^n) - f(U_j^n)}{U_{j+1}^n - U_j^n} \end{aligned} \tag{3.33}$$

a právě vyloženou metodu: všimněme si, že pokud nahradíme řešení případu s centrální vlnou zředění (poslední varianta v právě uvedeném výčtu) rázovou vlnou (předposlední varianta), tj. interpretujeme všechny vlny jako rázové, získáme totožné metody, kde $U_j^n = \bar{U}_j^n$, $F_{j+1/2}^n = \bar{F}_{j+1/2}^{n+1/2}$, $a_{j+1/2}^n = s_{j+1/2}^n$. Jinak řečeno diferenční metoda typu upwind je totožná s Godunovovou metodou s přibližným Riemannovým řešičem typu all-shock, tj. s řešičem, ve kterém všechny vlny považujeme za rázové. Obě dvě metody jsou formálně vzato metodami prvního řádu. V jedné jsou však použity aproximace bodových hodnot, ve druhé metodě používáme aproximace integrálních průměrů (je vhodné uvědomit si, jakou jsme použili rekonstrukci, a obecně také to, že rozdíl mezi integrálním průměrem a bodovou hodnotou je $O(\Delta x^2)$).Také si povšimněme, že podstatným prvkem obou přístupů je přesná linearizace

$$f(U_{j+1}^n) = f(U_j^n) + a_{j+1/2}^n (U_{j+1}^n - U_j^n).$$

Především v případě nehladkých dat není vhodné používat Taylorovskou linearizaci

$$f(U_{j+1}^n) = f(U_j^n) + f'(U_j^n)(U_{j+1}^n - U_j^n).$$

Kromě právě popsaného způsobu lze použít následující modifikaci, v rámci které se vyhneme řešení Riemannových problémů. Tentokrát vyjdeme z integrální bilance

$$\begin{aligned} \int\limits_{x_j}^{x_{j+1}} u(x, t_{n+1})\,\mathrm{d}x \;&=\; \int\limits_{x_j}^{x_{j+1}} u(x, t_n)\,\mathrm{d}x - \\ &- \int\limits_{t_n}^{t_{n+1}} f[u(x_{j+1}, t)]\,\mathrm{d}t + \int\limits_{t_n}^{t_{n+1}} f[u(x_j, t)]\,\mathrm{d}t, \end{aligned} \tag{3.34}$$

Tuto rovnost lze přepsat do následujícího tvaru

$$\bar{u}^{n+1}_{j+1/2} = \bar{u}^{n}_{j+1/2} - \frac{\Delta t}{\Delta x}(\bar{f}^{n+1/2}_{j+1} - \bar{f}^{n+1/2}_{j}), \tag{3.35}$$

kde

$$\bar{u}^{n}_{j+1/2} = \frac{1}{\Delta x}\int\limits_{x_j}^{x_{j+1}} u(x,t_n)\,\mathrm{d}x \tag{3.36}$$

a

$$\bar{f}^{n+1/2}_{j+1} = \frac{1}{\Delta t}\int\limits_{t_n}^{t_{n+1}} f(u(x_{j+1},t))\,\mathrm{d}t. \tag{3.37}$$

Zvolíme stejný typ rekonstrukce funkcí

$$\hat{U}^n(x) = \bar{U}^n_j \text{ pro } x \in (x_{j-1/2}, x_{j+1/2}).$$

Pak platí

$$\bar{u}^{n}_{j+1/2} \approx \bar{U}^{n}_{j+1/2} = \frac{\bar{U}^n_j + \bar{U}^n_{j+1}}{2}$$

Dalším krokem je stanovení (přibližné) hodnoty integrálů (3.30). K jejich stanovení však musíme – jak jsme již zmínili – znát přibližné hodnoty funkce $u(x_{j+1},t)$. Tuto hodnotu stanovíme jednoduše, protože pro dostatečně malé Δt není třeba řešit Riemannův problém – řešení úlohy je v bodě $x_{j+1/2}$ hladké (v našem případě konstantní).

$$u(x_{j+1},t) \approx \tilde{U}^n(x_{j+1},t) = \bar{U}^n_{j+1}$$

Velice jednoduše také stanovíme integrál

$$\bar{F}^{n+1/2}_{j+1} = \frac{1}{\Delta t}\int\limits_{t_n}^{t_{n+1}} f(\bar{U}^n_{j+1})\,\mathrm{d}t = \frac{1}{\Delta t} f(\bar{U}^n_{j+1})\int\limits_{t_n}^{t_{n+1}} \mathrm{d}t = f(\bar{U}^n_{j+1}).$$

Pak lze již použít rekurentní formuli

$$\bar{U}^{n+1}_{j+1/2} = \bar{U}^n_j - \frac{\Delta t}{\Delta x}(\bar{F}^{n+1/2}_{j+1} - \bar{F}^{n+1/2}_{j}). \tag{3.38}$$

Máme tedy metodu ve tvaru

$$\bar{U}^{n+1}_{j+1/2} = \frac{\bar{U}^n_j + \bar{U}^n_{j+1}}{2} - \frac{\Delta t}{\Delta x}[f(\bar{U}^n_{j+1}) - f(\bar{U}^n_j)]. \tag{3.39}$$

Dostali jsme tak metodu Godunovova typu, která je formálně totožná s Laxovou-Friedrichsovou metodou. Dostali jsme tak nejjednodušší centrální metodu. Do této skupiny zařazujeme metody, které ve velmi omezené míře využívají informace o šíření vln. Přesněji řečeno jde o metody, u kterých jsou kontrolní objemy voleny tak, že řešení počáteční úlohy s danou rekonstrukcí je v rámci časového kroku $\langle t_n, t_{n+1}\rangle$ nehladké pouze uvnitř těchto objemů. Nehladké části řešení tedy nemají vliv na příslušné integrální bilance.

V případě metod Godunovova typu je nutnou podmínkou konvergence opět Courantova-Friedrichsova-Lewyho podmínka. Je ji však třeba interpretovat jinak. Pokud platí

$$\max_u |f'(u)| \frac{\Delta t}{\Delta x} \leqq \frac{1}{2},$$

nedojde v rámci časového kroku ke kolizi vln v řešení sousedících Riemannových problémů. V tomto případě lze hodnotu integrálního průměru na nové časové vrtsvě stanovit také integrací řešení příslušných Riemannových problémů přes interval $\langle x_{j-1/2}, x_{j+1/2}\rangle$ v případě (3.32) a přes interval $\langle x_j, x_{j+1}\rangle$ v případě (3.32). Pokud platí

$$\max_u |f'(u)| \frac{\Delta t}{\Delta x} \leqq 1,$$

dojde sice k interakci vln v řešení sousedících Riemannových problémů, ale v případě první probírané metody Godunovova typu lze ukázat, že řešení nového Riemannova problému (a to i v případě, že použijeme obecnější rekonstrukci než po částech konstantní), neovlivní hodnoty řešení v bodech $x_{j+1/2}$.

Jak jsme již zmínili, právě popsané metody jsou prvního řádu. Metody, které jsou vyššího řádu lze získat různými způsoby. Nejdříve se ještě zastavme u toho, co vlastně míníme metodou, kterého je řádu vyššího než prvního. Poznamenejme, že některé z těchto poznámek lze vztáhnout i k již vyloženým modifikacím diferenčních metod:

- V případě zobecněných řešení pracujeme s nehladkými daty. Nelze tedy realizovat některé úvahy založené na Taylorově rozvoji. Obecně si lze samozřejmě klást otázku, jakého řádu je metoda, nelze však použít standardní přístupy užívané při analýze metody konečných diferencí. Pokud pro analýzu chyby metody použijeme formálně Taylorův rozvoj, hovoříme o formálním řádu metody.

- Metody, jejichž součástí je použití přepínacích funkcí, tj. metody typu (3.19), jsou i pro hladká data pouze formálně prvního řádu. V bodech, ve kterých má řešení na dané časové vrstvě extrém, je metoda totiž pouze prvního řádu. Pokud chceme tento negativní jev odstranit, získáme metody, které nemají vlastnost TVD (tak je tomu např. u ENO metod, které jsou, zjednodušeně řečeno, založeny na přepínání různých aproximací vyššího řádu).

- Řád metody může být také omezen volbou (přibližného) Riemannova řešiče. Pokud např. použijeme metodu s numerickým tokem založeným na přesném Riemannově řešiči (tento tok je pouze Lipschitzovsky spojitý), výsledná metoda může být formálně prvního řádu, přestože ostatní komponenty metody jsou řádu vyššího.

- Další otázkou je problém šíření chyb v řešení. Pokud na dané časové vrstvě aproximujeme řešení v některých bodech schématem formálně vyššího řádu a v některých bodech schématem formálně prvního řádu, může se příslušná chyba šířit do okolí těchto bodů. Chyba v okolí bodů, ve kterých je řešení nehladké, tak ovlivňuje i body, ve kterých je řešení hladké (a metoda řádu formálně vyššího).

V dalším textu se budeme zabývat metodami, které jsou formálně vyššího řádu. Přestože nejde o řád tak, jak je chápán při analýze metody konečných diferencí, má smysl tyto metody konstruovat, protože mohou dávat přesnější výsledky než metody řádu formálně prvního. Tyto metody však budeme dále nazývat metody s vysokým rozlišením.

3.4 Metody s vysokým rozlišením

Uveďme nejdříve, co rozumíme pojmem metoda s vysokým rozlišením:

- Metoda negeneruje umělé (numerické) oscilace.
- Metoda negeneruje řešení s velkou numerickou difúzí, tj. „rozmazané řešení".
- Metoda korektně aproximuje entropické řešení, tj. rázové vlny, vlny zředění a kontaktní nespojitosti.
- Získáme aproximací formálně vyššího řádu v bodech, ve kterých je řešení hladké.

Metody s vysokým rozlišením lze zkonstruovat různými způsoby. Existuje řada přístupů, z nichž vyjmenujme:

- Hybridní metody a metody umělé vazkosti. Tyto metody jsme již probírali v odstavci věnovaném modifikacím diferenčních metod. Zopakujme, že hybridní metody jsou založené na přepínání mezi schématem prvního a vyššího řádu (metody typu flux limiter). Schéma prvního řádu může být založeno na metodě Godunovova typu, která konverguje k entropickému řešení. V námi popsaném případě jsme vyšli z metody formálně prvního řádu a doplnili ji korekcí, abychom dosáhli formálně druhého řádu. Lze postupovat i tak, že vyjdeme z metody formálně druhého řádu a provádíme korekci na formálně první řád. Tyto metody nesou označení WAF (weighted average flux). Druhou často využívanou jsou metody umělé vazkosti. Do této skupiny lze také například zařadit Jamesonovu-Schmidtovu-Turkelovu metodu a metodu SLIP (symmetric limited positive), které jsou taloženy na aproximaci formálně druhého řádu, která je modifikována pomocí numerické vazkosti závislé na hladkosti dat.

- Metody založené na obecnější rekonstrukci přibližného řešení z integrálních průměrů. Existuje celá řada rekonstrukcí (např. MUSCL –monotone upstream-centred for conservation laws , PPM – piecewise parabolic method, ENO – essentially non-oscillatory , WENO – weighted essentially non-oscillatory).

  Zvolíme-li rekonstrukci, která obsahuje i polynomy stupně 1 nebo vyššího, získáme v bodech $x_{j+1/2}$ tzv. zobecněné Riemannovy problémy. Až na výjimky bohužel přesné řešení zobecněného Riemannova problému nelze určit. Proto řešíme tento problém přibližně. Opět existuje řada přístupů: Hancockova metoda, Colellova metoda PLM (piece-wise linear method), metoda zobecněného Riemannova problému (metoda GRP) navržená Ben-Artzim a Falcovitzem, metoda SLIC (slope-limiter centred), semidiskrétní metody (aplikace metody přímek), metody ADER [6].

3.4.1 Hybridní metody

Jednu z hybridních metod jsme již vyložili – jde o Hartenovu hybridní metodu. Jak jsme zmínili, v rámci této metody lze využít metody Godunovova typu. Jinak řečeno, tok prvního řádu v Hartenově hybridní metodě lze nahradit tokem

$$F^{n,L}_{j+1/2} = \begin{cases} f(Q) & \text{je-li } f'(U^n_j) < 0 \text{ a } f'(U^n_{j+1}) > 0, \\ \frac{1}{2}[f(U^n_j) + f(U^n_{j+1})] - \frac{1}{2}|a^n_{j+1/2}|(U^n_{j+1} - U^n_j) & \text{jinak,} \end{cases} \tag{3.40}$$

kde Q je řešením rovnice $f'(Q) = 0$. Obecně lze tok $F^{n,H}_{j+1/2}$ druhého řádu nahradit tokem řádu vyššího. Pro odvození je potřeba použít Taylorův polynom vyššího stupně než druhého, např.

$$u(x_j, t_{n+1}) \approx u(x_j, t_n) + \Delta t u_t(x_j, t_n) + \frac{\Delta t^2}{2} u_{tt}(x_j, t_n) + \frac{\Delta t^3}{6} u_{ttt}(x_j, t_n)$$

Členy s časovými derivacemi lze vyjádřit

$$\begin{aligned} u_t &= -[f(u)]_x, \\ u_{tt} &= 2f'f''u_x^2 + (f')^2 u_{xx}, \\ u_{ttt} &= -6(f'')^2 f' u_x^3 - 3(f')^2 f''' u_x^3 - 9(f')^2 f'' u_x u_{xx} - (f')^3 u_{xxx}. \end{aligned}$$

Aproximace těchto výrazů pomocí diferencí dostatečně vysokých řádů a současně zajištění korektní aproximace nehladkých řešení je komplikovaný problém. V případě aproximací vyšších derivací není jasné, jak konstruovat omezovací funkce (limitery). Problémy nastávají i v případě úloh se zdrojovým členem, soustav rovnic i úloh ve více prostorových dimenzích. Proto se v případě hybridních metod používá především kombinace schémat prvního a druhého řádu. Zde lze dosáhnout korektní aproximace nehladkých řešení a současně lze snadno ukázat, že pro řešení hladká je metoda druhého řádu (zde ztotožňujeme aproximace integrálních průměrů s bodovými aproximacemi hledané funkce).

3.4.2 Metody konečných objemů založené na rekonstrukcích vyššího řádu

Doposud jsme v rámci metody konečných objemů uvažovali po částech konstantní rekonstrukci hledané funkce na zvolené časové vrstvě. Lze samozřejmě zvolit i rekonstrukce vyššího řádu. Rekonstruovaná funkce nesmí však mít větší totální variaci než po částech konstantní rekonstrukce.

To vede k tomu, že nelze použít např. libovolnou lineární rekonstrukci, ale je nutné modifikovat směrnice této po částech lineární rekonstrukce. Lze ukázat, že výše zmíněnou vlastnost má např. následující volba

$$\hat{U}^n(x) = \bar{U}_i^n + \sigma_i^n(x - x_i) \text{ pro } x_{i-1/2} < x < x_{i+1/2}. \tag{3.41}$$

Jednou z možností volby σ_i^n je omezení směrnice typu minmod

$$\sigma_i^n = \text{minmod}\left(\frac{U_i^n - U_{i-1}^n}{\Delta x}, \frac{U_{i+1}^n - U_i^n}{\Delta x}\right), \tag{3.42}$$

kde

$$\text{minmod}(a,b) = \begin{cases} a & \text{pro } |a| < |b| \text{ a } ab > 0, \\ b & \text{pro } |b| < |a| \text{ a } ab > 0, \\ 0 & \text{pro } ab < 0. \end{cases}$$

Jinou možností je omezení směrnice typu superbee

$$\sigma_i^n = \text{maxmod}(\sigma_i^{(1)}, \sigma_i^{(2)}), \tag{3.43}$$

kde

$$\begin{aligned} \sigma_i^{(1)} &= \text{minmod}\left(\tfrac{U_{i+1}^n - U_i^n}{\Delta x}, 2\tfrac{U_i^n - U_{i-1}^n}{\Delta x}\right), \\ \sigma_i^{(1)} &= \text{minmod}\left(2\tfrac{U_{i+1}^n - U_i^n}{\Delta x}, \tfrac{U_i^n - U_{i-1}^n}{\Delta x}\right), \\ \text{maxmod}(a,b) &= \begin{cases} a & \text{pro } |a| > |b|, \\ b & \text{pro } |b| > |a|. \end{cases} \end{aligned}$$

Další možností je omezení směrnice typu monotonized central-difference (MC) ve tvaru

$$\sigma_i^n = \text{minmod}\left(\frac{U_{i+1}^n - U_{i-1}^n}{2\Delta x}, 2\frac{U_i^n - U_{i-1}^n}{\Delta x}, 2\frac{U_{i+1}^n - U_i^n}{\Delta x}\right), \tag{3.44}$$

kde použijeme zobecnění funkce minmod

$$\text{minmod}(a_1, \dots, a_m) = \begin{cases} \min\limits_i a_i & \text{pro } a_j > 0 \quad \forall j, \\ \max\limits_i a_i & \text{pro } a_j < 0 \quad \forall j, \\ 0 & \text{jinak.} \end{cases}$$

Pro takto rekonstruovaná data definujeme takzvané zobecněné Riemannovy problémy

$$\begin{aligned} u_t + [f(u)]_x &= 0, x \in R, t \in (t_n, t_{n+1}), \\ u(x, t_n) &= \begin{cases} \hat{U}^n(x), & x < x_{j+1/2}, \\ \hat{U}^n(x), & x > x_{j+1/2}. \end{cases} \end{aligned} \tag{3.45}$$

Tento problém nelze obecně přesně řešit. Proto se obvykle používá semidiskrétní přístup, tj. vyjde se ze vztahu (3.26), který je nahrazen diskrétní aproximací

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{U}_j(t) = -\frac{1}{\Delta x}\left[F_{j+1/2}(t) - F_{j-1/2}(t)\right]. \tag{3.46}$$

Numerické toky $F_{j+1/2}(t)$ nepředstavují nyní aproximace integrálních průměrů přes interval (t_n, t_{n+1}), ale jsou chápány jako bodové hodnoty. Tyto hodnoty v čase $t = t_n$ lze stanovit jako $F_{j+1/2}(t_*) = f[\tilde{U}(x_{j+1/2}, t_*+)]$, kde $\tilde{U}(x, t_*+)$ jsou řešením standardních Riemannových problémů

$$\begin{aligned} u_t + [f(u)]_x &= 0, \quad x \in R, \quad t > t_*, \\ u(x, t_n) &= \begin{cases} \hat{U}^*(x_{j+1/2}-), & x < x_{j+1/2}, \\ \hat{U}^*(x_{j+1/2}+), & x > x_{j+1/2}. \end{cases} \end{aligned} \tag{3.47}$$

Diskretizace v časové proměnné je pak realizována pomocí Rungových-Kuttových metod. Jako příklad zvolme metodu druhého řádu. Nejdříve zavedeme označení

$$[\mathcal{L}(\bar{U}^*)]_j = \frac{1}{\Delta x}[F_{j+1/2} - F_{j-1/2}], \tag{3.48}$$

kde $\bar{U}^*$ představuje soubor integrálních průměrů, kterých je využito k rekonstrukci, řešení zobecněných Riemannových problémů a vyčíslení numerických toků $F_{j+1/2} = F_{j+1/2}(t^*)$. Pak lze metodu druhého řádu zapsat ve tvaru

$$\begin{aligned} \bar{U}^* &= \bar{U}^n - \Delta t \mathcal{L}(\bar{U}^n), \\ \bar{U}^{**} &= \bar{U}^* - \Delta t \mathcal{L}(\bar{U}^*), \\ \bar{U}^{n+1} &= \tfrac{1}{2}(\bar{U}^n + \bar{U}^{**}). \end{aligned} \tag{3.49}$$

Má-li metoda (tj. Eulerova metoda)

$$\bar{U}^{n+1} = \bar{U}^n - \Delta t \mathcal{L}(\bar{U}^n) \tag{3.50}$$

vlastnost TVD, pak i metoda (3.49) má tuto vlastnost, protože $\mathrm{TV}_*(\bar{U}) \leqq \mathrm{TV}_n(\bar{U})$ a $\mathrm{TV}_{**}(\bar{U}) \leqq \mathrm{TV}_*(\bar{U})$, a tedy

$$\mathrm{TV}_{n+1}(\bar{U}) \leqq \frac{1}{2}[\mathrm{TV}_n(\bar{U}) + \mathrm{TV}_{**}(\bar{U})] \leqq \frac{1}{2}[\mathrm{TV}_n(\bar{U}) + \mathrm{TV}_n(\bar{U})] = \mathrm{TV}_n(\bar{U}).$$

Metody tohoto typu jsou dnes označovány jako SSP (strong stability-preserving).

Všechny doposud popisované způsoby omezení směrnic měly vlastnost TVD. To znamená, že pokud na dané časové vrstvě řešení nabývá v některém bodu extrému, získáme po částech konstantní aproximaci. Jak jsme již zmínili, lze realizovat i aproximaci, která je vyššího řádu v bodech, kde se nabývá extrém. Postupuje se obvykle tak, že se vygeneruje více rekonstrukcí stejného stupně využívajících více okolních dat, a pak je vybrána taková, která nejméně osciluje. Na této myšlence jsou založeny například metody typu ENO (essentially non-oscillating). Těmi se zde však nebudeme zabývat.

3.5 Přibližné Riemannovy řešiče

Jak jsme zmínili, je často místo přesného Riemannova řešiče používán přibližný Riemannův řešič. Ve skalárním případě je většinou velmi snadné implementovat v rámci metody Godunovova typu přímo přesný Riemannův řešič. Přesto na tomto místě z didaktických důvodů popišme ještě další možnosti. Získané poznatky lze pak využít v případě soustav parciálních diferenciálních rovnic hyperbolického typu.

Budeme předpokládat, že máme k dispozici obecnou rekonstrukci hledané funkce na časové vrstvě $t = t_*$. Zavedeme zjednodušené značení rekonstrukcí hledané funkce na časové vrstvě $t = t_*$. $\hat{U}^-_{j+1/2} = \hat{U}(x_{j+1/2}-)$, $\hat{U}^+_{j+1/2} = \hat{U}(x_{j+1/2}+)$ Jak jsme uvedli, přesný Riemannův řešič vede k numerickému toku ve tvaru uvedeném v (3.40).

Mírným zjednodušením je tzv. Murmanova metoda, pro kterou je

$$\begin{aligned} F_{j+1/2} &= \tfrac{1}{2}[f(\hat{U}^+_{j+1/2}) + f(\hat{U}^-_{j+1/2})]- \\ &- \tfrac{1}{2}|a_{j+1/2}(t)|[\hat{U}^+_{j+1/2} - \hat{U}^-_{j+1/2}], \end{aligned} \tag{3.51}$$

kde

$$a_{j+1/2} = \frac{f(\hat{U}^+_{j+1/2}) - f(\hat{U}^-_{j+1/2})}{\hat{U}^+_{j+1/2} - \hat{U}^-_{j+1/2}}. \tag{3.52}$$

Získáme tak metodu typu upwind. Zdůrazněme ještě, že tato metoda nerespektuje řešení ve tvaru vlny zředění. V pasáži věnované výkladu Godunovovy metody jsme uvedli, že metodu upwind lze interpretovat jako metodu s přibližným Riemannovým řešičem typu all-shock. Z tohoto důvodu máme zaručenu konvergenci pouze ke slabému řešení, nikoliv k řešení entropickému (neboť vlny zředění neaproximujeme korektně).

Další přibližným řešičem je tzv. Rusanovova metoda (někdy také označována jako lokální Laxova-Friecrichsova metoda), pro kterou

$$\begin{aligned} F_{j+1/2} &= \tfrac{1}{2}[f(\hat{U}^+_{j+1/2}) + f(\hat{U}^-_{j+1/2})]- \\ &- \tfrac{1}{2}|a_{j+1/2}|[U^+_{j+1/2} - U^-_{j+1/2}], \end{aligned} \tag{3.53}$$

kde

$$a_{j+1/2} = \max\{|f(\hat{U}^-_{j+1/2})|, |f(\hat{U}^+_{j-1/2})|\}. \tag{3.54}$$

V případě této metody je $a_{j+1/2}$ získáno jako v absolutní hodnotě největší rychlost šíření vln. Metoda nerespektuje rázové vlny, ale pouze vlny zředění. Právě uvedená metoda patří současně do skupiny metod centrálních. Tuto metodu lze odvodit podobně jako Laxovu-Friedrichsovu metodu. Lze však ukázat, že Rusanovova aproximace se vyznačuje menším množstvím numerické vazkosti (difúze), než má aproximace Laxova-Friedrichsova.

Další možností je Hartenova-Laxova-Leerova (HLL) metoda, pro kterou

$$\begin{aligned} F_{j+1/2} &= \tfrac{a^+_{j+1/2} f(\hat{U}^-_{j+1/2}) - a^-_{j+1/2} f(\hat{U}^+_{j+1/2})}{a^+_{j+1/2} - a^-_{j+1/2}} - \\ &- \tfrac{a^+_{j+1/2} a^-_{j+1/2}}{a^+_{j+1/2} - a^-_{j+1/2}} [U^+_{j+1/2} - U^-_{j+1/2}], \end{aligned} \tag{3.55}$$

kde

$$\begin{aligned} a^+_{j+1/2} &= \max\{f(\hat{U}^-_{j+1/2}), f(\hat{U}^+_{j-1/2}), 0\}, \\ a^-_{j+1/2} &= \min\{f(\hat{U}^-_{j+1/2}), f(\hat{U}^+_{j-1/2}), 0\}. \end{aligned} \tag{3.56}$$

Tato metoda je opět konstruována na podobném principu jako metoda Rusanovova a Laxova-Friedrichsova, lze však ukázat, že je založena na aproximaci s menším množstvím numerické vazkosti (difúze). Někdy také bývá označována jako metoda typu central-upwind. Tuto metodu lze opět interpretovat jako metodu centrální (bez Riemannových řešičů). V některých speciálních případech přechází tato metoda na metodu typu upwind (např. v případě, že $f'(u) > 0 \quad \forall u$).

Dále uveďme Engquist-Osherovu metodu, pro kterou

$$\begin{aligned} F_{j+1/2} &= \tfrac{1}{2}[f(\hat{U}^+_{j+1/2}) + f(\hat{U}^-_{j+1/2})] - \\ &- \tfrac{1}{2} \int\limits_{U^-_{j+1/2}}^{U^+_{j+1/2}} |f'(q)| \, dq. \end{aligned} \tag{3.57}$$

Tato metoda je založena na představě, že řešením Riemannova problému je buď vlna zředění, nebo kompresní vlna. Uvažujeme tedy pouze řešení spojitá (v případě kompresní vlny však musíme uvažovat místo řešení ve tvaru funkce řešení ve tvaru relace až s třemi „funkčními hodnotami“).

Jako poslední uveďme znovu Lax-Friedrichsovu metodu, kterou lze také vnímat jako přibližný Riemannův řešič

$$\begin{aligned} F_{j+1/2} &= \tfrac{1}{2}[f(\hat{U}^+_{j+1/2}) + f(\hat{U}^-_{j+1/2})] - \\ &- \tfrac{1}{2}|a_{j+1/2}|[U^+_{j+1/2} - U^-_{j+1/2}], \end{aligned} \tag{3.58}$$

kde

$$a_{j+1/2} = \frac{\Delta x}{\Delta t}. \tag{3.59}$$

Povšimněme si, že z CFL podmínky dostáváme $\max|f'(q)| \leqq \frac{\Delta x}{\Delta t}$. Opět tedy $a_{j+1/2}$ je horním odhadem pro v absolutní hodnotě největší rychlost šíření vln.

3.6 Konvergence k zobecnému řešení. Laxova-Wendroffova věta

Nyní shrneme některé základní poznatky týkající se analýzy metod. V první řadě je třeba zavést novou definici konzistence, která není závislá na hladkosti řešení. Podíváme-li se na stávající definici, lze zjednodušeně říci, že jde o konzistenci (soulad) mezi numerickou aproximací a diferenciální rovnicí. Nová definice je založena na konzistenci mezi tokovou funkcí $f = f(u)$ a numerickou tokovou funkcí (numerickým tokem) $F^n_{j+1/2}$.

Definice 3.1. Řekneme, že *numerický tok je konzistentní* s tokovou funkcí $f = f(u)$, jestliže při konstantní hodnotě funkce $u(x,t) = \bar{u}$ platí, že se hodnota numerického toku redukuje na hodnotu $f(\bar{u})$, tj.

$$F(\bar{u}, \bar{u}, \ldots, \bar{u}) = f(\bar{u}), \quad \forall \bar{u} \in \mathbb{R} \tag{3.60}$$

a dále pro všechna $u_{j-m}, \ldots, u_{j+k}, \bar{u} \in \langle u_{\min}, u_{\max}\rangle$, $u_{\min}, u_{\max} \in \mathbb{R}$ existuje $K > 0$:

$$|F(u_{j-m}, \ldots, u_{j+k}) - f(\bar{u})| \leqq K \max\{|u_{j-m} - \bar{u}|, \ldots, |u_{j+k} - \bar{u}|\}. \tag{3.61}$$

Je zřejmé, že z platnosti nerovnosti (3.61) automaticky plyne rovnost (3.60). Uveďme si nyní dva příklady numerických toků a ukažme si jejich konzistenci.

1. $F^n_{j+1/2} = \frac{1}{2}[f(U^n_j) + f(U^n_{j+1})]$,
2. $F^n_{j+1/2} = \frac{1}{2}[f(U^n_j) + f(U^n_{j+1})] - \frac{1}{2}|a^n_{j+1/2}|(U^n_{j+1} - U^n_j)$, $a^n_{j+1/2} = \frac{f(U^n_{j+1}) - f(U^n_j)}{U^n_{j+1} - U^n_j}$.

Pro jednoduchost nebudeme uvádět horní index n.

1.

$$\begin{aligned}
\left|F_{j+1/2} - f(\bar{u})\right| &= \left|\frac{1}{2}[f(U_j) + f(U_{j+1})] - f(\bar{u})\right| = \\
&= \left|\frac{1}{2}[f(U_j) - f(\bar{u})] + \frac{1}{2}[f(U_{j+1}) - f(\bar{u})]\right| \leqq \\
&\leqq \frac{1}{2}\left|f(U_j) - f(\bar{u})\right| + \frac{1}{2}\left|f(U_{j+1}) - f(\bar{u})\right|
\end{aligned}$$

Dále $\exists C > 0$ taková, že

$$|f(U_{j+1}) - f(U_j)| \leqq C|U_{j+1} - U_j|, \quad \forall U_{j+1}, U_j \in \langle U_{\min}, U_{\max}\rangle. \tag{3.62}$$

Potom

$$\frac{1}{2}\left|f(U_j) - f(\bar{u})\right| + \frac{1}{2}\left|f(U_{j+1}) - f(\bar{u})\right| \leqq \frac{1}{2}C\left|U_j - \bar{u}\right| + \frac{1}{2}C\left|U_{j+1} - \bar{u}\right| \leqq$$

$$\leqq C \max\{|U_j - \bar{u}|, |U_{j+1} - \bar{u}|\}.$$

Zde tedy můžeme položit $K = C$.

2.

$$\begin{aligned}
&\left|F_{j+1/2} - f(\bar{u})\right| = \\
= &\left|\tfrac{1}{2}[f(U_j) + f(U_{j+1})] - \tfrac{1}{2}\underbrace{\left|\frac{f(U_{j+1}) - f(U_j)}{U_{j+1} - U_j}\right|}_{\leqq C, \forall U_j, U_{j+1} \in \langle U_{\min}, U_{\max}\rangle} (U_{j+1} - U_j) - f(\bar{u})\right| \leqq \\
\leqq &\quad \tfrac{1}{2}C\,|U_j - \bar{u}| + \tfrac{1}{2}C\,|U_{j+1} - \bar{u}| + \tfrac{1}{2}C|U_{j+1} - U_j| \leqq \\
\leqq &\ \tfrac{1}{2}C\,|U_j - \bar{u}| + \tfrac{1}{2}C\,|U_{j+1} - \bar{u}| + \tfrac{1}{2}C\,|U_j - \bar{u}| + \tfrac{1}{2}C\,|U_{j+1} - \bar{u}| \leqq \\
\leqq &\quad 2C \max\{|U_j - \bar{u}|, |U_{j+1} - \bar{u}|\}.
\end{aligned}$$

V tomto případě položíme $K = 2C$.

Tím jsme ukázali konzistenci obou numerických toků s tokem $f = f(u)$.

Nyní budeme formulovat jednu ze základních vět týkajících se konvergence metod.

Věta 3.2 (Laxova-Wendroffova). Uvažujme pro $t \in \langle 0, T\rangle$ a $x \in \mathbb{R}$ diskretizační sítě s parametry Δt_l, Δx_l a označme příslušnou síťovou aproximaci $U_{\Delta t_l}$. Pokud je aproximace $U_{\Delta t_l}$ stanovena konzervativní metodou s konzistentním numerickým tokem a pokud $U_{\Delta t_l}$ konverguje k nějaké funkci $u = u(x,t)$ pro $l \to \infty$, tj. $\Delta t_l \to 0$ a $\Delta x_l \to 0$, potom $u = u(x,t)$ je slabé řešení.

Síťovou aproximací v předchozí větě chápeme jako po částech konstantní funkci

$$U_{\Delta t_l} = U_j^n \text{ pro } x \in (x_{j-1/2}, x_{j+1/2}) \times (t_n, t_{n+1}).$$

V předchozí větě hovoříme o konvergenci metody. Je však ještě třeba upřesnit, co konvergencí metody přesně míníme:

- Prvním požadavkem je

$$\int_0^T \int_a^b |U_{\Delta t_l}(x,t) - u(x,t)|\,\mathrm{d}x\,\mathrm{d}t \to 0$$

pro každou množinu $\Omega = \langle a, b\rangle \times \langle 0, T\rangle$.

-

$$\forall T > 0 \quad \exists R \in \mathbb{R} : \quad \mathrm{TV}_t(U_{\Delta t_l}(\cdot, t)) < R \quad \forall t \in \langle 0, T\rangle, \quad l = 1, 2, \dots,$$

kde

$$\mathrm{TV}_t(v) = \sup \sum_{j=1}^{M} |v(\xi_j) - v(\xi_{j-1})|$$

přes všechna dělení $-\infty < \xi_0 < \xi_1 < \cdots < \xi_M < \infty$.

Důkaz. Schéma v konzervativním tvaru lze přepsat na

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + \frac{F_{j+1/2}^n - F_{j-1/2}^n}{\Delta x} = 0.$$

Přenásobením hodnotami $\varphi_j^n = \varphi(x_j, t_n)$ testovací funkce $\varphi \in C_0^1(\mathbb{R}^+ \times \mathbb{R})$ a sečtením přes indexy n a j získáme

$$\sum_{n=0}^{\infty} \sum_{j=-\infty}^{\infty} \left[\varphi_j^n \frac{U_j^{n+1} - U_j^n}{\Delta t} + \varphi_j^n \frac{F_{j+1/2}^n - F_{j-1/2}^n}{\Delta x} \right] = 0, \tag{3.63}$$

kde $\varphi_j^n = \varphi(x_j, t_n)$. Tento vztah (3.63) můžeme za použití rovnosti (formule per partes)

$$\sum_{j=c}^{d} r_j(s_{j+1} - s_j) = -\sum_{j=c+1}^{d} s_j(r_j - r_{j-1}) - r_c s_c + r_d s_{d+1} \tag{3.64}$$

a následným vynásobením $\Delta x \Delta t$ přepsat na

$$\begin{aligned} &-\Delta x \Delta t \sum_{n=0}^{\infty} \sum_{j=-\infty}^{\infty} \left[U_j^{n+1} \tfrac{\varphi_j^{n+1} - \varphi_j^n}{\Delta t} + F_{j+1/2}^n \tfrac{\varphi_{j+1}^n - \varphi_j^n}{\Delta x} \right] - \\ &-\Delta x \sum_{j=-\infty}^{\infty} U_j^0 \varphi_j^0 = 0. \end{aligned} \tag{3.65}$$

Podle definice slabého řešení předpokládáme kompaktní nosič funkce φ, proto členy $\varphi_j^n = 0$ pro $|j| \to \infty$ nebo $n \to \infty$.

Nyní uvažujme $i \to \infty$, tedy $\Delta t^{(i)}, \Delta x^{(i)} \to 0$ a $U^{(i)} \to u$. Potom první člen (3.65) konverguje k $-\int\limits_0^\infty \int\limits_{-\infty}^\infty \varphi_t u \,\mathrm{d}x\mathrm{d}t$ a třetí člen konverguje k $-\int\limits_{-\infty}^\infty \varphi(x,0)u(x,0)\mathrm{d}x$. Dále využijme vlastnosti konzistentní metody. Například pro tříbodovou metodu platí (pro volbu $\bar{u} = U_j^{(i)n}$ v (3.61))

$$|F_{j+1/2}^{(i)n} - f(U_j^{(i)n})| \leqq K|U_j^{(i)n} - U_{j+1}^{(i)n}|, \qquad \text{pro konstantu } K > 0. \tag{3.66}$$

Protože $U^{(i)n}$ má omezenou totální variaci, platí

$$|F_{j+1/2}^{(i)n} - f(U_j^{(i)n})| \to 0, \qquad \text{když } i \to \infty \tag{3.67}$$

pro skoro všechna j. Využitím (3.67) a konvergence $U^{(i)} \to u$ lze ukázat konvergenci

$$\Delta x \Delta t \sum_{n=0}^{\infty} \sum_{j=-\infty}^{\infty} F_{j+1/2}^{(i)n} \frac{\varphi_{j+1}^{(i)n} - \varphi_j^{(i)n}}{\Delta x} \to \int\limits_0^\infty \int\limits_{-\infty}^\infty f(u)\varphi_x \mathrm{d}x\mathrm{d}t. \tag{3.68}$$

Platí-li předpoklady věty a $l \to \infty$, dostáváme

$$-\int_0^\infty \int_{-\infty}^\infty [u\varphi_t + f(u)\varphi_x]\,\mathrm{d}x\mathrm{d}t - \int_{-\infty}^\infty u(x,0)\varphi(x,0)\mathrm{d}x = 0. \tag{3.69}$$

Podle definice 1.2 je tedy funkce u slabým řešením úlohy (1.6). □

Povšimněme si, že věta má tvar implikace. V případě, že existuje klasické řešení, může konvergovat i nekonzervativní metoda. Věta také neříká nic o tom, za jakých okolností daná metoda konverguje. K tomu, abychom mohli formulovat konvergenční věty, je třeba zavést pojem stability.

Definice 3.3. Metoda se nazývá TV-stabilní, jestliže všechny aproximace řešení $U_{\Delta t}(x,t)$, $\Delta_t < \Delta t_0$, patří do množiny

$$\mathcal{K} = \{U_{\Delta t} \in \mathcal{L}_{1,T} : \quad \mathrm{TV}(U_{\Delta t}) \leqq R \text{ a } \mathrm{supp}(U_{\Delta t}) \subset \langle -M, M\rangle \quad \forall t \in \langle 0, T\rangle\}, \tag{3.70}$$

kde

$$||v||_{1,T} = \int_0^T \int_{-\infty}^\infty |v(x,t)|\,\mathrm{d}x\,\mathrm{d}t,$$

$$\mathcal{L}_{1,T} = \{v : \quad ||v||_{1,T} < \infty\},$$

$$\begin{aligned} \mathrm{TV}(u) \;&=\; \limsup_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_0^T \int_{-\infty}^\infty |u(x+\varepsilon,t) - u(x,t)|\,\mathrm{d}x\,\mathrm{d}t \\ &+\; \limsup_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_0^T \int_{-\infty}^\infty |u(x,t+\varepsilon) - u(x,t)|\,\mathrm{d}x\,\mathrm{d}t. \end{aligned}$$

Dále platí následující konvergenční věta.

Věta 3.4. Nechť je numerické řešení $U_{\Delta t}$ generováno konzervativní numerickou metodou s konzistentním numerickým tokem. Je-li metoda TV-stabilní, potom přibližné řešení získané touto metodou konverguje ke slabému řešení.

Důkaz. Dokažme tuto větu. Nejdříve zavedeme měřítko rozdílu mezi přibližným a slabým řešením

$$\mathrm{dist}(U_{\Delta t}, \mathcal{U}) = \inf_{u \in \mathcal{U}} ||U_{\Delta t} - u||_{1,T},$$

kde

$$\mathcal{U} = \{u : \quad u(x,t) \text{ je slabé řešení}\}.$$

Důkaz povedeme sporem. Nechť metoda nekonverguje ke slabému řešení. To znamená, že existuje posloupnost přibližných řešení, která nekonverguje ke slabému řešení. Tato posloupnost pak obsahuje podposloupnost $U_{\Delta t_1}, U_{\Delta t_2}, \ldots$ takovou, že pro $\Delta t_l \to 0$, $l \to \infty$ platí

$$\operatorname{dist}(U_{\Delta t_l}, \mathcal{U}) > \varepsilon \quad \forall l.$$

Protože je metoda ovšem TV-stabilní, pak $U_{\Delta t_l} \in \mathcal{K}$. $\mathcal{K}$ je však kompaktní množina. Existuje tedy podposloupnost předchozí podposloupnosti, která konverguje k nějaké funkci $q \in \mathcal{K}$, tj. platí

$$||U_{\Delta t_l} - q||_{1,T} < \varepsilon \text{ pro dostatečně velká l.}$$

Jelikož uvažujeme konzervativní metodu s konzistentním numerickým tokem, pak podle Laxovy-Wendroffovy věty je funkce $q = q(x,t)$ slabým řešením. Dospěli jsme tedy ke sporu. □

Tato věta již říká, za jakých okolností metoda konverguje. To, zda je metoda TV-stabilní, však nelze snadno ověřit. Je třeba ještě formulovat několik vět, které již povedou na relativně snadno prověřitelné vlastnosti.

Poznamenejme ještě, že pro po částech konstantní funkci $U_{\Delta t} = U_{\Delta t}(x,t)$ platí

$$\mathrm{TV}(U_{\Delta t}) = \sum_{n=0}^{T/\Delta t} \sum_{j=-\infty}^{\infty} (\Delta t |U_{j+1}^n - U_j^n| + \Delta x |U_j^{n+1} - U_j^n|). \tag{3.71}$$

Definice 3.5. Řekneme, že metoda má vlastnost TVD (total variation diminishing), platí-li

$$\mathrm{TV}_{n+1}(U_{\Delta t}) \leqq \mathrm{TV}_n(U_{\Delta t}),$$

kde $\mathrm{TV}_n(U_{\Delta t}) = \sum_{j=-\infty}^{\infty} |U_{j+1}^n - U_j^n|$.

Řekneme, že metoda má vlastnost TVB (total variation bounded), existuje-li $B > 0$ nezávislé na datech metody (může však záviset na datech úlohy) takové, že

$$\mathrm{TV}_n(U_{\Delta t}) \leqq B \quad \forall n \in \mathbb{N}_0.$$

Zde je dobré připomenout, že slabé entropické řešení má vlastnost analogickou vlastnosti TVD. Nyní uvedeme dvě věty. První z nich má pomocný charakter, druhá již může posloužit k analýze metod.

Věta 3.6. Předpokládejme, že

- pro numerický tok F existuje $L > 0$ takové, že $\forall U^n_{j+i} \in \mathbb{R}$

$$|F(U^n_{j-l}, \dots, U^n_{j+k}) - F(U^n_{j-l-1}, \dots, U^n_{j+k-1})| \leqq L \sum_{i=-l}^{k} |U^n_{j+i} - U^n_{j+i-1}|. \quad (3.72)$$

- Dále předpokládejme, že existuje $R_0 > 0$

$$\mathrm{TV}_n(U_{\Delta t}) \leqq R_0.$$

Pak pro příslušnou konzervativní metodu existuje $R > 0$

$$\mathrm{TV}(U_{\Delta t}) \leqq R.$$

Věta 3.7. Nechť metoda má vlastnost TVD a nechť numerický tok splňuje podmínku (3.72) z předchozí věty. Je-li $\mathrm{TV}_0(U_{\Delta T}) < \infty$ a je-li množina $\mathrm{supp} u_0(x)$ omezená, potom tato metoda je TV stabilní.

Dále uvedeme Hartenovu větu, která je již přímým nástrojem pro ověření vlastnosti TVD.

Věta 3.8 (Hartenova). Uvažujme metodu, kterou lze vyjádřit ve tvaru

$$U^{n+1}_j = U^n_j - C^-_{n,j-1/2}(U^n_j - U^n_{j-1}) + C^+_{n,j+1/2}(U^n_{j+1} - U^n_j). \quad (3.73)$$

Pokud pro každé $j \in \mathbb{Z}$ a $n \in \mathbb{N}$ platí

$$C^-_{n,j-1/2} \geqq 0, \quad C^+_{n,j+1/2} \geqq 0, \quad C^-_{n,j+1/2} + C^-_{n,j-1/2} \leqq 1, \quad (3.74)$$

potom má metoda vlastnost TVD.

Dále zavedeme několik dalších vlastností metod, které jsou také užitečné pro analýzu některých metod

Definice 3.9. Metoda se nazývá

- monotónní, jestliže pro každá dvě přibližná řešení $W_{\Delta t}$ a $V_{\Delta t}$ dané úlohy pro dvě počáteční funkce $w_0(x)$, $v_0(x)$ a pro každé $n \in \mathbb{N}_0$ platí

$$(W_j^n \geqq V_j^n \quad \forall j \in \mathbb{Z}) \Rightarrow (W_j^{n+1} \geqq V_j^{n+1} \quad \forall j \in \mathbb{Z}).$$

- l_1-kontraktivní, jestliže pro libovolná výše uvedená přibližná řešení a pro všechna $n \in \mathbb{N}_0$ platí

$$||W^{n+1} - V^{n+1}|| \leqq ||W^n - V^n||.$$

- zachovávající monotonii, jestliže má tuto vlastnost: je-li počáteční funkce $w_0(x)$ monotónní, potom příslušné přibližné řešení W^n zachovává monotonii pro všechna $n \in \mathbb{N}$.

- E-schéma, pokud je konzervativní a platí

$$\begin{aligned} F_{j+1/2}^n &\leqq f(u) \quad \forall u \in \langle U_j^n, U_j^{n+1} \rangle, \quad U_j^n < U_j^{n+1}, \\ F_{j+1/2}^n &\geqq f(u) \quad \forall u \in \langle U_j^n, U_j^{n+1} \rangle, \quad U_j^n > U_j^{n+1}. \end{aligned}$$

Výše uvedené vlastnosti spolu souvisí tak, jak uvádí následující věta.

Věta 3.10.

- Každé tříbodové, konzervativní, monotónní a konzistentní schéma je E-schéma.
- Každá monotónní metoda je l_1 kontraktivní.
- Každá l_1 kontraktivní metoda má vlastnost TVD.
- Každá TVD metoda zachovává monotonii.

Důkazy jednotlivých tvrzení lze nalézt např. v [24].
Další věta je věnována monotónním metodám.

Věta 3.11. Numerické řešení získané konzistentní monotónní metodou konverguje k entropickému řešení.

Věta 3.12. Lineární metoda zachovávající monotonii je (až na speciální výjimky) prvního řádu.

Nyní ukážeme, že Laxova-Friedrichsova metoda je monotónní pro

$$\max_{j,n} |f'(U_j^n| \frac{\Delta t}{\Delta x} \leqq 1. \tag{3.75}$$

Laxova-Friedrichsova metoda má tvar

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{\Delta t}{2\Delta x}\left[f(U_{j+1}^n) - f(U_{j-1}^n)\right].$$

Monotonii ověříme pomocí parciálních derivací

$$\begin{aligned} \frac{\partial \mathcal{H}(U^n)}{\partial U_{j-1}^n} &= \tfrac{1}{2}\left[1 + \tfrac{\Delta t}{\Delta x} f'(U_{j-1}^n)\right] \\ \frac{\partial \mathcal{H}(U^n)}{\partial U_{j+1}^n} &= \tfrac{1}{2}\left[1 - \tfrac{\Delta t}{\Delta x} f'(U_{j+1}^n)\right]. \end{aligned}$$

Pokud je splněna podmínka (3.75), pak jsou výše uvedené derivace nezáporné, a tedy metoda je monotónní.

Dále ukážeme, že metoda

$$\begin{aligned} U_j^{n+1} &= U_j^n - \tfrac{\Delta t}{\Delta x}\left(F_{j+1/2}^n - F_{j-1/2}^n\right), \\ F_{j+1/2}^n &= \tfrac{1}{2}\left[f(U_j^n) + f(U_{j+1}^n)\right] - \tfrac{1}{2}|a_{j+1/2}^n|\left(U_{j+1}^n - U_j^n\right), \end{aligned} \tag{3.76}$$

kde

$$a_{j+1/2}^n = \begin{cases} \frac{f(U_{j+1}^n) - f(U_j^n)}{U_{j+1}^n - U_j^n} & \text{pro } U_j^n \neq U_{j+1}^n, \\ f'(U_j^n) & \text{jinak.} \end{cases} \tag{3.77}$$

má za splnění podmínky (3.75) vlastnost TVD. Po dosazení konkrétního tvaru numerického toku do metody v konzervativním tvaru a po přičtení a odečtení výrazu $\frac{\Delta t}{\Delta x} f(U_j^n)$ dostaneme

$$\begin{aligned} U_j^{n+1} &= U_j^n + \tfrac{\Delta t}{2\Delta x}(|a_{j+1/2}^n| - a_{j+1/2}^n)(U_{j+1}^n - U_j^n) - \\ &- \tfrac{\Delta t}{2\Delta x}(|a_{j-1/2}^n| + a_{j-1/2}^n)(U_j^n - U_{j-1/2}^n). \end{aligned}$$

Z toho plyne, že

$$C_{n,j-1/2}^- = \frac{\Delta t}{2\Delta x}(|a_{j-1/2}^n| + a_{j-1/2}^n),$$

a

$$C_{n,j+1/2}^+ = \frac{\Delta t}{2\Delta x}(|a_{j+1/2}^n| - a_{j+1/2}^n) =$$

Jelikož platí (3.75), pak

$$\begin{aligned} C_{n,j-1/2}^- + C_{n,j+1/2}^+ &= \tfrac{\Delta t}{2\Delta x}(|a_{j+1/2}^n| - a_{j+1/2}^n) = \tfrac{\Delta t}{\Delta x}|a_{j+1/2}^n| \leqq \\ &\leqq \tfrac{\Delta t}{\Delta x} \max |f'(Q_j^n)| \leqq 1. \end{aligned}$$

Ukázali jsme, že některé metody prvního řádu konvergují ke slabému řešení. Podrobnější analýzou lze ukázat, že k entropickému slabému řešení konvergují Laxova-Friedrichsova, Rusanovova, Hartenova-Laxova-Leerova, Engquistova-Osherova a Godunovova metoda s přesným Riemannovým řešičem. Další otázkou je, jak zkonstruovat metody vyššího než prvního řádu, které konvergují ke slabému řešení. Z předchozí věty plyne následující tvrzení: je-li metoda vyššího než prvního řádu, pak je nelineární nebo nezachovává monotónii. Z předposlední věty taktéž plyne, že nezachovává-li metoda monotónii, pak není TVD. Metoda, která není TVD zřejmě nekonverguje ke slabému řešení. Jedinou možností je tedy použít nelineární metody. Lze ukázat, že takovou metodou je již zmíněná hybridní metoda, která je postavena na myšlence užít aproximaci druhého řádu v bodech, ve kterých je řešení hladké, a aproximaci prvního řádu v bodech, ve kterých je řešení nehladké. Připomeňme, že tento přístup je společný všem zatím diskutovaným metodám. Platí následující dvě věty.

Věta 3.13. Pokud $\Phi = \Phi(\Theta)$ je omezená funkce, pak je Hartenova hybridní metoda konzistentní.

Věta 3.14. Hartenova hybridní metoda je TVD, pokud platí

- $\frac{\Delta t}{\Delta x} \max |f'(u)| \leqq 1$,
- $\Phi(\Theta) = 0$ pro $\Theta \leqq 0$,
- $\left| \frac{\Phi(\Theta_2)}{\Theta_2} - \Phi(\Theta_1) \right| \leqq 2 \quad \forall \Theta_1, \Theta_2 \in \mathbb{R}$.

Ukázali jsme, že Laxova-Friedrichsova metoda konverguje k entropickému řešení. Nyní ukážeme, že metoda typu upwind k entropickému řešení obecně nekonverguje. Uvažujme počáteční úlohu pro Burgersovu rovnici

$$\begin{aligned} u_t + (u^2/2)_x &= 0, \quad x \in \mathbb{R}, \quad t \in \langle 0, T \rangle, \\ u(x,0) &= \begin{cases} -1 & x \leqq 0 \\ 1 & x > 0. \end{cases} \end{aligned}$$

Dále uvažujme diskretizaci $\Delta x = 1$, $x_j = j$. Tedy

$$U_j^0 = \begin{cases} -1 & j \leqq 0 \\ 1 & j > 0. \end{cases}$$

Lze snadno ukázat, že $F_{j+1/2}^0 = 1/2$. Z toho plyne $U_j^1 = U_j^0$ pro všechna $j \in \mathbb{Z}$. Stejnou úvahu lze učinit i pro další časové kroky. Numerickým řešením je tedy stacionární rázová vlna. Toto řešení je aproximací slabého řešení, není však aproximací řešení entropického.

Závěrem konstatujme, že máme k dispozici metody, které konvergují k entropickému řešení a jsou prvního řádu (Laxova-Friedrichsova metoda, Godunovova metoda). Dále máme k dispozici metody s vysokým rozlišením, které ovšem konvergují k řešení slabému (nemáme zaručenu konvergenci k řešení entropickému). Dostáváme se tak k dalšímu problému: lze zkostruovat metody s vysokým rozlišením, které k entropickému řešení konvergují. Ukazuje se, že takové metody zkonstruovat lze (H. Nessyahu & E. Tadmor (1990): Non-oscillatory central differencing for hyperbolic conservation laws Journal of Computational Physics 87, 1990, 408-463.). Často však užíváme i metody, u kterých nemáme teoreticky prokázanou konvergenci k entropickému řešení. Obvykle postupujeme tak, že doplňujeme příslušnou metodu o proceduru zvanou entropy-fix. Jde o stejný postup, jako jsme uplatnili v odstavci o hybridních metodách, kde jsme pro stanovení toku prvního řádu použili Godunovovu metodu (která k entropickému řešení konverguje).

Poznamenejme ještě, že několikrát zmiňované metody typu ENO nemají vlastnost TVD. Zachovávají však monotonii.

3.7 Numerické metody pro řešení soustav parciálních diferenciálních rovnic hyperbolického typu

3.7.1 Metody pro lineární soustavy s konstantními koeficienty

Uvažujme počáteční úlohu pro soustavu

$$\begin{aligned} \boldsymbol{u}_t + \boldsymbol{A} u_x &= \boldsymbol{0}, \quad x \in \mathbb{R}, \quad t \in (0, T) \\ \boldsymbol{u}(x,0) &= \boldsymbol{u}_0(x), \quad x \in \mathbb{R}. \end{aligned} \tag{3.78}$$

Předpokládáme, že soustava je silně hyperbolická. Z toho plyne, že existuje regulární matice $\boldsymbol{R}$ a diagonální matice $\boldsymbol{\Lambda}$ tak, že platí (podrobnosti lze nalézt v úvodní teoretické kapitole)

$$\boldsymbol{\Lambda} = \boldsymbol{R}^{-1} \boldsymbol{A} \boldsymbol{R}. \tag{3.79}$$

Po zavedení substituce $\boldsymbol{u} = \boldsymbol{R}\boldsymbol{v}$ lze soustavu upravit na tvar

$$\boldsymbol{v}_t + \boldsymbol{\Lambda} \boldsymbol{v}_x = \boldsymbol{0}. \tag{3.80}$$

Získali jsme tak soustavu m nezávislých rovnic, které můžeme řešit metodami uvedenými v předchozích kapitolách. My se nyní však zaměříme speciálně na metody, které lze použít i pro řešení nelineárních úloh.

3.7.2 Metoda umělé vazkosti

Uvažujme opět nejdříve lineární soustavu

$$\boldsymbol{u}_t + \boldsymbol{A} \boldsymbol{u}_x = \boldsymbol{0}.$$

Místo úlohy s touto soustavou budeme řešit úlohu se soustavou modifikovanou

$$\boldsymbol{u}_t + \boldsymbol{A} u_x = d(\Delta t, \Delta x) \boldsymbol{u}_{xx}. \tag{3.81}$$

Pro jednoduchost budeme dále uvažovat semidiskrétní metody (to znamená také $d = d(\Delta x)$) ve tvaru

$$\frac{\mathrm{d}\boldsymbol{U}_j}{\mathrm{d}t} = -\frac{1}{\Delta x}\left(\boldsymbol{F}_{j+1/2} - \boldsymbol{F}_{j-1/2}\right), \tag{3.82}$$

kde

$$\boldsymbol{F}_{j+1/2} = \frac{1}{2}\boldsymbol{A}\left(\boldsymbol{U}_j + \boldsymbol{U}_{j+1}\right). \tag{3.83}$$

Právě uvedená metoda je druhého řádu v prostorové proměnné. Je však nestabilní. Zvolíme-li

$$d(\Delta x) = \alpha\frac{\Delta x}{2}, \quad \alpha = \max_{p=1\ldots m}\{|\lambda_i|\}, \tag{3.84}$$

dostaneme následující aproximaci vztahu (3.81)

$$\frac{\mathrm{d}\boldsymbol{U}_j}{\mathrm{d}t} = -\frac{1}{\Delta x}\left(\boldsymbol{F}_{j+1/2} - \boldsymbol{F}_{j-1/2}\right) + \frac{\alpha}{2\Delta x}(\boldsymbol{U}_{j+1} - 2\boldsymbol{U}_j + \boldsymbol{U}_{j-1}). \tag{3.85}$$

Provedeme-li substituci $\boldsymbol{U}_j = \boldsymbol{R}\boldsymbol{V}_j$, že právě uvedená metoda odpovídá metodě

$$\frac{\mathrm{d}\boldsymbol{V}_j}{\mathrm{d}t} = -\frac{1}{\Delta x}\boldsymbol{\Lambda}\left(\boldsymbol{V}_{j+1} - \boldsymbol{V}_{j-1}\right) + \frac{\alpha}{2\Delta x}(\boldsymbol{V}_{j+1} - 2\boldsymbol{V}_j + \boldsymbol{V}_{j-1}). \tag{3.86}$$

Popsaný postup lze tedy interpretovat jako uplatnění metody umělé vazkosti na každou z nezávislých rovnic soustavy $\boldsymbol{v}_t + \boldsymbol{\Lambda}\boldsymbol{v}_x = \boldsymbol{0}$. Získaná metoda je formálně prvního řádu. Chceme-li zkonstruovat metodu s vysokým rozlišením, je třeba opět zavést – stejně jako ve skalárním případě – vazké členy vyššího řádu. V nelineárním případě lze nejjednodušší verzi metody zkonstruovat následujícím způsobem. Numerické toky volíme

$$\boldsymbol{F}_{j+1/2} = \frac{1}{2}\left(\boldsymbol{f}(\boldsymbol{U}_j) + \boldsymbol{f}(\boldsymbol{U}_{j+1})\right) \tag{3.87}$$

a metoda má tvar

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{U}_j}{\mathrm{d}t} &= -\tfrac{1}{\Delta x}\Big[\left(\boldsymbol{F}_{j+1/2} - \boldsymbol{F}_{j-1/2}\right) + \tfrac{\alpha_{j+1/2}}{2\Delta x}(\boldsymbol{U}_{j+1} - \boldsymbol{U}_j) - \\ &- \tfrac{\alpha_{j-1/2}}{2\Delta x}(\boldsymbol{U}_j - \boldsymbol{U}_{j-1})\Big],\end{aligned} \tag{3.88}$$

kde

$$\alpha_{j+1/2} = \max_{p=1\ldots m}\{|\lambda_{i,j}|, |\lambda_{i,j+1}|\} \tag{3.89}$$

a $\lambda_{i,j}$ jsou vlastní čísla matice

$$\boldsymbol{A}_j = \boldsymbol{f}'(\boldsymbol{U}_j).$$

3.7.3 Hartenova hybridní metoda

Metodu pro případ soustav rovnic $\boldsymbol{u}_t + \boldsymbol{A}\boldsymbol{u}_x = \boldsymbol{0}$ budeme konstruovat podobně jako ve skalárním případě. Konstruujeme opět metodu v konzervativním tvaru

$$\boldsymbol{U}_j^{n+1} = \boldsymbol{U}_j^n - \frac{\Delta t}{\Delta x}\left(\boldsymbol{F}_{j+1/2}^n - \boldsymbol{F}_{j-1/2}^n\right). \tag{3.90}$$

Toky pro metodu prvního a druhého řádu lze vyjádřit ve tvaru

$$\boldsymbol{F}_{j+1/2}^{n,L} = \frac{1}{2}\boldsymbol{A}(\boldsymbol{U}_j^n + \boldsymbol{U}_{j+1}^n) - \frac{1}{2}|\boldsymbol{A}|(\boldsymbol{U}_{j+1}^n - \boldsymbol{U}_j^n), \tag{3.91}$$

$$\boldsymbol{F}_{j+1/2}^{n,H} = \frac{1}{2}\boldsymbol{A}(\boldsymbol{U}_j^n + \boldsymbol{U}_{j+1}^n) - \frac{1}{2}\boldsymbol{A}^2\frac{\Delta t}{\Delta x}(\boldsymbol{U}_{j+1}^n - \boldsymbol{U}_j^n), \tag{3.92}$$

kde $|\boldsymbol{A}| = \boldsymbol{R}|\boldsymbol{\Lambda}|\boldsymbol{R}^{-1}$. Zavedeme rozklad, který nám umožní použít omezovací funkce na jednotlivé složky po transformaci na nezávislé rovnice

$$\boldsymbol{U}_{j+1}^n - \boldsymbol{U}_j^n = \sum_{i=1}^m \alpha_{i,j+1/2}^n \boldsymbol{r}_i. \tag{3.93}$$

Rozdíl toků prvního a druhého řádu lze zapsat

$$\begin{aligned}
\boldsymbol{F}_{j+1/2}^{n,H} - \boldsymbol{F}_{j+1/2}^{n,L} &= \tfrac{1}{2}|\boldsymbol{A}|\left(\boldsymbol{I} - \tfrac{\Delta t}{\Delta x}|\boldsymbol{A}|\right)(\boldsymbol{U}_{j+1}^n - \boldsymbol{U}_j^n) = \\
&= \tfrac{1}{2}|\boldsymbol{A}|\left(\boldsymbol{I} - \tfrac{\Delta t}{\Delta x}|\boldsymbol{A}|\right)\boldsymbol{R}\boldsymbol{\alpha} = \tfrac{1}{2}\boldsymbol{R}|\boldsymbol{\Lambda}|\left(\boldsymbol{I} - \tfrac{\Delta t}{\Delta x}|\boldsymbol{\Lambda}|\right)\boldsymbol{\alpha} = \\
&= \tfrac{1}{2}\sum_{i=1}^m |\lambda_i|\left(1 - \tfrac{\Delta t}{\Delta x}|\lambda_i|\right)\alpha_{i,j+1/2}^n \boldsymbol{r}_i.
\end{aligned} \tag{3.94}$$

Korekci toků uplatňujeme prostřednictvím koeficientů $\alpha_{i,j+1/2}^n$. Zavedeme

$$\Theta_{i,j+1/2}^n = \begin{cases} \frac{\alpha_{i,j-1/2}^n}{\alpha_{i,j+1/2}^n}, & \lambda_i > 0, \\ \frac{\alpha_{i,j+3/2}^n}{\alpha_{i,j+1/2}^n}, & \lambda_i < 0. \end{cases} \tag{3.95}$$

Definujeme dále

$$\tilde{\alpha}_{i,j+1/2}^n = \alpha_{i,j+1/2}^n \Phi(\Theta_{i,j+1/2}^n). \tag{3.96}$$

Celkový tok metody je pak definován

$$\boldsymbol{F}_{j+1/2}^n = \boldsymbol{F}_{j+1/2}^{n,L} + \frac{1}{2}\sum_{i=1}^m |\lambda_i|\left(1 - \frac{\Delta t}{\Delta x}|\lambda_i|\right)\tilde{\alpha}_{i,j+1/2}^n \boldsymbol{r}_i. \tag{3.97}$$

Zůstává otázkou, jak modifikovat metodu pro případ nelineárních soustav. Při výpočtu toku $\boldsymbol{F}_{j+1/2}^n$ nahradíme matici $|\boldsymbol{A}|$ (a příslušná vlastní čísla a vlastní vektory) maticí $|\boldsymbol{A}_{j+1/2}^n|$ (a odpovídajícími vlastními čísly a vektory). Tuto matici můžeme definovat různými způsoby, např.

- $|\boldsymbol{A}^n_{j+1/2}| = |\boldsymbol{f}'(\boldsymbol{U}^n_j)|$
- $|\boldsymbol{A}^n_{j+1/2}| = |\boldsymbol{f}'(\boldsymbol{U}^n_{j+1})|$
- $|\boldsymbol{A}^n_{j+1/2}| = \frac{1}{2}\left(|\boldsymbol{f}'(\boldsymbol{U}^n_j)| + |\boldsymbol{f}'(\boldsymbol{U}^n_{j+1})|\right)$
- $|\boldsymbol{A}^n_{j+1/2}| = \left|\boldsymbol{f}'\left(\frac{\boldsymbol{U}^n_j + \boldsymbol{U}^n_{j+1}}{2}\right)\right|$
- $|\boldsymbol{A}^n_{j+1/2}| = \operatorname{sign}\left[\boldsymbol{f}'\left(\frac{\boldsymbol{U}^n_j + \boldsymbol{U}^n_{j+1}}{2}\right)\right]\left(\boldsymbol{f}(\boldsymbol{U}^n_{j+1}) - \boldsymbol{f}(\boldsymbol{U}^n_j)\right)$

kde

$$\operatorname{sign}\boldsymbol{A} = \boldsymbol{R}(\operatorname{sign}\boldsymbol{\Lambda})\boldsymbol{R}^{-1}.$$

Dostáváme tak

$$\boldsymbol{F}^{n,L}_{j+1/2} = \frac{1}{2}\left(\boldsymbol{f}(\boldsymbol{U}^n_j) + \boldsymbol{f}(\boldsymbol{U}^n_{j+1})\right) - \frac{1}{2}|\boldsymbol{A}^n_{j+1/2}|(\boldsymbol{U}^n_{j+1} - \boldsymbol{U}^n_j), \tag{3.98}$$

a

$$\boldsymbol{F}^n_{j+1/2} = \boldsymbol{F}^{n,L}_{j+1/2} + \frac{1}{2}\sum_{i=1}^{m}|\lambda^n_{i,j+1/2}|\left(1 - \frac{\Delta t}{\Delta x}|\lambda^n_{i,j+1/2}|\right)\tilde{\alpha}^n_{i,j+1/2}\boldsymbol{r}^n_{i,j+1/2}. \tag{3.99}$$

Právě popsaná metoda je konzervativní. Nemáme zaručenu korektní aproximaci vln zředění, proto je ji třeba doplnit procedurou typu entropy fix. Dalším problémem je, že metoda je založena na linearizaci nelineárního problému. Z toho plynou některé nedostatky metody. Uveďme jeden příklad: zatímco některé složky řešení původního nelineárního problému jsou v řadě konkrétních případů nezáporné (např. hustota), přibližné řešení tuto vlastnost obecně nemá.

3.7.4 Metody Godunovova typu

Stejně jako ve skalárním případě lze zkonstruovat metody Godunovova typu. Struktura algoritmu zůstává stejná: rekonstrukce aproximace řešení z aproximací integrálních průměrů; řešení Riemannových problémů (pouze v případě, že je v bodech $x = x_{j+1/2}$ rekonstrukce nespojitá); stanovení numerických toků, tj. hodnoty integrálů z tokové funkce $\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{u})$ přes časový interval (t_n, t_{n+1}) a stanovení integrálních průměrů na časové vrstvě t_{n+1} pomocí bilanční formule

$$\bar{\boldsymbol{U}}^{n+1}_j = \bar{\boldsymbol{U}}^n_j - \frac{\Delta t}{\Delta x}\left(\bar{\boldsymbol{F}}^{n+1/2}_{j+1/2} - \bar{\boldsymbol{F}}^{n+1/2}_{j-1/2}\right), \tag{3.100}$$

$$\bar{\boldsymbol{F}}^{n+1/2}_{j+1/2} = \int_{t_n}^{t_{n+1}} \boldsymbol{f}(\tilde{\boldsymbol{U}}^n(x_{j+1/2}, t))\,\mathrm{d}t, \tag{3.101}$$

příp.

$$\bar{\boldsymbol{U}}_{j+1/2}^{n+1} = \bar{\boldsymbol{U}}_{j+1/2}^{n} - \frac{\Delta t}{\Delta x}\left(\bar{\boldsymbol{F}}_{j+1}^{n+1/2} - \bar{\boldsymbol{F}}_{j}^{n+1/2}\right), \tag{3.102}$$

$$\bar{\boldsymbol{F}}_{j}^{n+1/2} = \int\limits_{t_n}^{t_{n+1}} \boldsymbol{f}(\tilde{\boldsymbol{U}}^n(x_j,t))\,\mathrm{d}t, \tag{3.103}$$

kde $\tilde{\boldsymbol{U}}^n = \tilde{\boldsymbol{U}}^n(x,t)$ je řešení Riemannových problémů pro rekonstruovaná data na časové vrstvě $t = t_n$, popř. řešení počátečních úloh s hladkou počáteční podmínkou (pokud je rekonstrukce v okolí bodů $x = x_{j+1/2}$ hladká).

Po částech konstantní rekonstrukci lze realizovat po složkách. V případě (3.102) není součástí postupu řešení Riemannových problémů. Získáme stejně jako ve skalárním případě Laxovu-Friedrichsovu metodu ve tvaru

$$\bar{\boldsymbol{U}}_{j+1/2}^{n+1} = \frac{\bar{\boldsymbol{U}}_{j}^{n} + \bar{\boldsymbol{U}}_{j+1}^{n}}{2} - \frac{\Delta t}{\Delta x}\left(\boldsymbol{f}(\bar{\boldsymbol{U}}_{j+1}^{n}) - \boldsymbol{f}(\bar{\boldsymbol{U}}_{j}^{n})\right). \tag{3.104}$$

V případě (3.101) potřebujeme znát řešení Riemannova problému v bodech $x = x_{j+1/2}$. Přestože nepotřebujeme znát kompletní strukturu řešení, jedná se na rozdíl od skalárního případu o velice složitou úlohu (přestože potřebujeme částečnou informaci o řešení, musíme alespoň částečně prozkoumat řešení celé). Na druhou stranu je vhodné si uvědomit, že získané aproximace integrálních průměrů interpretujeme obvykle i jako aproximace hodnot řešení v diskretizačních bodech $x = x_j$. Proto je třeba volit relativně jemné dělení. Z tohoto důvodu je v bodech $x = x_{j+1/2}$ až na výjimky obvykle řešením hladká funkce, nebo samotná rázová vlna, vlna zředění, anebo kontaktní nespojitost. Struktura řešení Riemannova problému je tedy obvykle velmi jednoduchá. Toto lze vnímat jako podpůrný argument pro použití přibližných Riemannových řešičů. Pouze v případě, kdybychom volili hrubou diskretizační síť, museli bychom věnovat velmi velkou pozornost řešení Riemannova problému. Získali bychom tak přesnější aproximace integrálních průměrů řešení. Ty bychom ovšem již nemohli interpretovat jako bodové hodnoty (příp. by musela být výstupem rekonstrukce prováděná také na základě řešení Riemannova problému).

Na závěr kapitoly konstatujme, že nutnou podmínkou stability je

$$\max_{j,n}\left\{|\lambda_1(\bar{\boldsymbol{U}}_j^n)|,\ldots,|\lambda_m(\bar{\boldsymbol{U}}_j^n)|\right\}\frac{\Delta t}{\Delta x} \leqq \frac{1}{2}, \tag{3.105}$$

kde $\lambda_1(\bar{\boldsymbol{U}}_j^n)$ jsou vlastní čísla Jacobiho matice $\boldsymbol{A}(\bar{\boldsymbol{U}}_j^n) = \boldsymbol{f}'(\bar{\boldsymbol{U}}_j^n)$. Toto omezení na časový krok může být prodlouženo

$$s_{\max}\frac{\Delta t}{\Delta x} \leqq 1, \tag{3.106}$$

kde $s_{\max}$ je horním odhadem maximální velikosti rychlostí šíření vln pro danou úlohu.

3.7.5 Přibližné Riemannovy řešiče

Jak jsme uvedli, na rozdíl od skalárního případu, je řešení Riemannových problémů pro soustavy nelineárních rovnic velice složité. Proto v tomto případě věnujeme velkou pozornost řešičům přibližným. Budeme uvažovat semidiskrétní přístup. Pro zjednodušení popisu zavedeme následující označení hodnot rekonstruovaných funkcí $\hat{\boldsymbol{U}}^-_{j+1/2} = \hat{\boldsymbol{U}}(x_{j+1/2}-)$, $\hat{\boldsymbol{U}}^+_{j+1/2} = \hat{\boldsymbol{U}}(x_{j+1/2}+)$. Vynecháme také označení závislosti na časové proměnné.

Roeův řešič

Prvním z přibližných Riemannových řešičů, které jsme studovali byla Murmanova metoda. I v případě nelineárních soustav volíme přesnou linearizaci ve tvaru

$$\boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j+1/2}\right) - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) = \boldsymbol{A}_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right). \tag{3.107}$$

Je-li řešením problému samotná rázová vlna (příp. kontaktní nespojitost), je splněn Rankineův-Hugoniotův vztah ve tvaru

$$\boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j+1/2}\right) - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) = s_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right). \tag{3.108}$$

Z předchozích dvou vztahů plyne

$$\boldsymbol{A}_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right) = s_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right). \tag{3.109}$$

Zvolená linearizace má tedy následující vlastnost: je-li řešením nelineárního problému samotná rázová vlna, je řešením linearizovaného problému také rázová vlna, která se navíc šíří totožnou rychlostí. Matice $\boldsymbol{A}_{j+1/2}$ se obvykle nazývá Roeova matice a je funkcí hodnot $\hat{\boldsymbol{U}}^-_{j+1/2}$ a $\hat{\boldsymbol{U}}^+_{j+1/2}$. Na Roeovu matici klademe dva požadavky:

- Roeova matice je stejně jako Jacobiho matice tokové funkce diagonalizovatelná a má navzájem různá reálná vlastní čísla;
- pro $\hat{\boldsymbol{U}}^-_{j+1/2} \to \hat{\boldsymbol{U}}$ a $\hat{\boldsymbol{U}}^+_{j+1/2} \to \hat{\boldsymbol{U}}$ platí $\boldsymbol{A}_{j+1/2} \to \boldsymbol{A}(\hat{\boldsymbol{U}})$.

Otázkou je, jakým způsobem získat konkrétní tvar Roeovy matice. Jednou z možností je integrovat Jacobiho matici po vhodné integrační cestě ve stavovém prostoru mezi body $\hat{\boldsymbol{U}}^-_{j+1/2}$ a $\hat{\boldsymbol{U}}^+_{j+1/2}$ (vzhledem k tomu, že pracujeme se zákonem zachování, lze ukázat, že si lze zvolit libovolnou integrační cestu). Zvolíme tedy jednoduchou variantu – úsečku

$$\boldsymbol{u}(\xi) = \hat{\boldsymbol{U}}^-_{j+1/2} + [\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}]\xi, \tag{3.110}$$

kde $0 \leqq \xi \leqq 1$. Pak lze psát

$$\begin{aligned} & \boldsymbol{f}\left(\hat{\boldsymbol{U}}^{+}_{j+1/2}\right) - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^{-}_{j+1/2}\right) = \int\limits_0^1 \frac{\mathrm{d}\,\boldsymbol{f}(\boldsymbol{u}(\xi))}{\mathrm{d}\,\xi}\mathrm{d}\,\xi = \\ & = \int\limits_0^1 \frac{\mathrm{d}\,\boldsymbol{f}(\boldsymbol{u}(\xi)))}{\mathrm{d}\,\boldsymbol{u}}\boldsymbol{u}'(\xi)\mathrm{d}\,\xi = \\ & = \left[\int\limits_0^1 \boldsymbol{f}'(\boldsymbol{u}(\xi)))\mathrm{d}\,\xi\right]\left(\hat{\boldsymbol{U}}^{+}_{j+1/2} - \hat{\boldsymbol{U}}^{-}_{j+1/2}\right). \end{aligned} \tag{3.111}$$

Z právě provedeného odvození a z (3.107) dostáváme

$$\boldsymbol{A}_{j+1/2} = \int\limits_0^1 \boldsymbol{f}'(\boldsymbol{u}(\xi)))\mathrm{d}\,\xi. \tag{3.112}$$

Použijeme-li tento vztah pro konstrukci Roeovy matice, je pochopitelně splněn vztah (3.107). Je také automaticky splněn druhý požadavek kladený na Roeovu matici. Splnění prvního požadavku však není zaručeno (i když Jacobiho matice tokové funkce tuto vlastnost má). Roe navrhl postup, jak zkonstruovat po něm nazvanou matici, která splňuje oba dva požadavky. Podrobnosti nalezne čtenář např. v [9]. Používáme opět algoritmus typu REA (reconstruct – evolve – average). Místo původního Riemannova problému, jehož přesné řešení příliš komplikované, řešíme přesně linearizované Riemannovy problémy ve tvaru

$$\begin{aligned} \boldsymbol{u}_t + \boldsymbol{A}_{j+1/2}\boldsymbol{u}_x &= 0, \quad x \in \mathbb{R}, \quad t > t_*, \\ \boldsymbol{u}(x, t_*) &= \begin{cases} \hat{\boldsymbol{U}}^{-}_{j+1/2}, & x < x_{j+1/2}, \\ \hat{\boldsymbol{U}}^{+}_{j+1/2}, & x > x_{j+1/2}. \end{cases} \end{aligned} \tag{3.113}$$

Při kontrukci numerické tokové funkce vycházíme však z původního zákona zachování, abychom zajistili konzervativitu i v případě numerického schématu (formulace (3.113) konzervativní není). Získaná toková funkce

$$\boldsymbol{A}^{+} = \frac{1}{2}(\boldsymbol{A} + |\boldsymbol{A}|), \quad \boldsymbol{A}^{-} = \frac{1}{2}(\boldsymbol{A} - |\boldsymbol{A}|),$$

je formálně shodná s tokovou funkcí Murmanovy metody

$$\boldsymbol{F}_{j+1/2} = \frac{1}{2}\left[\boldsymbol{f}\left(\hat{\boldsymbol{U}}^{-}_{j+1/2}\right) + \boldsymbol{f}\left(\hat{\boldsymbol{U}}^{+}_{j+1/2}\right)\right] - \frac{1}{2}|\boldsymbol{A}_{j+1/2}|\left(\hat{\boldsymbol{U}}^{+}_{j+1/2} - \hat{\boldsymbol{U}}^{-}_{j+1/2}\right). \tag{3.114}$$

Lokální Laxova-Friedrichsova metoda, semidiskrétní centrální metoda

Popíšeme metodu, která je nazývána lokální Laxova-Friedrichsova metoda. Současně se jedná o jednu z nejjednodušších metod, kterou lze odvodit přístupem používaným v případě tzv. centrálních metod Godunovova typu s vysokým rozlišením. Jde o vektorovou verzi lokální Laxovy-Friedrichsovy metody pro skalární nelineární případ.

Uvažujeme pouze největší velikost rychlosti šíření vlny (stejně jako ve skalárním případě nebereme v potaz rázové vlny a kontaktní nespojitosti). Označme tuto velikost rychlosti $s^{\max}_{j+1/2}$. Obvykle volíme

$$s^{\max}_{j+1/2} = \max_{i=1,\dots,m} \{|\lambda(\hat{\boldsymbol{U}}^-_{j+1/2})|, |\lambda(\hat{\boldsymbol{U}}^+_{j+1/2})|\}. \tag{3.115}$$

Touto volbou zajistíme, že nehladká část řešení Riemannova problému je uzavřena v podintervalu $\langle x_{j+1/2,L}, x_{j+1/2,R}\rangle$, kde $x_{j+1/2,L} = x_{j+1/2} - s^{\max}_{j+1/2}\Delta t$, $x_{j+1/2,R} = x_{j+1/2} + s^{\max}_{j+1/2}\Delta t$. Pak lze zapsat bilance v intervalech $\langle x_{j+1/2,L}, x_{j+1/2}\rangle$ a $\langle x_{j+1/2}, x_{j+1/2,R}\rangle$ takto

$$\boldsymbol{F}_{j+1/2} - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) = s^{\max}_{j+1/2}\left(\hat{\boldsymbol{U}}^*_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right), \tag{3.116}$$

$$\boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j+1/2}\right) - \boldsymbol{F}_{j+1/2} = s^{\max}_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^*_{j+1/2}\right). \tag{3.117}$$

Sečtením předchozích dvou rovností dostaneme

$$\boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j+1/2}\right) - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2})\right) = s^{\max}_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right). \tag{3.118}$$

Tuto metodu lze interpretovat jako přibližný Riemannův řešič s jediným mezistavem $\hat{\boldsymbol{U}}^*_{j+1/2}$ mezi $\hat{\boldsymbol{U}}^-_{j+1/2}$ a $\hat{\boldsymbol{U}}^+_{j+1/2}$. Z (3.116) a (3.117) také dostáváme

$$\boldsymbol{F}_{j+1/2} = \frac{1}{2}\left[\boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) + \boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j+1/2}\right)\right] - \frac{1}{2}s^{\max}_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right). \tag{3.119}$$

Hartenova-Laxova-Leerova metoda, metoda typu central-upwind

Jde o metodu, kterou lze chápat jako přibližný Riemannův řešič, obvykle nazývaný Hartenův-Laxův-Leerův, nebo jako centrální metodu Godunovova typu, která částečně využívá informaci o směru šíření vln. Opět jde o vektorovou verzi metody již popsané pro skalární nelineární případ. Uvažujeme pouze největší a nejmenší rychlosti šíření vln (stejně jako ve předchozím případě nebereme v potaz rázové vlny a kontaktní nespojitosti). Označme tyto rychlosti $s^{\min}_{j+1/2}$ a $s^{\max}_{j+1/2}$. Obvykle volíme

$$s^{\min}_{j+1/2} = \min_{i=1,\dots,m} \{\lambda(\hat{\boldsymbol{U}}^-_{j+1/2}), \lambda(\hat{\boldsymbol{U}}^+_{j+1/2}), 0\}, \tag{3.120}$$

$$s^{\max}_{j+1/2} = \max_{i=1,\dots,m} \{0, \lambda(\hat{\boldsymbol{U}}^-_{j+1/2})), \lambda(\hat{\boldsymbol{U}}^+_{j+1/2})\}. \tag{3.121}$$

Touto volbou zajistíme, že nehladká část řešení Riemannova problému je uzavřena v podintervalu $\langle x_{j+1/2,L}, x_{j+1/2,R}\rangle$, kde $x_{j+1/2,L} = x_{j+1/2} - s^{\min}_{j+1/2}\Delta t$, $x_{j+1/2,R} = x_{j+1/2} + s^{\max}_{j+1/2}\Delta t$. Pak lze zapsat bilance v intervalech $\langle x_{j+1/2,L}, x_{j+1/2}\rangle$ a $\langle x_{j+1/2}, x_{j+1/2,R}\rangle$ takto

$$\boldsymbol{F}_{j+1/2} - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) = s^{\min}_{j+1/2}\left(\hat{\boldsymbol{U}}^*_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right), \tag{3.122}$$

$$\boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j+1/2}\right) - \boldsymbol{F}_{j+1/2} = s^{\max}_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^*_{j+1/2}\right). \tag{3.123}$$

Sečtením předchozích dvou rovností tentokrát ovšem dostaneme

$$\boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j+1/2}\right) - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) = s^{\min}_{j+1/2}\left(\hat{\boldsymbol{U}}^*_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right) + s^{\max}_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^*_{j+1/2}\right).. \tag{3.124}$$

Tuto metodu lze interpretovat jako přibližný Riemannův řešič s jediným mezistavem $\hat{\boldsymbol{U}}^*_{j+1/2}$ mezi $\hat{\boldsymbol{U}}^-_{j+1/2}$ a $\hat{\boldsymbol{U}}^+_{j+1/2}$. Na rozdíl od centrální metody je zde však více respektován směr šíření vln (v některých speciálních případech dostaneme přímo metodu typu upwind).

Další verzí této metody je Hartenova-Laxova-Leerova-Einfeldtova metoda, která je totožná s HLL metodou až na volbu

$$s^{\min}_{j+1/2} = \min_{i=1,\dots,m}\{0, \lambda(\hat{\boldsymbol{U}}^-_{j+1/2}), \lambda(\hat{\boldsymbol{U}}^+_{j+1/2}), \lambda_{i,j+1/2}\}, \tag{3.125}$$

$$s^{\max}_{j+1/2} = \max_{i=1,\dots,m}\{0, \lambda(\hat{\boldsymbol{U}}^-_{j+1/2}), \lambda(\hat{\boldsymbol{U}}^+_{j+1/2}), \lambda_{i,j+1/2}\}, \tag{3.126}$$

kde $\lambda_{i,j+1/2}$ jsou vlastní čísla Roeovy matice. Právě uvedená modifikace dává obvykle ostřejší výsledky, než dostáváme v případě standardní HLL metody.

Z (3.122) a (3.123) dostáváme

$$\boldsymbol{F}_{j+1/2} = \frac{s^{\max}_{j+1/2}\boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) - s^{\min}_{j+1/2}\boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j+1/2}\right)}{s^{\max}_{j+1/2} - s^{\min}_{j+1/2}} - \frac{s^{\max}_{j+1/2}s^{\min}_{j+1/2}}{s^{\max}_{j+1/2} - s^{\min}_{j+1/2}}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right). \tag{3.127}$$

Další přibližné Riemannovy řešiče

Existuje celá řada přibližných Riemannových řešičů. Poznamenejme, že lze zkonstruovat verzi Engquistovy-Osherovy metody pro soustavy rovnic. Podrobnosti nalezne čtenář např. v [6]. Dále uvedeme pouze metody štěpení toku. Nejdříve však konstatujme, že všechny doposud popsané přibližné Riemannovy řešiče lze zapsat ve tvaru

$$\begin{aligned}\frac{\mathrm{d}\bar{\boldsymbol{U}}_j}{\mathrm{d}t} = -\tfrac{1}{\Delta x}\big[&\boldsymbol{A}^+(\hat{\boldsymbol{U}}^-_{j-1/2}, \hat{\boldsymbol{U}}^+_{j-1/2}) + \\ &+\boldsymbol{A}(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j-1/2}) + \boldsymbol{A}^+(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j+1/2})\big]\end{aligned} \tag{3.128}$$

Pro jednotlivé přibližné Riemannovy řešiče dostáváme:

- Roeova metoda

$$\boldsymbol{A}^+(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j+1/2}) = \sum_{i=1}^{m}\max\{0, \lambda_{i,j+1/2}\}\alpha_{i,j+1/2}\boldsymbol{r}_{i,j+1/2},$$

$$\boldsymbol{A}(\hat{\boldsymbol{U}}^+_{j-1/2}, \hat{\boldsymbol{U}}^-_{j+1/2}) = \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j-1/2}\right),$$

$$\boldsymbol{A}^-(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j+1/2}) = \sum_{i=1}^{m}\min\{0, \lambda_{i,j+1/2}\}\alpha_{i,j+1/2}\boldsymbol{r}_{i,j+1/2},$$

kde $\boldsymbol{\alpha}_{j+1/2} = \boldsymbol{R}^{-1}_{j+1/2}\left(\hat{\boldsymbol{U}}^+_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right)$

- Lokální Laxova-Friedrichsova metoda

$$\boldsymbol{A}^+(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j+1/2}) = s^{\max}_{j+1/2}\left(\hat{\boldsymbol{U}}^-_{j+1/2} - \hat{\boldsymbol{U}}^*_{j+1/2}\right),$$

$$\boldsymbol{A}(\hat{\boldsymbol{U}}^+_{j-1/2}, \hat{\boldsymbol{U}}^-_{j+1/2}) = \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j-1/2}\right),$$

$$\boldsymbol{A}^-(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j+1/2}) = -s^{\max}_{j+1/2}\left(\hat{\boldsymbol{U}}^*_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right),$$

- Hartenova-Laxova-Leerova metoda

$$\boldsymbol{A}^+(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j+1/2}) = s^{\max}_{j+1/2}\left(\hat{\boldsymbol{U}}^-_{j+1/2} - \hat{\boldsymbol{U}}^*_{j+1/2}\right),$$

$$\boldsymbol{A}(\hat{\boldsymbol{U}}^+_{j-1/2}, \hat{\boldsymbol{U}}^-_{j+1/2}) = \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j-1/2}\right),$$

$$\boldsymbol{A}^-(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j+1/2}) = s^{\min}_{j+1/2}\left(\hat{\boldsymbol{U}}^*_{j+1/2} - \hat{\boldsymbol{U}}^-_{j+1/2}\right),$$

Je zřejmé, že platí

$$\begin{aligned}&\boldsymbol{f}\left(\hat{\boldsymbol{U}}^+_{j+1/2}\right) - \boldsymbol{f}\left(\hat{\boldsymbol{U}}^-_{j+1/2}\right) = \\ &= \boldsymbol{A}^-_{j+1/2}\left(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j+1/2}\right) + \boldsymbol{A}^+_{j+1/2}\left(\hat{\boldsymbol{U}}^-_{j+1/2}, \hat{\boldsymbol{U}}^+_{j+1/2}\right),\end{aligned} \tag{3.129}$$

tj. rozdíl toků je rozštěpen na kladnou (šířící se doprava) a zápornou (šířící se doleva) část. Ve všech případech lze tak metodu (3.128) interpretovat pomocí štěpení diferencí toků takto: časová změna hledané funkce na kontrolním objemu $\langle x_{j-1/2}, x_{j+1/2}\rangle$ je dána diferencí toků v bodech $x_{j+1/2}$ a $x_{j-1/2}$, dále kladnou částí tokové diference v bodě $x_{j-1/2}$ a zápornou částí tokové diference v bodě $x_{j+1/2}$.

Další variantou je místo štěpení diference toků realizovat štěpení toku samotného. Zavedeme tedy štěpení

$$\boldsymbol{f}(\boldsymbol{u}) = \boldsymbol{f}^-(\boldsymbol{u}) + \boldsymbol{f}^+(\boldsymbol{u}). \tag{3.130}$$

Numerický tok zavádíme následujícím způsobem

$$\boldsymbol{F}_{j+1/2} = \boldsymbol{f}^+(\boldsymbol{U}^-_{j+1/2}) + \boldsymbol{f}^-(\boldsymbol{U}^+_{j+1/2}). \tag{3.131}$$

Základní otázkou je, jak volit $\boldsymbol{f}^-(\boldsymbol{u})$, $\boldsymbol{f}^+(\boldsymbol{u})$. V případě, že toková funkce je homogenní stupně 1, tj. pro libovolné $k \in \mathbb{R}$ platí $\boldsymbol{f}(k\boldsymbol{u}) = k\boldsymbol{f}(\boldsymbol{u})$, lze ukázat, že

$$\boldsymbol{f}(\boldsymbol{u}) = \boldsymbol{A}(\boldsymbol{u})\boldsymbol{u}, \quad \boldsymbol{A}(\boldsymbol{u}) = \boldsymbol{f}'(\boldsymbol{u}). \tag{3.132}$$

Pak lze zavést

$$\boldsymbol{f}^+(\boldsymbol{u}) = \boldsymbol{A}^+(\boldsymbol{u})\boldsymbol{u}, \quad \boldsymbol{A}^+(\boldsymbol{u}) = \boldsymbol{R}(\boldsymbol{u})\boldsymbol{\Lambda}^+(\boldsymbol{u})\boldsymbol{R}^{-1}(\boldsymbol{u}) \tag{3.133}$$

a

$$\boldsymbol{f}^-(\boldsymbol{u}) = \boldsymbol{A}^-(\boldsymbol{u})\boldsymbol{u}, \quad \boldsymbol{A}^-(\boldsymbol{u}) = \boldsymbol{R}(\boldsymbol{u})\boldsymbol{\Lambda}^-(\boldsymbol{u})\boldsymbol{R}^{-1}(\boldsymbol{u}). \tag{3.134}$$

Metoda založená na popsaném rozkladu se nazývá Stegerova-Warmingova metoda. Lze však realizovat i jiné rozklady, např. rozklad Vijayasundaramův ve tvaru

$$\boldsymbol{F}_{j+1/2} = \boldsymbol{A}^{+}\left(\frac{\boldsymbol{U}^{-}_{j+1/2}+\boldsymbol{U}^{+}_{j+1/2}}{2}\right)\boldsymbol{U}^{-}_{j+1/2} + \boldsymbol{A}^{-}\left(\frac{\boldsymbol{U}^{-}_{j+1/2}+\boldsymbol{U}^{+}_{j+1/2}}{2}\right)\boldsymbol{U}^{+}_{j+1/2}. \quad (3.135)$$

Chceme-li získat metodu s vysokým rozlišením je potřeba zvolit rekonstrukci hledané funkce vyššího řádu. Tu ovšem nelze realizovat po složkách. Musíme postupovat analogicky jako v případě Hartenovy hybridní metody, ve které přepínání toků nebylo realizováno pomocí skoků v aproximacích funkcí $\boldsymbol{u}$ soustavy

$$\boldsymbol{u}_t + \boldsymbol{A}\boldsymbol{u}_x = \boldsymbol{0},$$

ale pomocí skoků v aproximacích funkcí $\boldsymbol{v}$

$$\boldsymbol{v}_t + \boldsymbol{\Lambda}\boldsymbol{v}_x = \boldsymbol{0},$$

kde $\boldsymbol{v} = \boldsymbol{R}^{-1}\boldsymbol{u}$, $\boldsymbol{\Lambda} = \boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{R}$. Touto problematikou se zde již nebudeme zabývat. Pouze uvedeme, že v některých případech realizujeme rekonstrukci po složkách, přestože se jedná o postup, který není po teoretické stránce plně korektní.

Kapitola 4

Numerické metody pro parabolické rovnice

Průvodce studiem

V této kapitole se čtenář seznámí s parabolickými rovnicemi a některými metodami na jejich řešení. Jedná se o rovnice obsahující druhé prostorové derivace neznámé funkce. Budou ukázány rozdíly mezi explicitním a implicitním přístupem.

Cíle

Po prostudování této kapitoly budete schopni:

- popsat některé fyzikální děje pomocí parabolických rovnic
- konstruovat explicitní metody
- konstruovat implicitní metody
- posoudit vhodnost použití různých typů metod

4.1 Parabolické rovnice

Budeme se zde zabývat diferenčními metodami pro řešení počátečních úloh založených na parabolických rovnicích. Uvažujeme tedy i zde diskretizaci prostorové a časové proměnné a aproximaci neznámé funkce pomocí „uzlových“ hodnot U_j^n, $j = 0, 1, \dots, J$, $n = 0, 1, \dots, N$, jak bylo popsáno na začátku kapitoly 2.

Konstrukci numerických metod si přiblížíme na řešení evolučních úloh obsahujících difúzi. Nejjednodušším takovým příkladem je lineární počátečně-okrajová úloha

$$\begin{array}{rcll} u_t & = & Du_{xx}, & t \in (0,T), x \in (0,1), D \in \mathbb{R}^+, T > 0, \\ u(x,0) & = & u_0(x), & x \in \langle 0,1 \rangle, \\ u(0,t) & = & g_0(t), & u(1,t) = g_1(t), \qquad t \in (0,T) \end{array} \tag{4.1}$$

popisující například difúzi neznámé funkce $u = u(x,t)$, $D =$ konst. je difúzní koeficient. Tato rovnice však může rovněž popisovat šíření tepla v tenké tyči a další fyzikální jevy. V případě, že difúzní koeficient je závislý na prostorové proměnné $D = D(x)$, má počáteční úloha tvar

$$\begin{aligned} u_t &= (D(x)u_x)_x, \qquad t \in (0,T), x \in (0,1), D \in \mathbb{R}^+, T > 0, \\ u(x,0) &= u_0(x), \qquad x \in \langle 0,1\rangle, \\ u(0,t) &= g_0(t), \qquad u(1,t) = g_1(t), \qquad t \in (0,T). \end{aligned} \tag{4.2}$$

Tyto úlohy mohou mimo jiné popisovat rovněž přenos tepla v tenké tyči, kde koeficient D zastupuje tepelnou difuzivitu.

4.2 Explicitní metody

4.2.1 Eulerova metoda

Nejjednodušší metodu na řešení úloh popsaných v předchozím odstavci je možné sestrojit tak, že časovou i prostorovou derivaci nahradíme vhodným diferenčním podílem. Použijeme-li jednostrannou diferenci pro časovou proměnnou

$$u_t(x_j, t_n) = \frac{u_j^{n+1} - u_j^n}{\Delta t} + O(\Delta t) \tag{4.3}$$

a druhou poměrnou diferenci pro prostorovou proměnnou

$$u_{xx}(x_j, t_n) = \frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{\Delta x^2} + O(\Delta x^2) \tag{4.4}$$

získáme diferenční rovnost

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = D\frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{\Delta x^2} \tag{4.5}$$

pro $j = 1, \ldots, J-1$. Z této rovnosti (4.5) je možné jednoduše získat explicitní vztah pro hodnotu na nové časové vrstvě t_{n+1}

$$U_j^{n+1} = rU_{j-1}^n + (1-2r)U_j^n + rU_{j+1}^n, \tag{4.6}$$

kde pro jednoduchost zápisu zavádíme parametr r

$$r = \frac{D\Delta t}{\Delta x^2}. \tag{4.7}$$

Explicitní diferenční formule metod lze obecně zapsat ve tvaru

$$\boldsymbol{U}^{(n+1)} = \boldsymbol{\mathcal{H}}(\boldsymbol{U}^{(n)}), \tag{4.8}$$

kde $\boldsymbol{\mathcal{H}}$ je operátor přechodu z časové vrstvy n na vrstvu $n+1$. Lineární diferenční schéma pak lze psát ve tvaru

$$\boldsymbol{U}^{(n+1)} = \boldsymbol{H}\boldsymbol{U}^{(n)}, \tag{4.9}$$

kde $\boldsymbol{H}$ je lineární operátor . Formálně má tedy diferenční formule stejný tvar, jako v případě hyperbolických schémat popisovaných v kapitole 2, proto i některé základní definice jsou shodné.

Definice 4.1. *Lokální diskretizační chybu* $\boldsymbol{L}^{(n)}$ metody (4.8) definujeme předpisem

$$\boldsymbol{L}^{(n)} = \frac{1}{\Delta t}\left[\boldsymbol{U}^{(n+1)} - \mathcal{H}\boldsymbol{U}^{(n)}\right], \tag{4.10}$$

kde $\boldsymbol{U}^{(n)}$ je posloupnost o členech $u_j^n = u(x_j, t_n)$.

Definice 4.2. Diferenční formule se nazývá *konzistentní* s příslušnou aproximovanou rovnicí, jestliže pro všechna j a n taková, že $T \geqq n\Delta t$ platí

$$L_j^n \to 0 \text{ pro } \Delta t \to 0. \tag{4.11}$$

Definice 4.3. Největší přirozené číslo p, pro které existuje konstanta $C_L > 0$ taková, že pro všechna j a n taková, že $T \geqq n\Delta t$ platí

$$L_j^n \leqq C_L(\Delta t)^p \text{ pro } \Delta t \to 0, \tag{4.12}$$

se nazývá *řád konzistence.*

Podle předchozích definic lze ukázat, že metoda (4.6) aproximuje diferenciální rovnici (4.1) s přesností $O(\Delta t + \Delta x^2)$. Lokální diskretizační chyba je dána výrazem

$$L_j^n = \frac{u_j^{n+1} - u_j^n}{\Delta t} - D\frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{\Delta x}. \tag{4.13}$$

Předpokládáme hladkou funkci u a použijeme Taylorův rozvoj k určení hodnot řešení v čase t_{n+1}. Potom lze zapsat

$$L_j^n = \left(u_t + \frac{1}{2}\Delta t u_{tt} + \frac{1}{6}\Delta t^2 u_{ttt} + \dots\right) - D\left(u_{xx} + \frac{1}{12}\Delta x^2 u_{xxxx} + \dots\right). \tag{4.14}$$

Jelikož $u_t = Du_{xx}$ snadno zjistíme, že $u_{tt} = Du_{xxt} = Du_{xxxx}$ a proto

$$L_j^n = \left(\frac{1}{2}\Delta t - \frac{1}{12}\Delta x^2\right) u_{xxxx} + O(\Delta t^2 + \Delta x^4). \tag{4.15}$$

4.2.2 Stabilita explicitní metody

Obdobně jako u explicitních metod pro hyperbolické úlohy i zde je explicitní metoda stabilní vzhledem k počáteční podmínce pouze pro některé délky časového kroku.

Definice 4.4. Řekneme, že lineární metoda (4.9) je *lineárně stabilní,* je-li posloupnost $\{\|\boldsymbol{H}^n\|\}$ omezená, tzn. pro každé T existuje číslo $C_S > 0$ takové, že platí

$$\|\boldsymbol{H}^n\| \leqq C_S \tag{4.16}$$

pro všechna n taková, že $n\Delta t \leqq T$.

Příklad 4.5. Řešme explicitní metodou (4.6) počátečně-okrajovou úlohu (4.1), kde $D = 1$ a počáteční podmínka je nulová všude s výjimkou jediného uzlu, kde je hodnota neznámé funkce rovna jedné. Okrajové podmínky jsou také nulové $g_0(t) = g_1(t) = 0$. Zvolme

$$r = 1, \quad \text{tj. } \Delta t = \frac{\Delta x^2}{D}. \tag{4.17}$$

Řešení. Metoda má v tomto případě tvar

$$U_j^{n+1} = U_{j-1}^n - U_j^n + U_{j+1}^n. \tag{4.18}$$

Řešení získané pomocí této metody je následující

$$\begin{array}{lcccccccccccc} U^0 = & 0 & \dots & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ U^1 = & 0 & \dots & 0 & 0 & 1 & -1 & 1 & 0 & 0 & \dots & 0 \\ U^2 = & 0 & \dots & 0 & 1 & -2 & 3 & -2 & 1 & 0 & \dots & 0 \end{array} \tag{4.19}$$

Z tohoto je patrné, že řešení nesplňuje princip maxima (maximální hodnoty by řešení mělo nabývat na krajích intervalu nebo v počátku v čase t_0). Dochází tak z zesilování chyb v počáteční podmínce a řešení se chová nestabilně. ▲

Jak jsme viděli v předchozím příkladu, je třeba sestavit takovou metodu, aby během výpočtu nedocházelo k zesilování chyb v počáteční podmínce. Pro jednoduchost uvažujme diskretizovanou úlohu, která obsahuje pouze tři uzly x_0, x_1, x_2. Opět volíme nulovou okrajovou podmínku $U_0 = U_2 = 0$. Počáteční hodnota v prostředním uzlu je zvolena jedna $U_1^0 = 1$. Metoda bude tedy stabilní, pokud bude platit, že $0 \leqq U_1^n \leqq 1$ pro všechna n. Numerická metoda pro tuto zjednodušenou úlohu má tvar

$$U_1^{n+1} = (1 - 2r)U_1^n \tag{4.20}$$

V čase t_n získáme řešení

$$U_1^n = (1 - 2r)U_1^{n-1} = \ldots = (1 - 2r)^n U - 1^0 = (1 - 2r)^n \tag{4.21}$$

Aby $0 \leqq U_1^n \leqq 1$, musí platí $0 \leqq (1-2r)^n \leqq 1$. Z tohoto plyne požadavek

$$0 \leqq r \leqq \frac{1}{2} \tag{4.22}$$

Lze ukázat, že tato podmínka (4.22) je postačující podmínkou stability explicitní metody i pro obecnější úlohy. délka časového kroku je tedy omezena vztahem

$$0 \leqq \Delta t \leqq \frac{\Delta x^2}{2D} \tag{4.23}$$

Tato ukázka vyšetřování stability byla zjednodušením Fourierovy metody, která je založena na studiu šíření chyby v počátečních datech během výpočtu. Tato chyba je přitom vyjádřena ve formě Fourierovy řady.

Fourierova metoda vyšetřování stability

Předpokládejme, že chybu ve vstupních datech lze vyjádřit ve tvaru konečné Fourierovy řady

$$\sum_{m=0}^{M} c_m \mathrm{e}^{i\nu_m x}. \tag{4.24}$$

Šíření této chyby během výpočtu budeme analyzovat pomocí jediného členu $\mathrm{e}^{i\nu x}$, kde $\nu \in \mathbb{R}$. Počáteční chyba v bodě x_j má tedy tvar

$$Z_j^0 = \mathrm{e}^{i\nu x_j} = \mathrm{e}^{i\nu j\Delta x}, \qquad j = 0, 1, \ldots, J. \tag{4.25}$$

Jelikož řešená počáteční úloha je lineární, platí pro tuto chybu stejná diferenční rovnice (4.6) jako pro přibližné řešení U_j^n. V čase t_n pak danou chybu lze na základě Fourierovy metody řešení diferenciálních rovnic vyjádřit ve tvaru

$$Z_j^n = q^n Z_j^0 = q^n \mathrm{e}^{i\nu j\Delta x}, \tag{4.26}$$

kde $q = q(\nu)$. Výraz (4.26) dosadíme do (4.6) a upravíme

$$\begin{aligned}
q^{n+1}\mathrm{e}^{i\nu j\Delta x} &= rq^n\mathrm{e}^{i\nu(j-1)\Delta x} + (1-2r)q^n\mathrm{e}^{i\nu j\Delta x} + rq^n\mathrm{e}^{i\nu(j+1)\Delta x}\\
q^n\mathrm{e}^{i\nu(j-1)\Delta x}q\mathrm{e}^{i\nu\Delta x} &= q^n\mathrm{e}^{i\nu(j-1)\Delta x}r + q^n\mathrm{e}^{i\nu(j-1)\Delta x}(1-2r)\mathrm{e}^{i\nu\Delta x} + q^n\mathrm{e}^{i\nu(j-1)\Delta x}r\mathrm{e}^{i\nu\Delta x}\\
q\mathrm{e}^{i\nu\Delta x} &= r + (1-2r)\mathrm{e}^{i\nu\Delta x} + r\mathrm{e}^{i\nu\Delta x}\\
q &= r\mathrm{e}^{-i\nu\Delta x} + (1-2r) + r\mathrm{e}^{i\nu\Delta x}\\
q &= 2r\cos(\nu\Delta x) + 1 - 2r\\
q &= 1 - 2r(1-\cos(\nu\Delta x)\\
q &= 1 - 4r\sin^2\frac{\nu\Delta x}{2}
\end{aligned}$$

Aby se chyba během výpočtu nezvětšovala, požadujeme, aby $|q| \leqq 1$ pro všechna $\nu \in \mathbb{R}$. Jednoduchými úpravami uvidíme, že tento požadavek bude splněn při platnosti nerovnosti

$$0 \leqq r \leqq \frac{1}{2\sin^2\frac{\nu\Delta x}{2}}. \tag{4.27}$$

Z toho plyne $0 \leqq r \leqq \frac{1}{2}$ a omezení časového kroku (4.23).

Maticová analýza stability

Dalším způsobem vyšetření stability je maticová analýza, kdy postupujeme přímo ve smyslu definice 4.4, přesněji řečeno na základě vlastností matice $\boldsymbol{H}$. Chybu v datech lze vyjádřit jako rozdíl přesné hodnoty a hodnoty zatížené chybou $\boldsymbol{Z} = \boldsymbol{U} - \tilde{\boldsymbol{U}}$. Aby se tato chyba nezvětšovala, musí v lineárním případě platit

$$\|\boldsymbol{H}(\boldsymbol{U}) - \boldsymbol{H}(\tilde{\boldsymbol{U}})\| = \|\boldsymbol{H}(\boldsymbol{U} - \tilde{\boldsymbol{U}})\| \leqq \|\boldsymbol{H}\|\|\boldsymbol{U} - \tilde{\boldsymbol{U}}\|. \tag{4.28}$$

Za stabilní metodu lze tedy považovat takovou, pro kterou vhodně zvolená norma $\|\boldsymbol{H}\| \leqq 1$. Diferenční schéma (4.6) lze vyjádřit v maticovém tvaru (uvažujeme-li okrajové podmínky 4.1)

$$\begin{bmatrix} U_1^{n+1} \\ U_2^{n+1} \\ \vdots \\ U_{J-2}^{n+1} \\ U_{J-1}^{n+1} \end{bmatrix} = \begin{bmatrix} 1-2r & r & 0 & \dots & 0 \\ r & 1-2r & r & \dots & 0 \\ \vdots & & \ddots & \dots & 0 \\ 0 & \dots & r & 1-2r & r \\ 0 & \dots & 0 & r & 1-2r \end{bmatrix} \begin{bmatrix} U_1^n \\ U_2^n \\ \vdots \\ U_{J-2}^n \\ U_{J-1}^n \end{bmatrix} + \begin{bmatrix} rg_0(t_n) \\ 0 \\ \vdots \\ 0 \\ rg_1(t_n) \end{bmatrix}, \tag{4.29}$$

zkráceně

$$\boldsymbol{U}^{n+1} = \boldsymbol{H}\boldsymbol{U}^n + \boldsymbol{F}^n. \tag{4.30}$$

Pro symetrické matice lze zvolit například $\|\boldsymbol{H}\| = \max\limits_k |\lambda_k|$. Vlastní čísla matice $\boldsymbol{H}$ jsou

$$\lambda_k = 1 - 4r\sin^2\frac{k\pi}{2J}. \tag{4.31}$$

Požadujeme, aby $\max\limits_k |\lambda_k| \leqq 1$, z čehož plyne, že $0 \leqq r \leqq \frac{1}{2}$ a omezení časového kroku je opět (4.23).

Důležitou otázkou u numerických metod je konvergence přibližného řešení k přesnému. O konvergenci založené na konzistenci a stabilitě hovoří Laxova věta [1], zde si uvedeme její speciální tvar vhodný pro zde popisované metody.

Věta 4.6. *[o konvergenci]* Nechť má počátečně-okrajová úloha (4.1) řešení u, které je čtyřikrát spojitě diferencovatelné podle x a dvakrát podle t. Označme U_j^n přibližné řešení této úlohy získané explicitní metodou (4.6) v bodě x_j a čase t_n a platí $0 \leqq r \leqq \frac{1}{2}$. Pak pro normu chyby přibližného řešení platí odhad

$$\|u_j^n - U_j^n\| \equiv \max_{j,n} |u_j^n - U_j^n| = O(\Delta t + \Delta x^2). \tag{4.32}$$

Důkaz.

$$u_j^{n+1} = ru_{j-1}^n + (1-2r)u_j^n + ru_{j+1}^n + \Delta t \cdot O(\Delta t + \Delta x^2) \tag{4.33}$$

pro $j = 1, \dots, J-1$. Zavedeme chybu pro přibližné řešení $\mathrm{e}_j^n = u_j^n - U_j^n$. Jelikož hodnoty přibližného řešení splňují diferenciální rovnici (4.1) přesně, pro takto zavedenou chybu platí vztah

$$\mathrm{e}_j^{n+1} = r\mathrm{e}_{j-1}^n + (1-2r)\mathrm{e}_j^n + r\mathrm{e}_{j+1}^n + \Delta t \cdot O(\Delta t + \Delta x^2) \tag{4.34}$$

pro $j = 1, \dots, J-1$. Na hranici obdélníku $\langle 0, 1\rangle \times \langle 0, T\rangle$ platí

$$\begin{aligned} \mathrm{e}_j^n &= 0, j = 0, \dots, J \\ \mathrm{e}_0^n = \mathrm{e}_J^n &= 0, n = 0, \dots, N. \end{aligned} \tag{4.35}$$

Pokud je splněna podmínka stability, tj. $0 < r \leqq \frac{1}{2}$, jsou všechny tři koeficienty na pravé straně (4.34) nezáporné a jejich součet je roven jedné. Proto platí odhad

$$\begin{aligned} |\mathrm{e}_j^{n+1}| \leqq r|\mathrm{e}_{j-1}^n| + (1-2r)|\mathrm{e}_j^n| + r|\mathrm{e}_{j+1}^n| + \Delta t \cdot O(\Delta t + \Delta x^2) \leqq \\ \leqq \|\mathrm{e}^n\| + \Delta t \cdot O(\Delta t + \Delta x^2), \end{aligned} \tag{4.36}$$

kde

$$\|\mathrm{e}^n\| = \max_j |\mathrm{e}_j^n|. \tag{4.37}$$

Vzhledem k (4.35) platí také

$$\|\mathrm{e}^{n+1}\| \leqq \|\mathrm{e}^n\| + \Delta t \cdot O(\Delta t + \Delta x^2) \tag{4.38}$$

Vzhledem k evoluci řešení v čase t_n a k tomu, že $\|\mathrm{e}^0\| = 0$ máme

$$\begin{aligned} \|\mathrm{e}^n\| \leqq \|\mathrm{e}^{n-1}\| + \Delta t \cdot O(\Delta t + \Delta x^2) \leqq \dots \leqq \|\mathrm{e}^0\| + n\Delta t \cdot O(\Delta t + \Delta x^2) = \\ = T \cdot O(\Delta t + \Delta x^2) = O(\Delta t + \Delta x^2). \end{aligned} \tag{4.39}$$

□

Jako varovný příklad nestabilní numerické metody nezávisle na hodnotě parametru r si uveďme následující metodu

$$\frac{U_j^{n+1} - U_j^{n-1}}{2\Delta t} = D\frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{\Delta x^2}. \tag{4.40}$$

Oproti metodě (4.5) jsme časovou jednostrannou diferenci nahradili diferencí centrální. Tím jsme docílili zvýšení řádu přesnosti na $O(\Delta t^2 + \Delta x^2)$. Přesto, že se jedná o „malou“ změnu, výsledná numerická metoda je nestabilní a to pro libovolnou hodnotu parametru r. Z toho je vidět, že diskretizaci diferenciálních úloh nelze dělat zcela libovolně, ale třeba se zajímat o teoretické vlastnosti sestrojených metod.

4.3 Implicitní metody

Výhodu explicitních metod je především jejich jednoduchá algoritmizace díky explicitnímu vyjádření hodnot aproximace neznámé funkce na nové časové vrstvě. Nevýhodou je však jejich podmíněná stabilita, která zapříčiňuje omezení délky časového (viz (4.23), navíc si všimněme kvadratické závislosti Δt na velikosti Δx), které není například s ohledem na charakter řešení nutné. To pak výrazně zpomaluje výpočet, neboť je třeba hledat řešení na příliš velkém množství časových vrstev.

Naším cílem tedy bude sestrojit metodu, která bude bezpodmínečně stabilní a bude tak možné volit délku časového kroku nezávisle na zachování stability metody. Takovou metodu můžeme sestrojit například tak, že použijeme postup použitý k sestrojení metody (4.5) pouze s tím rozdílem, že druhou prostorovou derivaci nebudeme aproximovat na časové vrstvě t_n, ale na vrstvě t_{n+1}. Výsledkem je diferenční metoda

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = D\frac{U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}}{\Delta x^2}, \tag{4.41}$$

pro $j = 1, \dots, J-1$, jejíž diskretizační chyba je opět $O(\Delta t + \Delta x^2)$. Vztah pro výpočet řešení na nové časové vrstvě t_{n+1} je však nyní implicitní, což vede ke zkomplikování výpočtu. Po úpravě získáme

$$-rU_{j-1}^{n+1} + (1+2r)U_j^{n+1} - rU_{j+1}^{n+1} = U_j^n, \tag{4.42}$$

pro $j = 1, \dots, J-1$, kde r je opět definováno výrazem (4.7). Tyto rovnice, kterých je o dvě méně než neznámých, je ještě třeba doplnit o okrajové podmínky

$$U_0^{n+1} = g_0(t_{n+1}), \qquad U_J^{n+1} = g_1(t_{n+1}). \tag{4.43}$$

Potom lze danou soustavu pro neznámé hodnoty $U_1^{n+1}, \dots, U_{J-1}^{n+1}$ zapsat v maticové podobě

$$\begin{bmatrix} 1+2r & -r & 0 & \dots & 0 \\ -r & 1+2r & -r & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -r & 1+2r & -r \\ 0 & \dots & 0 & -r & 1+2r \end{bmatrix} \begin{bmatrix} U_1^{n+1} \\ U_2^{n+1} \\ \vdots \\ U_{J-2}^{n+1} \\ U_{J-1}^{n+1} \end{bmatrix} = \begin{bmatrix} U_1^n \\ U_2^n \\ \vdots \\ U_{J-2}^n \\ U_{J-1}^n \end{bmatrix} + \begin{bmatrix} rg_0(t_{n+1}) \\ 0 \\ \vdots \\ 0 \\ rg_1(t_{n+1}) \end{bmatrix}, \tag{4.44}$$

zkráceně

$$\boldsymbol{A}\boldsymbol{U}^{n+1} = \boldsymbol{U}^n + \boldsymbol{F}^{n+1}. \tag{4.45}$$

Této metodě se říká ryze implicitní metoda. Jak je patrné, na každé časové vrstvě musíme řešit soustavu lineárních algebraických rovnic. Tyto soustavy jsou jednoznačně řešitelné, neboť matice $\boldsymbol{A}$ je ostře diagonálně dominantní. Podstatnou výhodou této metody je její bezpodmínečná stabilita vzhledem k počáteční podmínce. Nejsou kladeny žádné požadavky na velikost parametru r a tím ani na délku časového kroku Δt.

4.3.1 Crankova-Nicolsonova metoda

Nyní bude naším cílem sestrojit metodu, která bude bezpodmínečně stabilní a navíc bude druhého řádu přesnosti i v časové proměnné. Toho lze docílit například tak, že jednostrannou časovou diferenci s chybou $O(\Delta t)$ nahradíme centrální diferencí s chybou $O(\Delta t^2)$. Bude však třeba postupovat jinak než v případě (4.40), kde byla výsledná metoda nestabilní.

Diferenční podíl $\frac{u(x_j,t_{n+1})-u(x_j,t_n)}{\Delta t}$ aproximuje pro dostatečně hladké funkce $u(x,t)$ sice derivace $u_t(x_j,t_n)$ a $u_t(x_j,t_{n+1})$ s přesností řádu $O(\Delta t)$, ale zároveň aproximuje derivaci $u_t(x_j,t_{n+1/2})$ s přesností řádu $O(\Delta t^2)$. Z tohoto důvodu, budeme diferenciální rovnici aproximovat v „mezivrstvě“ $t_{n+1/2}$. Kromě časové diference je třeba zvolit ještě aproximaci druhé prostorové derivace u_{xx}. Máme k dispozici pouze hodnoty v uzlech diskretizační sítě, tedy na časových vrstvách t_n a t_{n+1}). Rozumnou volbou je tak průměr druhých poměrných diferencích na těchto sousedních vrstvách. Získáme tak diferenční rovnici

$$\frac{U_j^{n+1}-U_j^n}{\Delta t}=\frac{1}{2}\left[D\frac{U_{j+1}^{n+1}-2U_j^{n+1}+U_{j-1}^{n+1}}{\Delta x^2}+D\frac{U_{j+1}^{n}-2U_j^{n}+U_{j-1}^{n}}{\Delta x^2}\right] \tag{4.46}$$

pro $j=1,\dots,J-1$. Lze ukázat, že toto schéma je skutečně řádu přesnosti $O(\Delta t^2 + \Delta x^2)$ a je bezpodmínečně stabilní vzhledem k počáteční podmínce. Jednoduchou úpravou ihned uvidíme, že se opět jedná o implicitní schéma

$$-\frac{1}{2}rU_{j-1}^{n+1}+(1+r)U_j^{n+1}-\frac{1}{2}rU_{j+1}^{n+1}=\frac{1}{2}rU_{j-1}^{n+1}+(1-r)U_j^{n+1}+\frac{1}{2}rU_{j+1}^{n+1} \tag{4.47}$$

pro $j=1,\dots,J-1$. Stejně jako u ryze implicitní metody je třeba doplnit rovnice o vztahy z okrajových podmínek

$$U_0^{n+1}=g_0(t_{n+1}),\qquad U_J^{n+1}=g_1(t_{n+1}) \tag{4.48}$$

$$U_0^{n}=g_0(t_{n}),\qquad U_J^{n}=g_1(t_{n}) \tag{4.49}$$

Vektor řešení na nové časové vrstvě $\boldsymbol{U}^{n+1}=[U_1^{n+1},\dots,U_{J-1}^{n+1}]$ pak opět získáme vyřešením soustavy lineárních algebraických rovnic ve tvaru

$$\boldsymbol{A}\boldsymbol{U}^{n+1}=\boldsymbol{B}\boldsymbol{U}^n+\boldsymbol{F}^{n+1}, \tag{4.50}$$

kde $\boldsymbol{A}$ je třídiagonální symetrická matice řádu $J-1$ tvaru

$$\boldsymbol{A}=\begin{bmatrix} 1+r & -\frac{1}{2}r & 0 & \dots & 0\\ -\frac{1}{2}r & 1+r & -\frac{1}{2}r & \ddots & \vdots\\ 0 & \ddots & \ddots & \ddots & 0\\ \vdots & \ddots & -\frac{1}{2}r & 1+r & -\frac{1}{2}r\\ 0 & \dots & 0 & -\frac{1}{2}r & 1+r \end{bmatrix}, \tag{4.51}$$

$\boldsymbol{B}$ je třídiagonální symetrická matice řádu $J-1$ tvaru

$$\boldsymbol{B} = \begin{bmatrix} 1-r & \frac{1}{2}r & 0 & \dots & 0 \\ \frac{1}{2}r & 1-r & \frac{1}{2}r & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \frac{1}{2}r & 1-r & \frac{1}{2}r \\ 0 & \dots & 0 & \frac{1}{2}r & 1-r \end{bmatrix} \tag{4.52}$$

a

$$\boldsymbol{F} = \left[\frac{1}{2}(g_0(t_{n+1}) + g_0(t_n)), 0, \dots, 0, \frac{1}{2}(g_1(t_{n+1}) + g_1(t_n))\right]^T. \tag{4.53}$$

Matice $\boldsymbol{A}$ je opět ostře diagonální dominantní, proto je soustava na každé časové vrstvě jednoznačně řešitelná. Zkonstruovaná metoda s nazývá Crankova-Nicolsonova metoda je bezpodmínečně stabilní a druhého řádu přesnosti v časové i prostorové proměnné.

4.3.2 Theta metoda

Čtenář se může setkat i s dvouvrstvými metodami, které se konstruují obdobně jako Crankova-Nicolsonova metoda s tím rozdílem, že druhá derivace u_{xx} se aproximuje nikoli průměrem, ale váženým průměrem diferencí na sousedních časových vrstvách s váhami θ a $1-\theta$, $0 \leqq \theta \leqq 1$. Získáme tak diferenční rovnici

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \theta \frac{U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}}{\Delta x^2} + (1-\theta) D \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{\Delta x^2}. \tag{4.54}$$

pro $j = 1, \dots, J-1$. Pro některé hodnoty parametru θ získáme již dříve konstruované metody (viz tabulka 4.1). Z toho je patrné, že na volbě parametru θ závisí vlastnosti dané metody jako je stabilita a velikost chyby aproximace.

θ	metoda	stabilní	chyba
0	explicitní	podmínečně	$O(\Delta t + \Delta x^2)$
$\frac{1}{2}$	Crankova-Nicolsonova	bezpodmínečně	$O(\Delta t^2 + \Delta x^2)$
1	ryze implicitní	bezpodmínečně	$O(\Delta t + \Delta x^2)$

Tab. 4.1 Dvoukrokové metody

4.4 Semidiskrétní metody

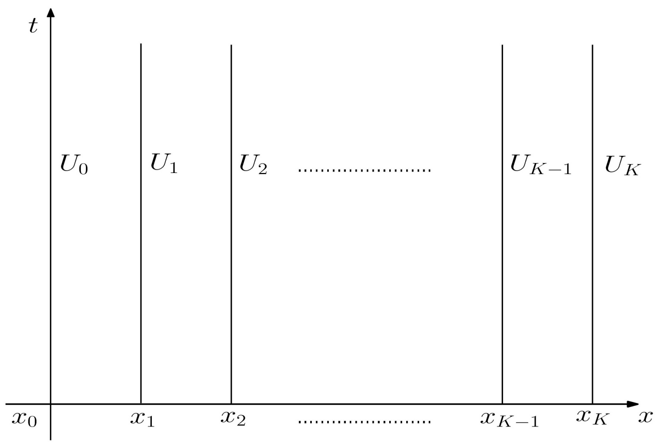
V této části se zaměříme na metody, které diskretizují jednotlivé proměnné parabolických úloh s výjimkou jedné. Tím vzniká systém obyčejných diferenciálních rovnic, které lze řešit příslušnými standardními algoritmy a metodami, kterých existuje celá řada.

4.4.1 Metoda přímek

Tato metoda pro evoluční úlohy diskretizuje všechny proměnné vyjma časové proměnné. Diskretizujeme prostorovou proměnou a časovou ponecháme spojitou. Označíme $u_j(t) = u(x_j, t)$ restrikci funkce $u(x,t)$ na polopřímce $x = x_k$ (odtud název metoda přímek). Získáme tak soustavu obyčejných diferenčních rovnic

$$\frac{\partial U_j}{\partial t} = F_j(t, U_{j-1}, U_j, U_{j+1}), \tag{4.55}$$

pro $j = 1, \ldots, K-1$, kde $F_j(U_{j-1}, U_j, U_{j+1})$ je diskretizační operátor, který obsahuje poměrné prostorové diference, a který může mít tvar jako v případě implicitní metody, Crankovy-Nicolsonovy metody, atd.



Obr. 4.1 Diskretizace metody přímek

Hodnoty $U_0(t)$ a $U_K(t)$ jsou dány pomocí okrajových podmínek, čímž získáme stejný počet rovnic jako neznámých. Počáteční podmínky jsou zřejmě dány $U_j(0) = \phi(x_j)$.

Vlastnosti vzniklých počátečních úloh

Počáteční úlohy vzniklé na základě metody přímek jsou často úlohy s tzv. velkým tlumením (říkáme také stiff-systém). V podstatě to znamená, že řešení obsahuje složky s různými nároky na délku časového kroku kvůli stabilitě. Tuto skutečnost je třeba mít na zřeteli při numerickém řešení. Například jsme si již ukázali, že vlastní čísla matice $\boldsymbol{H}$ úlohy 4.1 jsou (viz(4.29)

$$\lambda_k = 1 - 4r \sin^2 \frac{k\pi}{2J}, \qquad k = 1, 2, \ldots, J-1. \tag{4.56}$$

Podobně u metody přímek, kterou můžeme zapsat ve tvaru

$$\boldsymbol{U}' = \frac{1}{\Delta x^2} \boldsymbol{A}\boldsymbol{U}, \tag{4.57}$$

kde matice $\boldsymbol{A}$ má tvar

$$\boldsymbol{A} = \begin{bmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & -2 & 1 \\ 0 & \dots & 0 & 1 & -2 \end{bmatrix} \tag{4.58}$$

a její vlastní čísla jsou

$$\lambda_k = -4\sin^2\frac{k\pi\Delta x}{2}, \qquad k = 1, 2, \dots, J-1. \tag{4.59}$$

Je zřejmé, že $0 < \lambda_1 < \lambda_2 < \dots < \lambda_{J-1}$ a uvažujeme-li, že hledáme řešení na intervalu $x \in \langle 0, 1\rangle$, platí $J\Delta x = 1$. Potom při $\Delta x \to 0$ velikosti prvního a posledního vlastní čísla jsou

$$|\lambda_1| = 4\sin^2\frac{\pi\Delta x}{2}, \qquad |\lambda_{J-1}| = 4\cos^2\frac{\pi\Delta x}{2} \tag{4.60}$$

a jejich podíl se tak blíží $+\infty$. Jde tedy o typický příklad stiff-systému.

4.4.2 Metoda časové diskretizace

V této metodě diskretizujeme pouze časovou proměnnou, prostorové necháme spojité. Označíme $u^n(x) = u(x, t_n)$ restrikci funkce $u(x, t)$ na úsečce $t = t_n$. Uvažujme pro jednoduchost lineární rovnici

$$u_t = Du_{xx}. \tag{4.61}$$

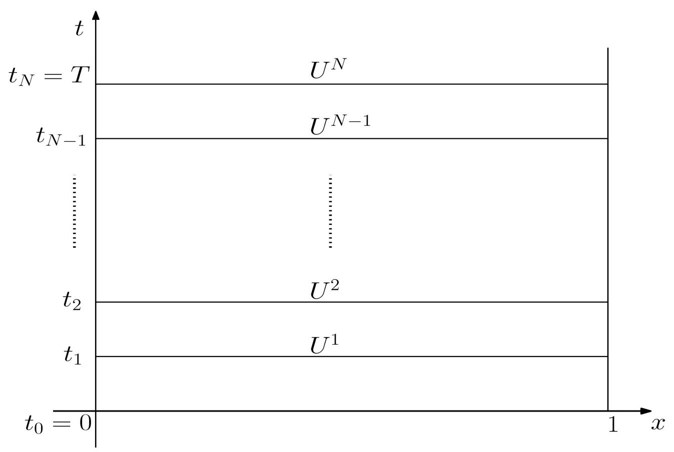
Použijeme jednostrannou diferenci na časovou proměnnou a zmíněné restrikce na jednotlivé časové vrstvy, získáme tak posloupnost semidiskrétních vztahů

$$\frac{U^{n+1} - U^n}{\Delta t} = D(U^{n+1})_{xx}. \tag{4.62}$$

Tento vztah upravíme, a získáme posloupnost okrajových úloh pro obyčejné diferenciální rovnice druhého řádu

$$U^{n+1} - \Delta t D(U^{n+1})_{xx} = U^n \tag{4.63}$$

Funkce U^0 je opět dána počáteční podmínkou. Postupně tak získáme řešení na časových vrstvách s vyšším indexem pomocí řešení okrajové úlohy z předcházející vrstvy.



Obr. 4.2 Diskretizace metody časové diskretizace

Kapitola 5

Numerické Metody štěpení

Průvodce studiem

V této kapitole se čtenář seznámí s metodami štěpení. Tyto metody jsou konstruovány pro řešení úloh ve více prostorových dimenzích úloh, které obsahují zdrojové členy a úloh obsahujících zároveň advekční, difúzní, či další členy různého typu. Budou uvedeny přístupy založená na štěpení na více jednodimenzionálních úloh.

Cíle

Po prostudování této kapitoly budete schopni:

- porozumět pojmu štěpení a jeho využití
- konstruovat lokálně jednodimenzionální metody
- konstruovat metody střídavých směrů

5.1 Proč štěpení

Pro řadu úloh založených na parciálních diferenciálních rovnicích může být nemožné či neefektivní použití jednoho přístupu pro řešení komplexního problému. Příkladem mohou být úlohy s advekčními i difuzními členy, kdy je lépe diskretizovat každou část jiným způsobem. Stejně tak použití implicitního přístupu na řešení kompletní diferenciální rovnice, může vést na řešení příliš velké nelineární algebraické soustavy. Z tohoto pohledu je velice efektivní použití některé metody štěpení, která umožní řešit každou část problému efektivním způsobem.

Předpokládejme tedy, že řešíme úlohu založenou na parciální diferenciální rovnici ve tvaru $\boldsymbol{u}_t = \boldsymbol{h}(\boldsymbol{u}, t)$, kde lze operátor na pravé straně štěpit na s operátorů

$$\boldsymbol{h}(\boldsymbol{u}, t) = \boldsymbol{h}_1(\boldsymbol{u}, t) + \boldsymbol{h}_2(\boldsymbol{u}, t) + \ldots + \boldsymbol{h}_s(\boldsymbol{u}, t) \tag{5.1}$$

Metody štěpení lze například využít na úlohy se zdrojovými členy. Jako příklad uveďme nehomogenní rovnici hyperbolického typu ve tvaru

$$\boldsymbol{u}_t + [\boldsymbol{f}(\boldsymbol{u})]_x = \boldsymbol{\psi}(\boldsymbol{u}, x). \tag{5.2}$$

Tuto úlohu lze řešit ve dvou krocích. Na každé časové vrstvě nejprve vyřešíme homogenní rovnici a získané řešení použijeme jako počáteční podmínku pro druhý krok, kde řešíme obyčejnou diferenciální rovnici

1. $\boldsymbol{u}_t + [\boldsymbol{f}(\boldsymbol{u})]_x = \boldsymbol{0}$,

2. $\boldsymbol{u}_t = \boldsymbol{\psi}(\boldsymbol{u}, x)$.

Zde tedy máme $\boldsymbol{h}_1(\boldsymbol{u}, t) = -[\boldsymbol{f}(\boldsymbol{u})]_x$ a $\boldsymbol{h}_2(\boldsymbol{u}, t) = \boldsymbol{\psi}(\boldsymbol{u}, x)$. Jak již bylo řečeno, výhodou tohoto postupu je fakt, že každý krok lze řešit naprosto odlišnou metodou, například jeden implicitní a druhý explicitní. Navíc díky tomu, že podproblémy představují jednodušší úlohy než původní problém (homogenní úloha a obyčejná diferenciální rovnice), existuje větší množství metod na jejich řešení.

Mohou zde však nastat problémy především v situacích, kdy se blížíme k ustálenému stavu, ve kterém $[\mathbf{f}(\mathbf{u})]_x = \boldsymbol{\psi}(\mathbf{u}, x)$. V takových případech je tato metoda nevhodná, neboť produkuje značné výchylky od ustáleného stavu a je tedy obtížné jej numericky určit. Dalším problémem jsou okrajové podmínky, které jsou dané pro původní úlohu se zdrojovými členy. Není vždy jednoduché na jejich základě stanovit, jaké okrajové podmínky zvolit pro obě nově sestavené úlohy. Problémy mohou být také se stabilitou celé metody štěpení, přestože metody použité na oba kroky budou stabilní, jejich kombinování může přinést nestabilitu. Podrobnosti o této problematice je možné nalézt například v [9].

Jako další příklad uveďme například advekčně difuzní rovnici

$$u_t + \left(\frac{1}{2}u^2\right)_x = u_{xx} \tag{5.3}$$

Pokud bychom rovnici (5.3) řešili nějakou explicitní metodou, dospěli bychom celkem přirozeně k jednoduchému schématu jako kombinace některé explicitní metody pro řešení advekční rovnice a metody pro řešení difuzní rovnice. Je třeba si však uvědomit, že každá část, by si kladla nároky na jinou délku časového kroku. Zatímco advekční část by vzhledem ke stabilitě omezovala časový krok řádově $\Delta t \approx \Delta x$, difuzní má vyšší nároky řádu $\Delta t \approx \Delta x^2$. Došlo by tedy k řešení advekční části na příliš mnoha časových vrstvách, přestože to není nutné vzhledem ke stabilitě. Jak už jsme zmínili v úvodu, použití plně implicitní metody na rovnici (5.3) také není příliš výhodný, zbavíme se sice závislosti na délce časového kroku, získáme však soustavu nelineárních algebraických rovnic, jejíž řešení může být pochopitelně značně složité.

Z těchto důvodů se ukazuje použití metody štěpení jako efektivní způsob řešení (5.3). Řešení probíhá na každé časové vrstvě ve dvou krocích. Advekční část je řešena explicitní metodou, neboť nemá tak omezující požadavky na délku časového kroku. Difuzní část je pak řešena nějakou vhodnou implicitní metodou, například některou uvedenou v kapitole 5.

1. $u_t = -\left(\frac{1}{2}u^2\right)_x$,
2. $u_t = u_{xx}$.

Toto štěpení (5.1) se také někdy nazývá Godunovovo štěpení a formálně je prvního řádu přesnosti v časové proměnné. Zvýšení na formální druhý řád přesnosti je možné provést tak, že jednotlivé kroky budou realizovány s polovičním časovým krokem $\Delta t/2$. Tomuto přístupu se někdy říká Strangovo štěpení. Rovnice (5.3) by tímto přístupem byla řešena ve třech krocích

1. $u_t = -\left(\frac{1}{2}u^2\right)_x$,
2. $u_t = u_{xx}$,
3. $u_t = -\left(\frac{1}{2}u^2\right)_x$,

Opět, řešení z jednotlivých kroků jsou použita, jako počáteční podmínky pro kroky následující. Krok 1 je ovšem řešen s polovičním časovým krokem $\Delta t/2$, krok 2 s časovým krokem Δt a krok 3 opět s polovičním krokem $\Delta t/2$.

5.2 Vícedimenzionální úlohy

Metody štěpení lze s výhodou využít také při řešení úloh ve více prostorových dimenzích. Budeme zde předpokládat rovnoměrnou síť s diskretizačními kroky Δx a Δy. Na dané časové vrstvě t_n tak budeme pracovat s aproximacemi neznámé funkce $\boldsymbol{U}_{jk}^n \approx \boldsymbol{u}(x_j, y_k, t_n)$, kde $x_j = j\Delta x$ a $y_k = k\Delta y$.

Nejprve si však na problémy dvojrozměrného proudění v mělkých vodách ukažme pro porovnání použití metody bez štěpení, konkrétně metodu typu central-upwind, jejíž varianta pro jednodimenzionální problém byla uvedena v [1]. Uvažujeme úlohu dvojrozměrného proudění v mělkých vodách popsaného Saint-Venantovými rovnicemi, které ve dvou prostorových dimenzích mají tvar

$$\begin{aligned} h_t + (hu)_x + (hv)_y &= 0, \\ (hu)_t + \left(hu^2 + \frac{1}{2}gh^2\right)_x + (huv)_y &= -ghB_x \\ (hv)_t + (huv)_x + \left(hv^2 + \frac{1}{2}gh^2\right)_y &= -ghB_y, \end{aligned} \tag{5.4}$$

kde $h = h(x, y, t)$ je neznámá funkce popisující hloubku vody od dna, $u = u(x, y, t)$ a $v = v(x, y, t)$ jsou neznámé funkce reprezentující kolmé složky rychlosti proudění ve směrech os x a y, $g = 9.81$ je gravitační konstanta a $B = B(x, y)$ je funkce popisující tvar dna.

Soustavu budeme stručně zapisovat

$$\mathbf{u}_t + [\mathbf{f}(\mathbf{u})]_x + [\mathbf{g}(\mathbf{u})]_y = \boldsymbol{\psi}(\mathbf{u}, x, y), \tag{5.5}$$

kde

$$\mathbf{u} = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix}, \qquad \mathbf{f}(\mathbf{u}) = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{bmatrix}, \qquad \mathbf{g}(\mathbf{u}) = \begin{bmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{bmatrix},$$

$$\boldsymbol{\psi}(\mathbf{u}, x, y) = \begin{bmatrix} 0 \\ -ghB_x \\ -ghB_y \end{bmatrix}. \tag{5.6}$$

Zde $\mathbf{u} = \mathbf{u}(x, y, t)$ je neznámá funkce, $\mathbf{f}(\mathbf{u})$ a $\mathbf{g}(\mathbf{u})$ jsou tokové funkce ve směrech osy x a y.

Metodu central-upwind lze zapsat v podobě

$$\boldsymbol{U}_{j,k}^{n+1} = \boldsymbol{U}_{j,k}^{n} - \frac{\Delta t}{\Delta x}[\mathbf{F}_{j+1/2,k} - \mathbf{F}_{j-1/2,k}] - \frac{\Delta t}{\Delta y}[\mathbf{G}_{j,k+1/2} - \mathbf{G}_{j,k-1/2}] + \boldsymbol{\Psi}_{j,k}. \tag{5.7}$$

Numerické toky $\mathbf{F}_{j+1/2,k}$ a $\mathbf{G}_{j,k+1/2}$ jsou pak dány výrazy

$$\begin{aligned} \mathbf{F}_{j+1/2,k} = & \quad \tfrac{a^+_{j+1/2,k}\mathbf{f}(\mathbf{U}^-_{j+1/2,k}) - a^-_{j+1/2,k}\mathbf{f}(\mathbf{U}^+_{j+1/2,k})}{a^+_{j+1/2,k} - a^-_{j+1/2,k}} + \\ & + \tfrac{a^+_{j+1/2,k}a^-_{j+1/2,k}}{a^+_{j+1/2,k} - a^-_{j+1/2,k}}\left[\mathbf{U}^+_{j+1/2,k} - \mathbf{U}^-_{j+1/2,k}\right], \end{aligned} \tag{5.8}$$

$$\begin{aligned} \mathbf{G}_{j,k+1/2} = & \quad \tfrac{b^+_{j,k+1/2}\mathbf{g}(\mathbf{U}^-_{j,k+1/2}) - b^-_{j,k+1/2}\mathbf{g}(\mathbf{U}^+_{j,k+1/2})}{b^+_{j,k+1/2} - b^-_{j,k+1/2}} + \\ & + \tfrac{b^+_{j,k+1/2}b^-_{j,k+1/2}}{b^+_{j,k+1/2} - b^-_{j,k+1/2}}\left[\mathbf{U}^+_{j,k+1/2} - \mathbf{U}^-_{j,k+1/2}\right], \end{aligned} \tag{5.9}$$

kde

$$\mathbf{f}(\mathbf{U}^-_{j+1/2,k}) = \begin{bmatrix} H^-_{j+1/2,k}U_{j,k} \\ H^-_{j+1/2,k}U^2_{j,k} + \frac{1}{2}g(H^-_{j+1/2,k})^2 \\ H^-_{j+1/2,k}U_{j,k}V_{j,k} \end{bmatrix} \tag{5.10}$$

$$\mathbf{f}(\mathbf{U}^+_{j+1/2,k}) = \begin{bmatrix} H^+_{j+1/2,k}U_{j+1,k} \\ H^+_{j+1/2,k}U^2_{j+1,k} + \frac{1}{2}g(H^+_{j+1/2,k})^2 \\ H^+_{j+1/2,k}U_{j+1,k}V_{j+1,k} \end{bmatrix}, \tag{5.11}$$

$$\mathbf{g}(\mathbf{U}^-_{j,k+1/2}) = \begin{bmatrix} H^-_{j,k+1/2}V_{j,k} \\ H^-_{j,k+1/2}U_{j,k}V_{j,k} \\ H^-_{j,k+1/2}V^2_{j,k} + \frac{1}{2}g(H^-_{j,k+1/2})^2 \end{bmatrix}, \tag{5.12}$$

$$\mathbf{g}(\mathbf{U}^+_{j,k+1/2}) = \begin{bmatrix} H^+_{j,k+1/2}V_{j,k+1} \\ H^+_{j,k+1/2}U_{j,k+1}V_{j,k+1} \\ H^+_{j,k+1/2}V^2_{j,k+1} + \frac{1}{2}g(H^+_{j+1/2,k})^2 \end{bmatrix} \tag{5.13}$$

a dolní a horní odhady šíření lokálních vln jsou definovány

$$\begin{aligned}
a^+_{j+1/2,k} &= \max\left\{\lambda^N\left(\mathbf{f}'(\mathbf{U}^-_{j+1/2,k})\right), \lambda^N\left(\mathbf{f}'(\mathbf{U}^+_{j+1/2,k})\right), 0\right\}, \\
a^-_{j+1/2,k} &= \min\left\{\lambda^1\left(\mathbf{f}'(\mathbf{U}^-_{j+1/2,k})\right), \lambda^1\left(\mathbf{f}'(\mathbf{U}^+_{j+1/2,k})\right), 0\right\}, \\
b^+_{j,k+1/2} &= \max\left\{\lambda^N\left(\mathbf{g}'(\mathbf{U}^-_{j,k+1/2})\right), \lambda^N\left(\mathbf{g}'(\mathbf{U}^+_{j,k+1/2})\right), 0\right\}, \\
b^-_{j,k+1/2} &= \min\left\{\lambda^1\left(\mathbf{g}'(\mathbf{U}^-_{j,k+1/2})\right), \lambda^1\left(\mathbf{g}'(\mathbf{U}^+_{j,k+1/2})\right), 0\right\}.
\end{aligned} \tag{5.14}$$

Zde $N = 3$ a předpokládáme, že máme seřazená vlastní čísla $\lambda^1 < \lambda^2 < \lambda^3$, přičemž

$$\begin{aligned}
&\lambda^1(\mathbf{f}(\mathbf{u})') = u - \sqrt{gh}, \quad \lambda^2(\mathbf{f}(\mathbf{u})') = u, \quad \lambda^3(\mathbf{f}(\mathbf{u})') = u + \sqrt{gh}, \\
&\lambda^1(\mathbf{g}(\mathbf{u})') = v - \sqrt{gh}, \quad \lambda^2(\mathbf{g}(\mathbf{u})') = v, \quad \lambda^3(\mathbf{g}(\mathbf{u})') = v + \sqrt{gh}.
\end{aligned} \tag{5.15}$$

Hodnoty neznámých veličin $H^{\pm}_{j+1/2,k}$ je opět možné volit několika způsoby, například pomocí polynomiálních rekonstrukcí, jako tomu bylo v jednodimenzionálním případě. Nevýhodou těchto rekonstrukcí je ovšem jejich časová náročnost, která se pochopitelně daleko více projevuje v případě 2D úlohy. Volíme proto jiný rychlejší postup, který navíc zaručí robustní výpočet nových hodnot neznámých veličin i v místech, kde je $h = 0$ v některé z okolních buněk (v případě proudění v říčním korytě se jedná o břehovou linii) a zachová pozitivní semidefinitnost schématu. Rekonstruované veličiny mají následující tvar

$$H^-_{j+1/2,k} = \max(0, H_{j,k}+B_{j,k}-B_{j+1/2,k}),\, H^+_{j+1/2,k} = \max(0, H_{j+1,k}+B_{j+1,k}-B_{j+1/2,k}), \tag{5.16}$$

$$H^-_{j,k+1/2} = \max(0, H_{j,k}+B_{j,k}-B_{j,k+1/2}),\, H^+_{j,k+1/2} = \max(0, H_{j,k+1}+B_{j,k+1}-B_{j,k+1/2}), \tag{5.17}$$

kde

$$B_{j+1/2,k} = \max(B_{j,k}, B_{j+1,k}), \qquad B_{j,k+1/2} = \max(B_{j,k}, B_{j,k+1}). \tag{5.18}$$

Pro zachování ustáleného stavu typu klid v jezeře, kde $u = v = 0$ a $h + B =$ konst. je možné zvolit speciální diskretizaci pravé strany (viz [7]). Za tohoto ustálené stavu totiž platí

$$\left(\frac{1}{2}gh^2\right)_x = -ghB_x. \tag{5.19}$$

$$\left(\frac{1}{2}gh^2\right)_y = -ghB_y. \tag{5.20}$$

Integrováním těchto vztahů v obou směrech přes buňku na pozici $[j, k]$ získáme aproximaci zdrojového členu

$$-\int_{x_{j-1/2}}^{x_{j+1/2}} ghB_x \mathrm{d}x \approx \frac{1}{2}g({H^-}_{j+1/2,k})^2 - \frac{1}{2}g({H^+}_{j-1/2,k})^2, \tag{5.21}$$

$$-\int\limits_{y_{k-1/2}}^{y_{k+1/2}} ghB_y \mathrm{d}x \approx \frac{1}{2}g({H^-}_{j,k+1/2})^2 - \frac{1}{2}g({H^+}_{j,k-1/2})^2. \tag{5.22}$$

Zdrojový člen má proto tvar

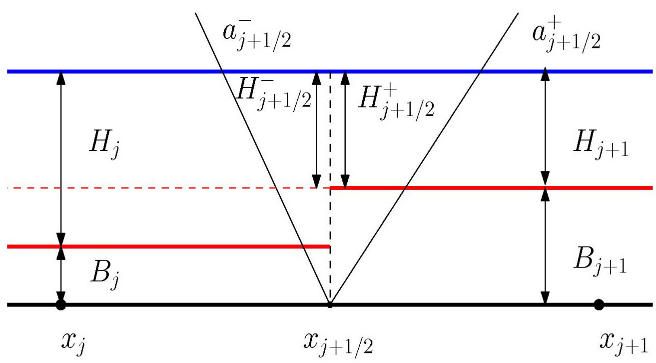
$$\mathbf{\Psi}_{j,k} = \left[\begin{array}{c} 0 \\ \frac{g}{2\Delta x}\left(({H^-}_{j+1/2,k})^2 - ({H^+}_{j-1/2,k})^2\right) \\ \frac{g}{2\Delta y}\left(({H^-}_{j,k+1/2})^2 - ({H^+}_{j,k-1/2})^2\right) \end{array} \right]. \tag{5.23}$$

Za předpokladu, že numerické toky a aproximace zdrojového členu jsou konzistentní s homogenní rovnicí, lze ukázat (viz [7]), že toto schéma zachovává ustálený stav typu klid v jezeře. To je v souladu i s tím, co jsme si uváděli o zachování ustálených stavů této metody v [1]. Požadovali jsme, aby prostorové derivace neznámých byly nulové. Toho je při použití rekonstrukce (5.16) dosaženo, neboť platí (ve směru osy x, indexy k vypouštíme)

$$\left.\begin{array}{rcl} H^-_{j+1/2} = & \max(0, H_j + B_j - \max(B_j, B_{j+1})) & = H_{j+1} \\ H^+_{j+1/2} = & \max(0, H_{j+1} + B_{j+1} - \max(B_j, B_{j+1})) & = H_{j+1} \end{array}\right\} \text{ pokud } B_j < B_{j+1}, \tag{5.24}$$

$$\left.\begin{array}{rcl} H^-_{j+1/2} = & \max(0, H_j + B_j - \max(B_j, B_{j+1})) & = H_j \\ H^+_{j+1/2} = & \max(0, H_{j+1} + B_{j+1} - \max(B_j, B_{j+1})) & = H_j \end{array}\right\} \text{ pokud } B_j > B_{j+1}. \tag{5.25}$$

Předpokládáme, že $h > 0$. Obdobně lze postupovat i ve směru osy y. Situace pro stav $B_j < B_{j+1}$ je ilustrována na obrázku 5.1. Konzistenci aproximace zdrojového



Obr. 5.1Rekonstrukce hodnot $H^{\pm}_{j+1/2}$

členu požadujeme ve smyslu splnění následující rovnosti (podrobnosti v [2])

$$\Delta x \mathbf{\Psi}_j = -gh_j \Delta B_j + O(\Delta B_j), \tag{5.26}$$

když $B_j \to B_{j+1}$. V našem případě

$$\begin{aligned}
&\Delta x \boldsymbol{\Psi}_j = \tfrac{g}{2}(H^-_{j+1/2})^2 - \tfrac{g}{2}(H^+_{j-1/2})^2 = \\
&= \tfrac{g}{2}\left[\left(H_j + (B_j - B_{j+1/2})\right)^2 - \left(H_j + (B_j - B_{j-1/2})\right)^2\right] = \\
&= \tfrac{g}{2}\left[2H_j(-B_{j+1/2} + B_{j-1/2}) + (B_{j+1/2} + B_{j-1/2})(B_j - B_{j+1/2} + B_j - B_{j-1/2})\right] = \\
&= -gH_j(B_{j+1/2} - B_{j-1/2}) - \tfrac{g}{2}\left(B_{j+1/2} - B_{j-1/2}\right)\left((B_j - B_{j+1/2}) + (B_j - B_{j-1/2})\right).
\end{aligned} \tag{5.27}$$

Z výrazu (5.27) je patrné, že naše aproximace druhé složky zdrojového členu (5.23) vztah (5.26) splňuje. Obdobně lze konzistenci ukázat i pro třetí složku zdrojového členu (5.23).

Pokud bychom na stejnou úlohu použili metody štěpení, vypadala by například takto:

1. $\mathbf{u}^x_t + [\mathbf{f}(\mathbf{u}^x)]_x = \mathbf{0}, \qquad \mathbf{u}^x(0) = \mathbf{u}_0,$
2. $\mathbf{u}^y_t + [\mathbf{g}(\mathbf{u}^y)]_y = \mathbf{0}, \qquad \mathbf{u}^y(0) = \mathbf{u}^x,$
3. $\mathbf{u}_t = \boldsymbol{\psi}(\mathbf{u}, x, y), \qquad \mathbf{u}(0) = \mathbf{u}^y.$

První a druhý krok bychom řešili jednodimenzionální metodou typu central-upwind, třetí krok nějakou vhodnou metodou pro řešení obyčejných diferenciálních rovnic.

V dalších částech si uvedeme dva typy metod štěpení - lokálně jednodimenzionální metody (LOD) a metody střídavých směrů (ADI). Jedním z rozdílů v přístupu těchto metod je to, že zatímco ADI metody jsou konzistentní s aproximovanou diferenciální rovnicí ve všech krocích na dané časové vrstvě, LOD metody zachovávají konzistenci až po uskutečnění posledního kroku.

5.3 LOD - Lokálně jednodimenzionální metody

Tyto metody byly vyvinuty především pro řešení vícedimenzionálních úloh tak, aby v každém kroku řešili jednodimenzionální problém ve směru dané souřadné osy.

5.3.1 LOD-implicitní Eulerova metoda

Za předpokladu štěpení diferenciálního operátoru (5.1) a využitím implicitní Eulerovy metody, získáme schéma ve tvaru

$$\boldsymbol{u}_0 = \boldsymbol{u}^n, \tag{5.28}$$

$$\begin{aligned}
\boldsymbol{u}_i &= \boldsymbol{u}_{i-1} + \Delta t \boldsymbol{h}_i(\boldsymbol{u}_i, t_{n+1}), \qquad i = 1, \dots s \\
\boldsymbol{u}^{n+1} &= \boldsymbol{u}_s.
\end{aligned} \tag{5.29}$$

Jako příklad si vezměme skalární Laplaceovu rovnici

$$u_t = u_{xx} + u_{yy} \tag{5.30}$$

Metoda se pak skládá z řešení dvou implicitních Eulerových metod

1. $U_{i,j}^{n+1,*} = U_{i,j}^{n} + \frac{\Delta t}{(\Delta x)^2}\left(U_{i-1,j}^{n+1,*} - 2U_{i,j}^{n+1,*} + U_{i+1,j}^{n+1,*}\right)$,

2. $U_{i,j}^{n+1} = U_{i,j}^{n+1,*} + \frac{\Delta t}{(\Delta y)^2}\left(U_{i,j-1}^{n+1} - 2U_{i,j}^{n+1} + U_{i,j+1}^{n+1}\right)$

Samozřejmě lze zkonstruovat i LOD-explicitní Eulerovu metodu, tu lze však jednoduchým dosazením zapsat ve tvaru bez štěpení

$$U_{i,j}^{n+1} = U_{i,j}^{n} + \frac{\Delta t}{(\Delta x)^2}\left(U_{i-1,j}^{n} - 2U_{i,j}^{n} + U_{i+1,j}^{n}\right) + \frac{\Delta t}{(\Delta y)^2}\left(U_{i,j-1}^{n} - 2U_{i,j}^{n} + U_{i,j+1}^{n}\right). \tag{5.31}$$

Diskretizační chyba metody (5.31) je $O(\Delta t + \Delta x^2 + \Delta y^2)$ a podmínka stability má tvar

$$\frac{\Delta t}{\Delta x^2} + \frac{\Delta t}{\Delta y^2} \leqq \frac{1}{2}. \tag{5.32}$$

5.3.2 LOD Crankova-Nicolsonova metoda

Jako další příklad lokálně jednodimenzionální metody si uvedeme LOD Crankovu--Nicolsonovu metodu. Opět využijeme štěpení diferenciálního operátoru (5.1) a tvaru samotné Crankovy-Nicolsonovy metody pro jednodimenzionální problémy. Dané schéma pak lze zapsat ve tvaru

$$\begin{aligned}
\boldsymbol{u}_0 &= \boldsymbol{u}_n, && (5.33)\\
\boldsymbol{u}_i &= \boldsymbol{u}_{i-1} + \frac{1}{2}\Delta t\boldsymbol{h}_i(\boldsymbol{u}_{i-1}, t_n + c_{i-1}\Delta t) + \frac{1}{2}\Delta t\boldsymbol{h}_i(u_i, t_n + c_i\Delta t), \qquad i = 1, \dots s\\
\boldsymbol{u}_{n+1} &= \boldsymbol{u}_s, && (5.34)
\end{aligned}$$

kde $c_0 = 0$ a $c_s = 1$, zatímco ostatní $c_i = \frac{1}{2}$. Použití této metody na řešení Laplaceovy rovnice (5.1) vede opět na konstrukci dvou kroků

1. $U_{i,j}^{n+1,*} = U_{i,j}^{n} + \frac{1}{2}\frac{\Delta t}{(\Delta x)^2}\left(U_{i-1,j}^{n} - 2U_{i,j}^{n} + U_{i+1,j}^{n}\right) + \frac{1}{2}\frac{\Delta t}{(\Delta x)^2}\left(U_{i-1,j}^{n+1,*} - 2U_{i,j}^{n+1,*} + U_{i+1,j}^{n+1,*}\right)$,

2. $U_{i,j}^{n+1} = U_{i,j}^{n+1,*} + \frac{1}{2}\frac{\Delta t}{(\Delta y)^2}\left(U_{i,j-1}^{n+1} - 2U_{i,j}^{n+1} + U_{i,j+1}^{n+1}\right) + \frac{1}{2}\frac{\Delta t}{(\Delta y)^2}\left(U_{i-1,j}^{n+1,*} - 2U_{i,j}^{n+1,*} + U_{i+1,j}^{n+1,*}\right)$,

5.4 ADI - Metody střídavých směrů

Stejně jako LOD metody uvedené v předchozí části, lze využít i metody střídavých směrů pro řešení vícedimenzionálních úloh. My si zde však ukážeme jejich použití při štěpení advekčně-difuzních problémů, při kterých dochází na štěpení na advekční a difuzní část. Často se zde využívá kombinace explicitních a implicitních metod, kdy každá část diferenciální rovnice klade jiné nároky na délku časového kroku.

5.4.1 Douglasova metoda

Za předpokladu štěpení diferenciálního operátoru (5.1) má Douglasova metoda následující tvar

$$\begin{aligned} \boldsymbol{u}_0 &= \boldsymbol{u}_n + \Delta t \boldsymbol{h}(\boldsymbol{u}_n, t_n), \\ \boldsymbol{u}_i &= \boldsymbol{u}_{i-1} + \theta \Delta t (\boldsymbol{h}_i(\boldsymbol{u}_i, t_{n+1}) - \boldsymbol{h}_i(\boldsymbol{u}_n, t_n)), \qquad i = 2, \dots s \\ \boldsymbol{u}_{n+1} &= \boldsymbol{u}_s. \end{aligned} \tag{5.35}$$

Tato metoda je prvního řádu přesnosti, když $\theta = 1$ a druhého řádu pokud $\theta = \frac{1}{2}$ a zároveň $\boldsymbol{h}_1 = \boldsymbol{0}$. Všimněme si, že tato metoda je skutečně konzistentní s diferenciální rovnicí v každém kroku.

Speciálním případem Douglasovy metody je Peacemanova-Rachfordova metoda, která předpokládá rozštěpení diferenciálního operátoru pouze na dvě části, tedy $s = 2$ a volbu $\theta = \frac{1}{2}$. Lze ji tedy použít například na řešení Laplaceovy rovnice (5.30).

$$U_{i,j}^{n+1/2} = U_{i,j}^n + \frac{\Delta t}{2} \left[\frac{U_{i-1,j}^{n+1/2} - 2U_{i,j}^{n+1/2} + U_{i+1,j}^{n+1/2}}{(\Delta x)^2} + \frac{U_{i,j-1}^n - 2U_{i,j}^n + U_{i,j+1}^n}{(\Delta y)^2} \right] \tag{5.36}$$

$$U_{i,j}^{n+1} = U_{i,j}^{n+1/2} + \frac{\Delta t}{2} \left[\frac{U_{i-1,j}^{n+1/2} - 2U_{i,j}^{n+1/2} + U_{i+1,j}^{n+1/2}}{(\Delta x)^2} + \frac{U_{i,j-1}^{n+1} - 2U_{i,j}^{n+1} + U_{i,j+1}^{n+1}}{(\Delta y)^2} \right] \tag{5.37}$$

Nyní si ukažme použití Douglasovy metody na řešení advekčně-difuzní rovnice ve třech prostorových dimenzích. Jako testovací úloha nám poslouží řešení šíření koncentrace nečistot v ovzduší. Tato úloha může být popsána následující rovnicí

$$q_t = -\nabla(\boldsymbol{v}q) + D\triangle q, \tag{5.38}$$

kde $q = q(x, y, z)$ představuje koncentraci nečistot v ovzduší, $\boldsymbol{v} = \boldsymbol{v}(x, y, z) = (V_1(x, y, z), V_2(x, y, z), V_3(x, y, z))$ je funkce popisující proudění vzduchu ve směrech souřadných os x, y a z a $D = D(x, y, z)$ je difúzní koeficient. Advekční část je možno diskretizovat některou vhodnou metodou popsané v kapitole 3, difúzní část pak zase některou metodou v kapitoly 5.

Řešení Douglasovou metodou při volbě $\theta = 1$ probíhá ve čtyřech krocích. Pro advekční část uvažuje metodu v konzervativním tvaru, pro advekční volíme diskretizaci jako v Eulerově metodě.

1.

$$\begin{aligned} Q_{ijk}^{n+1,0} = Q_{ijk}^n &- \frac{\Delta t}{\Delta x} \left(F_{i+1/2,jk}^n - F_{i-1/2,jk}^n \right) - \frac{\Delta t}{\Delta y} \left(F_{i,j+1/2,k}^n - F_{i,j-1/2,k}^n \right) - \\ &- \frac{\Delta t}{\Delta z} \left(F_{ij,k+1/2}^n - F_{ij,k-1/2}^n \right) + \frac{D\Delta t}{\Delta x^2} (Q_{i-1,jk}^n - 2Q_{ijk}^n + Q_{i+1,jk}^n) + \\ &+ \frac{D\Delta t}{\Delta y^2} (Q_{i,j-1,k}^n - 2Q_{ijk}^n + Q_{i,j+1,k}^n) + \frac{D\Delta t}{\Delta z^2} (Q_{ij,k-1}^n - 2Q_{ijk}^n + Q_{ij,k+1}^n) \end{aligned}$$

2.

$$Q_{ijk}^{n+1,1} = Q_{ijk}^{n+1,0} - \frac{D\Delta t}{\Delta x^2}(Q_{i-1,jk}^n - 2Q_{ijk}^n + Q_{i+1,jk}^n) + \frac{D\Delta t}{\Delta x^2}(Q_{i-1,jk}^{n+1,1} - 2Q_{ijk}^{n+1,1} + Q_{i+1,jk}^{n+1,1})$$

3.

$$Q_{ijk}^{n+1,2} = Q_{ijk}^{n+1,1} - \frac{D\Delta t}{\Delta y^2}(Q_{i,j-1,k}^n - 2Q_{ijk}^n + Q_{i,j+1,k}^n) + \frac{D\Delta t}{\Delta y^2}(Q_{i,j-1,k}^{n+1,1} - 2Q_{ijk}^{n+1,1} + Q_{i,j+1,k}^{n+1,1})$$

4.

$$Q_{ijk}^{n+1} = Q_{ijk}^{n+1,2} - \frac{D\Delta t}{\Delta z^2}(Q_{ij,k-1}^n - 2Q_{ijk}^n + Q_{ij,k+1}^n) + \frac{D\Delta t}{\Delta z^2}(Q_{ij,k-1}^{n+1} - 2Q_{ijk}^{n+1} + Q_{ij,k+1}^{n+1})$$

Velikost časového kroku pro zachování stability musí splňovat omezení pouze s ohledem na advekční část, tedy obdobu CFL podmínky v jednodimenzionálním případě s ohledem na konkrétní metodu.

5.5 Závěrečné poznámky

V této kapitole jsme ukázali možnosti řešení úloh ve více prostorových dimenzích, či úloh, obsahujících zároveň advekční i difúzní členy. Pro tyto úlohy se jako velmi efektivní ukazují metody štěpení, které umožňují převést daný problém na více jednodušších úloh, které lze řešit metodami uvedenými v předchozích částech skript. štěpení však zároveň přináší i některé nevýhody. Jednou z nich je problematické udržení ustálených stavů, které ze své podstaty vyžaduje rovnost členů, jež se díky štěpení neobjevují ve stejné rovnici.

Dalším problémem, který nastává u vícedimenzionálních úloh je požadavek na TVD vlastnost. U jednodimenzionálních úloh tato vlastnost zaručovala, že se v místech nespojitého řešení netvořili nežádoucí oscilace a navíc v místech hladkého řešení bylo možné dosáhnout vyššího řádu přesnosti. Vezměme si však úlohu ve dvou prostorových dimenzích, kde je možné totální variaci definovat následujícím způsobem

$$TV(U) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \left(\Delta x|U_{i+1,j} - U_{ij}| + \Delta y|U_{i,j+1} - U_{ij}|\right). \tag{5.39}$$

Lze ukázat, že až na některé triviální případy, každá metoda, která má vlastnost TVD ve dvou prostorových dimenzích je nejvýše prvního řádu [9]. To však neznamená, že není možné konstruovat metody vyššího řádu, které dávají kvalitativně dobrá řešení, jako u jednodimenzionálních úloh. Například metody štěpení dávají z tohoto pohledu dobré výsledky, pokud aplikujeme metody typu high-resolution na každý krok štěpení. Navíc, pokud použijeme například metody Strangova štěpení a vhodné limitery, můžeme konstruovat metody druhého řádu, které neprodukují nežádoucí oscilace. Problém je ovšem v tom, že ani s těmito prostředky nejsme schopni dokázat, že totální variace (5.39) neroste.

Literatura

[1] Brandner, M., Egermaier, J., Kopincová, H.: *Numerické modelování v hydrologii*

[2] Bouchut, F.: *Nonlinear Stability of Finite Volume Methods for Hyperbolic Conservation Laws*, Birkhäuser Verlag, 2004

[3] Crossley, A., J.: *Accurate and Efficient Numerical Solutions for the Saint Venant Equations of Open Channel Flow*, University of Nottingham, Ph.D. Thesis, 1999.

[4] Serre, D.: *Systems of Conservation laws I: Hyperbolicity, Entropies, Shock Waves*, Cambridge University Press, 1999, Cambridge, ISBN: 0-521-58233-4.

[5] Toro, E., F.: *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer-Verlag, 2009, Berlin, ISBN: 978-3-540-25202-3.

[6] Toro, E., F.: *Shock-Capturing Methods for Free-Surface Shallow Flows*, John Wiley & sons, Ltd, 2001, West Sussex, ISBN: 0-471-98766-2.

[7] Audusse, E., Bouchut, F., Bristeau, M.-O., Klein, R., Perthame, B.: *A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow eater flows*, SIAM Journal on Scientific Computing, Volume 25, Issue 6, pages: 2050 - 2065, 2004

[8] Kufner, A., John, O., Fučík, S.: *Function Spaces*, Academia, Praha and Noordhoff, Leyden, 1977, ISBN: 90-286-0015-9.

[9] LeVeque, R., J.: *Finite Volume Methods for Hyperbolic Problems*, Cambridge University Press, 2002, Cambridge, ISBN: 0-521-81087-6.

[10] Crossley, A. J. *Accurate and efficient numerical solutions for Saint Venant equations of open channel flow.*

[11] Straškraba, I. *Theory and Praxis of Mathematical Modelling of Hydrodynamical Processes.*

[12] Kurganov, A. and Levy, D.: *Central-Upwind Schemes for the Saint-Venant System.* Mathematical Modeling and Numerical **36** (2002), 397–425

[13] Kurganov, A. and Tadmor, E.: *New High-Resolution Central Schemes for Nonlinear Conservation Laws and Convection-Diffusion Equations.* Journal of Computational Physics **160** Issue 1 (2000), 241–282

[14] Črnjarič-Zič, N., Vukovič, S. and Sopta, L.: *Balanced finite volume WENO and central WENO schemes for the shallow water and the open-channel flow equations.* Journal of Computational Physics **200** Issue 2 (2004), 512–548

[15] George, D.L.: *Finite Volume Methods and Adaptive Refinement for Tsunami Propagation and Innundation.* University of Washington, Ph.D. Thesis, (2006)

[16] Wolf-Gladrow, D.A.: *Lattice gas celluar automata and lattice Boltzmann models: an introduction*, Lecture notes in mathematics **1725** (2000)

[17] Gueyffier, D., Li., J., Nadin, A, scardovelli, R. and zaleski, S.: *Volume-of-Fluid Interface Tracking with Smoothed Surface Stress Methods for Three-Dimensional Flows.* Journal of Computational Physics **152** (1999), 423–456

[18] Li, T.-T., Yu, W.: *Boundary Value Problem for Quasilinear Hyperbolic Systems*, Duke University Matematics Series V., 1985.

[19] Le Floch, P., G., Raviart, P. A.: *An Asymtotic Expansion of the Solution of the Generalized Riemann Problem. Part I: General Theory*, Annales de Inst. Henri Poincare, Analyse non Lineáre, 1988, vol. 5, no. 2, pp. 179 – 209.

[20] Li, T.-T., Yu, W.: *Boundary Value Problem for the First-Order Quasi-Linear Hyperbolic Systems and Their Applications*, Journal of Differential Equations, 1981, vol. 41, no. 1, pp. 1 – 26.

[21] Lax, P., D.: *Hyperbolic Systems of Conservation Laws II*, Communications on Pure and Applied Mathematics, 1957, vol. 10, no. 4, pp. 537 – 566.

[22] LeVeque, R., J.: *Numerical Methods for Conservation Laws*, Birkhauser Verlag, 1990, Basel, ISBN: 3-7643-2464-3.

[23] Jameson, A.: *Positive Schemes and Shock Modelling for Compressible Flows*, International Journal for Numerical Methods in Fluids, 1995, vol. 20, pp. 743 – 776 (1995)

[24] Shu, Ch.-W.: *Numerical Methods for Hyperbolic Conservation Laws*, Brown, 2006

[25] Wu, A., Yin, J., Wang, C.: *Elliptic and parabolic equations*, Word Scientific Publishing, 2006, ISBN: 981-270-026-9.

Rejstřík