KVATERNION 1–2 (2024), 3–18

RÔZNE POHĽADY NA ROVNICU VEDENIA TEPLA

MATEJ BENKO

ABSTRAKT. V tomto článku diskutujeme tri rôzne interpretácie rovnice vedenia tepla (parabolická parciálna diferenciálna rovnica). Najskôr ukážeme, že pozdĺž každého jej slabého riešenia klesá funkcionál, ktorý reprezentuje energiu systému. Pokračujeme úvahou, že pokiaľ jej riešenie reprezentujeme ako hustotu pravdepodobnosti a uvažujeme priestor pravdepodobnostných mier, riešenie môžeme interpretovať ako krivky, ktoré generujú gradientný tok po funkcionáli, ktorý reprezentuje zápornú Boltzmannovu entropiu. Toto zodpovedá fyzikálnym zákonom. Na záver si ukážeme, že riešenie tejto rovnice vieme tiež interpretovať ako hustotou pravdepodobnosti riešenia stochastického procesu (konkrétne Brownovho pohybu).

1. ÚVOD

Budeme študovať rovnicu vedenia tepla s počiatočnou podmienkou, uvažujeme teda úlohu

$$\begin{cases} \dfrac{\partial \varrho}{\partial t} = \Delta \varrho, \\ \varrho(\cdot, 0) = \varrho_0. \end{cases} \tag{1.1}$$

Funkcia ϱ predstavuje teplotu v bode x a v čase t. Riešenie úlohy (1.1) hľadáme na množine $\mathbb{R}^d \times [0, T)$, kde $T < +\infty$, a uvažujeme ho v *slabom zmysle.*

Rovnicu z každej strany vynásobíme testovacou funkciou φ z priestoru $C_c^\infty(\mathbb{R}^d \times [0, T))$, ktorý obsahuje funkcie s kompaktným nosičom majúce parciálne derivácie všetkých rádov. Túto funkciu spojito dodefinujeme tak, že $\varphi(\cdot, T) = 0$. Obe strany zintegrujeme a dostaneme

$$\int_0^T \int_{\mathbb{R}^d} \varphi \frac{\partial \varrho}{\partial t} \,\mathrm{d}x \,\mathrm{d}t = \int_0^T \int_{\mathbb{R}^d} \varphi \,\Delta \varrho \,\mathrm{d}x \,\mathrm{d}t \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^d \times [0, T)).$$

Na ľavej strane zameníme poradie integrácie a integrujeme per partes:

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} \varphi \frac{\partial \varrho}{\partial t} \,\mathrm{d}x \,\mathrm{d}t &= \int_{\mathbb{R}^d} \int_0^T \varphi \frac{\partial \varrho}{\partial t} \,\mathrm{d}t \,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} \left(\varrho(\cdot, T)\varphi(\cdot, T) - \varrho(\cdot, 0)\varphi(\cdot, 0) - \int_0^T \varrho \frac{\partial \varphi}{\partial t} \,\mathrm{d}t \right) \mathrm{d}x \end{aligned}$$

2010 *MSC.* Primární 60H30, 60J65, 35K05; Sekundární 49J50, 60H05.

Klíčová slova. Gradientný tok, parabolická parciálna diferenciálna rovnica, stochastická diferenciálna rovnica.

$$= -\int_{\mathbb{R}^d} \varphi(\cdot,0)\varrho_0 \,\mathrm{d}x - \int_0^T \int_{\mathbb{R}^d} \varrho \frac{\partial \varphi}{\partial t} \,\mathrm{d}x \,\mathrm{d}t.$$

Podobne na pravej strane použijeme Greenovu vetu a získame

$$\int_0^T \int_{\mathbb{R}^d} \varphi \,\Delta \varrho \,\mathrm{d}x \,\mathrm{d}t = -\int_0^T \int_{\mathbb{R}^d} \nabla\varphi \cdot \nabla\varrho \,\mathrm{d}x \,\mathrm{d}t.$$

Porovnaním upravenej ľavej a pravej strany získame slabú formuláciu problému. Riešením úlohy (1.1) budeme chápať „rozumnú“ funkciu ϱ, ktorá spĺňa

$$-\int_{\mathbb{R}^d} \varphi(\cdot,0)\,\varrho_0 \,\mathrm{d}x = \int_0^T \int_{\mathbb{R}^d} \varrho \frac{\partial \varphi}{\partial t} \,\mathrm{d}x \,\mathrm{d}t - \int_0^T \int_{\mathbb{R}^d} \nabla\varphi \cdot \nabla\varrho \,\mathrm{d}x \,\mathrm{d}t \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^d \times [0,T)). \tag{1.2}$$

V prvej z našich úvah slovami „rozumná funkcia“ chápeme funkciu ϱ takú, že pre skoro každé t je funkcia $\varrho(\cdot,t)$ z priestoru $W_0^{1,2}(\mathbb{R}^d)$. Je to *Sobolevov priestor* funkcií na $\mathbb{R}^d$, ktoré v „nekonečne“ idú k nule a ich slabé derivácie prvého rádu sú Lebesgueovsky integrovateľné v druhej mocnine. Je známe, že také riešenie úlohy (1.1) existuje, ak $\varrho_0 \in W_0^{1,2}(\mathbb{R}^d)$.

V ďalších dvoch úvahách budeme hľadať riešenie ϱ tak, že pre skoro každé t bude $\varrho(\cdot,t)$ hustotou pravdepodobnosti. Hustoty musia dostatočne rýchlo klesať v „nekonečne“ k nule, aby sa zabezpečil jednotkový integrál hustoty cez celý priestor.

V našej prvej úvahe chceme rozumieť funkcii teploty ϱ tak, že príroda sa snaží minimalizovať energiu uvažovaného systému. To znamená, že riešenie ϱ budeme interpretovať ako krivku v abstraktnom priestore, ktorá predstavuje najrýchlejšie klesanie energie telesa. Takúto krivku budeme nazývať *gradientný tok*.

Zavedieme pojem gradientného toku na Hilbertovom priestore. Dostaneme sa k prvému výsledku. Ten hovorí, že na Hilbertovom priestore rovnica vedenia tepla generuje gradientný tok po Dirichletovej energii (ktorú reprezentujeme ako funkcionál na Hilbertovom priestore).

Následne definujeme gradientný tok na metrickom priestore náhodných veličín s tzv. Wassersteinovou metrikou (tento priestor nie je lineárny) a ukážeme, že riešenie rovnice vedenia tepla generuje gradientný tok po záporne vzatej Boltzmannovej entropii. V tomto prípade môžeme interpretovať riešenie ϱ v rôznych časoch ako hustoty pravdepodobnosti.

Na záver nadviažeme na myšlienku reprezentácie riešenia ako pravdepodobnostnej miery, resp. náhodnej veličiny. Ukážeme si, že riešenie ϱ v slabej formulácii je ekvivalentné riešeniu stochastickej diferenciálnej rovnice. V tomto prípade Brownovmu pohybu, ktorý sa v teórii stochastických procesov tiež nazýva Wienerov proces.

V nasledujúcej kapitole diskutujeme gradientný tok na Hilbertovom priestore $L^2(\mathbb{R}^d)$. Pre názornosť si teraz uvedieme gradientný tok na Euklidovom priestore. Nech má funkcia $f\colon \mathbb{R}^d \to \mathbb{R}$ spojité parciálne derivácie prvého rádu (t.j. $f \in C^1(\mathbb{R}^d)$). Potom pre každé $x_0 \in \mathbb{R}^d$ je fázová trajektória riešenia $x\colon [0,+\infty) \to \mathbb{R}^d$

počiatočnej úlohy

$$\begin{cases} \dot{x} = -\nabla f(x), \\ x(0) = x_0 \end{cases} \tag{1.3}$$

krivkou v priestore $\mathbb{R}^d$, ktorú nazývame *gradientný tok* na $\mathbb{R}^d$ po funkcii f.

Všimnime si, že dotykový vektor $\dot{x}(t)$ gradientného toku v bode $x(t)$ je rovný záporne vzatému gradientu $-\nabla f(x(t))$ a preto určuje smer najstrmšieho klesania funkcie f.

2. Gradientný tok na $L^2(\mathbb{R}^d)$

Pojem gradietný tok na $\mathbb{R}^d$ zovšeobecníme, zavedieme pojem gradietný tok na priestore $L^2(\mathbb{R}^d)$. Pre súlad s literatúrou (a tiež kvôli tomu, že smerujeme k popisu gradientného toku na abstraktnejších priestoroch), budeme namiesto s pojmom gradient funkcie pracovať s pojmom subdiferenciál funkcionálu.

Definujme teda najskôr subdiferenciál na všeobecnom Hilbertovom priestore H.

Definícia 2.1 (Fréchetov subdiferenciál na H)**.** Nech je H Hilbertov priestor a $F\colon H \to \mathbb{R} \cup \{+\infty\}$ je funkcionál definovaný na celom H. *Fréchetovym subdiferenciálom* nazývame viachodnotový operátor $\partial F\colon H \to 2^H$ definovaný vzťahom

$$\partial F[u] = \Big\{ \xi \in H;\ F[v] \geq F[u] + \langle \xi, v-u \rangle\ \forall v \in H \Big\} \quad \text{pre } u \in H.$$

Teraz, definujeme gradientný tok na $L^2(\mathbb{R}^d)$.

Definícia 2.2 (Gradientný tok na $L^2(\mathbb{R}^d)$)**.** Abstraktnú funkciu $u\colon [0,T) \to L^2(\mathbb{R}^d)$ nazveme *gradientným tokom* na $L^2(\mathbb{R}^d)$ po funkcionáli $F\colon L^2(\mathbb{R}^d) \to \mathbb{R}$ s počiatočným bodom $u_0 \in L^2(\mathbb{R}^d)$, ak pre skoro všetky $t \in [0,T)$ existuje $\dot{u}(t)$, $\partial F[u(t)] \neq \emptyset$ a platí

$$-\dot{u}(t) \in \partial F[u(t)] \quad \text{pre s. v. } t \in [0,T), \quad u(0) = u_0.$$

Pristúpime k výsledku našej prvej úvahy a to, že každé riešenie rovnice vedenia tepla generuje gradientný tok po tzv. Dirichletovej energii.

Veta 2.3. *Na $L^2(\mathbb{R}^d)$ uvažujme funkcionál*

$$F[u] := \begin{cases} \frac{1}{2} \int_{\mathbb{R}^d} \|\nabla u\|^2 \,\mathrm{d}x & \text{pre } u \in W_0^{1,2}(\mathbb{R}^d), \\ +\infty & \text{inak,} \end{cases} \tag{2.1}$$

ktorý popisuje tzv. Dirichletovu energiu. Potom pre každé $u \in W_0^{1,2}(\mathbb{R}^d)$ platí

$$\partial F[u] \neq \emptyset \quad \Leftrightarrow \quad \Delta u \in L^2(\mathbb{R}^d).$$

Ak $u \in W_0^{1,2}(\mathbb{R}^d)$ a $\Delta u \in L^2(\mathbb{R}^d)$, potom

$$\partial F[u] = \{-\Delta u\}.$$

Dôsledok 2.4. *Každý gradientný tok po funkcionáli F definovaným vzťahom (2.1) na priestore $L^2(\mathbb{R}^d)$ generuje nejaké riešenie rovnice vedenia tepla v slabom zmysle a naopak.*

Pre ukážku aparátu používaného v teórii gradientných tokov na priestore $L^2(\mathbb{R}^d)$ uvedieme dôkaz vety 2.3.

Dôkaz vety 2.3. „$\Rightarrow$“: Nech $u \in W_0^{1,2}(\mathbb{R}^d)$ a funkcia $\xi \in L^2(\mathbb{R}^d)$ je prvkom subdiferenciálu $\partial F[u]$. Potom podľa definície 2.1 platí

$$F[v] \geq F[u] + \langle \xi, v - u \rangle \quad \forall v \in L^2(\mathbb{R}^d). \tag{2.2}$$

Zoberieme funkciu $v = u + \varepsilon w$, kde $w \in W_0^{1,2}(\mathbb{R}^d)$ a $\varepsilon > 0$. Dosadíme do nerovnosti v (2.2), všimneme si, že $v - u = \varepsilon w$ a získame vzťah

$$\frac{1}{2}\int_{\mathbb{R}^d} \|\nabla(u + \varepsilon w)\|^2 \,\mathrm{d}x - \frac{1}{2}\int_{\mathbb{R}^d} \|\nabla u\|^2 \,\mathrm{d}x \geq \varepsilon \int_{\mathbb{R}^d} \xi w \,\mathrm{d}x. \tag{2.3}$$

Úpravou

$$\|\nabla(u + \varepsilon w)\|^2 = \|\nabla u + \varepsilon \nabla w\|^2 = \|\nabla u\|^2 + 2\varepsilon \nabla u \cdot \nabla w + \varepsilon^2 \|\nabla w\|^2$$

a dosadením do (2.3) získame

$$\varepsilon \int_{\mathbb{R}^d} \nabla u \cdot \nabla w \,\mathrm{d}x + \frac{\varepsilon^2}{2}\int_{\mathbb{R}^d} \|\nabla w\|^2 \,\mathrm{d}x \geq \varepsilon \int_{\mathbb{R}^d} \xi w \,\mathrm{d}x.$$

Vydelíme ε, prejdeme k limite pre $\varepsilon \to 0$ a získame tak

$$\int_{\mathbb{R}^d} \nabla u \cdot \nabla w \,\mathrm{d}x \geq \int_{\mathbb{R}^d} \xi w \,\mathrm{d}x \quad \forall w \in W_0^{1,2}(\mathbb{R}^d),$$

čo je možné s využitím Greenovej vety na ľavej strane nerovnosti prepísať do tvaru

$$-\int_{\mathbb{R}^d} w \Delta u \,\mathrm{d}x \geq \int_{\mathbb{R}^d} \xi w \,\mathrm{d}x \quad \forall w \in W_0^{1,2}(\mathbb{R}^d). \tag{2.4}$$

Keďže nerovnosť v (2.4) platí pre ľubovoľnú funkciu z priestoru $W_0^{1,2}(\mathbb{R}^d)$, môžeme uvažovať tiež funkciu $-w \in W_0^{1,2}(\mathbb{R}^d)$ (pretože je lineárnym priestorom) a získame

$$\int_{\mathbb{R}^d} w \Delta u \,\mathrm{d}x \geq -\int_{\mathbb{R}^d} \xi w \,\mathrm{d}x \quad \forall w \in W_0^{1,2}(\mathbb{R}^d). \tag{2.5}$$

Porovnaním nerovností (2.4) a (2.5) získame

$$\int_{\mathbb{R}^d} w \Delta u \,\mathrm{d}x = -\int_{\mathbb{R}^d} \xi w \,\mathrm{d}x \quad \forall w \in W_0^{1,2}(\mathbb{R}^d).$$

Odtiaľ plynie $\xi = -\Delta u$, a preto $-\Delta u \in \partial \mathcal{F}[u]$ a $\Delta u \in L^2(\mathbb{R}^d)$.

„$\Leftarrow$“: Predpokladajme, že $u \in W_0^{1,2}(\mathbb{R}^d)$ je také, že $\Delta u \in L^2(\mathbb{R}^d)$. Potom podľa (2.1) pre každé $w \in W_0^{1,2}(\mathbb{R}^d)$ platí

$$\begin{aligned} F[u + w] - F[u] &= \frac{1}{2}\int_{\mathbb{R}^d} \|\nabla(u + w)\|^2 \,\mathrm{d}x - \frac{1}{2}\int_{\mathbb{R}^d} \|\nabla u\|^2 \,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} \nabla u \cdot \nabla w \,\mathrm{d}x + \frac{1}{2}\int_{\mathbb{R}^d} \|\nabla w\|^2 \,\mathrm{d}x \\ &\geq \int_{\mathbb{R}^d} \nabla u \cdot \nabla w \,\mathrm{d}x = -\int_{\mathbb{R}^d} w \Delta u \,\mathrm{d}x. \end{aligned}$$

V prípade, že $w \in L^2(\mathbb{R}^d)$, avšak $w \notin W_0^{1,2}(\mathbb{R}^d)$, z (2.1) je zrejmé, že

$$F[u+w] = +\infty \geq F[u] - \int_{\mathbb{R}^d} w\Delta u \,\mathrm{d}x.$$

Dokázali sme, že

$$F[u+w] - F[u] \geq -\int_{\mathbb{R}^d} w\Delta u \,\mathrm{d}x = \langle -\Delta u, w\rangle \quad \forall w \in L^2(\mathbb{R}^d)$$

a teda podľa definície subdiferenciálu je $-\Delta u \in \partial F[u]$. □

3. Wassersteinova metrika

V tejto kapitole zavedieme metriku na priestore pravdepodobnostných mier, ktorý označíme $\mathscr{P}(\mathbb{R}^d)$. Pravdepodobnostné miery určujú náhodné veličiny na $\mathbb{R}^d$, t.j. $X \sim \mu$, kde $\mu \in \mathscr{P}(\mathbb{R}^d)$. Poznačme, že je možné (nad rámec tohoto článku) nahradiť $\mathbb{R}^d$ Hilbertovym priestorom H. V tomto prípade miery určujú rozdelenie stochastických procesov. Priestor H však nie je tzv. *lokálne kompaktný* a teda tieto miery nemajú hustotu, resp. pravdepodobnostnú funkciu.

Kvôli prehľadnosti sa obmedzíme na priestor absolútne spojitých pravdepodobnostných mier voči Lebesgueovej miere, ktorý budeme značiť $\mathscr{P}^{\mathrm{ac}}(\mathbb{R}^d)$. Tento podpriestor určuje rozdelenie spojitých náhodných veličín. Toto má niekoľko dôvodov. Po prvé, popísať vlastnosti, ktoré plánujeme na tomto podpriestore je názornejšie a stručnejšie, ale je možné (avšak technicky náročnejšie) popis rozšíriť na všetky pravdepodobnostné miery. Po druhé, je možné ukázať, že riešenie rovnice vedenia tepla v slabej formulácii určuje v skoro každom čase hustotu absolútne spojitej pravdepodobnostnej miery.

Uvažujeme mieru $\mu \in \mathscr{P}^{\mathrm{ac}}(\mathbb{R}^d)$ s hustotou ϱ. Potom pre každú merateľnú množinu $A \subseteq \mathbb{R}^d$ máme

$$\mu(A) = \int_A \varrho(x)\,\mathrm{d}x\,; \qquad \varrho = \frac{\mathrm{d}\mu}{\mathrm{d}x}\,, \tag{3.1}$$

kde výraz $\mathrm{d}\mu/\mathrm{d}x$ značí tzv. *Radonovu-Nikodýmovu deriváciu* [3, Section 1.6.1]. Keďže hustota ϱ jednoznačne určuje mieru μ, stotožníme označenie miery a hustoty.

Uvedieme si pojem moment náhodnej veličiny, s ktorým budeme pracovať.

Definícia 3.1 (Moment náhodnej veličiny)**.** Nech je rozdelenie náhodnej veličiny X určené hustotou ϱ. Potom jej *moment cez funkciu* $f\colon \mathbb{R}^d \to \mathbb{R}$ definujeme vzťahom

$$\mathbb{E}[f(X)] := \int_{\mathbb{R}^d} f(x)\,\varrho(x)\,\mathrm{d}x.$$

Uvažujeme miery s konečným druhým momentom, $\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d) \subset \mathscr{P}(\mathbb{R}^d)$, t.j.

$$\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d) := \left\{\varrho \in \mathscr{P}^{\mathrm{ac}}(\mathbb{R}^d);\ \int_{\mathbb{R}^d} \|x\|^2\,\varrho(x)\,\mathrm{d}x < +\infty\right\},$$

ekvivalentne

$$\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d) := \left\{\varrho \in \mathscr{P}^{\mathrm{ac}}(\mathbb{R}^d);\ \varrho = \mathrm{Law}(X),\ \mathbb{E}\big[\|X\|^2\big] < +\infty\right\}.$$

Teraz si popíšeme, čo je *push-forward operátor.* Určuje transformáciu mier, resp. zobrazenie medzi nimi. V reči klasickej štatistiky sa jedná zobrazenie náhodných veličín. Ak máme náhodnú veličinu $X \sim \varrho$ a jej transformáciu danú zobrazením T, potom hustotou pravdepodobnosti náhodnej veličiny $T(X)$ značíme $T_{\#}\varrho$ a operátor $T_{\#}$ nazývame push-forward operátor cez zobrazenie T. To znamená, že $T_{\#}$ popisuje transformáciu náhodných veličín v reči pravdepodobnostných mier.

Veta 3.2 (Zmena premenných)**.** *Nech je $T : X \to Y$ zobrazenie, $\varrho \in \mathscr{P}^{\mathrm{ac}}(\mathbb{R}^d)$ hustota pravdepodobnosti a $T_{\#}\varrho$ je jej push-forward hustota pravdepodobnosti cez zobrazenie T. Potom pre všetky prípustné funkcie $f : Y \to \mathbb{R} \cup \{+\infty\}$ platí*

$$\int_{\mathbb{R}^d} f(x)\, T_{\#}\varrho(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(T(x))\, \varrho(x)\,\mathrm{d}x.$$

Popíšeme si formu charakterizácie push-forward operátora pomocou vety o substitúcii integrálu. Túto charakterizáciu využijeme neskôr pri určovaní subdiferenciálu konkrétneho funkcionálu (entropie), viď str. 11.

Veta 3.3 (Zmena premenných cez jakobián)**.** *Nech je $T : \mathbb{R}^d \to \mathbb{R}^d$ zobrazenie, $\varrho \in \mathscr{P}^{\mathrm{ac}}(\mathbb{R}^d)$ je hustota pravdepodobnosti a $\varphi = T_{\#}\varrho$. Ak $\det \nabla T > 0$, potom platí*

$$\varphi(T(x))\, \det \nabla T(x) = \varrho(x) \quad \forall x \in \mathbb{R}^d.$$

Dôkaz. Z vety 3.2 o zmene premenných plynie pre všetky prípustné funkcie $f : Y \to \mathbb{R} \cup \{+\infty\}$

$$\int_{\mathbb{R}^d} f(y)\,\varphi(y)\,\mathrm{d}y = \int_{\mathbb{R}^d} f(y)\,\mathrm{d}T_{\#}\varrho(y)\,\mathrm{d}y = \int_{\mathbb{R}^d} f(T(y))\,\varrho(y)\,\mathrm{d}y.$$

Vykonáme transformáciu $y = T(x)$ v integráli na ľavej strane predchádzajúceho vzťahu a z vety o substitúcii tak dostaneme

$$\int_{\mathbb{R}^d} f(T(x))\varphi(T(x)) \det \nabla T(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} f(T(x))\varrho(x)\,\mathrm{d}x.$$

Keďže funkcia f je ľubovoľná, získame $\varphi(T(x)) \det \nabla T(x) = \varrho(x)$ pre každé $x \in \mathbb{R}^d$. □

Definovali sme všetky potrebné pojmy a môžeme pristúpiť k definícii Wassersteinovej vzdialenosti.

Definícia 3.4 (Wassersteinova vzdialenosť)**.** *Wassersteinova vzdialenosť* medzi dvoma pravdepodobnostnými mierami $\varrho^1, \varrho^2 \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ je definovaná vzťahom

$$W_2(\varrho^1, \varrho^2) := \sqrt{\inf\left\{\mathbb{E}\left[\|X - Y\|^2\right];\ X \sim \varrho^1, Y \sim \varrho^2\right\}}.$$

Poznámka 3.5 (W_2 je optimálna cena prepravy)*.* Wassersteinovu vzdialenosť W_2 vieme interpretovať ako optimálnu cenu prepravy tovaru, ktorý je rozdelený podľa hustoty ϱ^1 na nové miesta (napr. z tovární do predajní), ktoré sú rozdelené podľa hustoty ϱ^2. Keďže pracujeme s hustotami pravdepodobnosti, množstvo tovaru je stále konštantné (integrál pravdepodobnostnej miery cez celý priestor je 1).

Uvažovaním diskrétnych mier (nie náš prípad) by integrálna formulácia prešla na sumu a variačný problém na klasický problém optimalizácie prepravy tovaru s cenou prepravy, ktorá zodpovedá druhej mocnine Euklidovskej vzdialenosti medzi miestami $\|x_i - y_i\|^2$.

Veta 3.6 ([2, Theorem 2.2])**.** *Wassersteinova vzdialenosť W_2 je metrika na priestore pravdepodobnostných mier s konečným druhým momentom $\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$.*

4. Subdiferenciál na priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$

V predchádzajúcej kapitole sme ukázali, že na množine $\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ vieme zaviesť metriku W_2 a získame metrický priestor. Priestor $\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ však nie je ani lineárnym priestorom (pre $d > 1$) a je preto potrebné definovať subdiferenciál iným spôsobom než v kapitole 2.

Felix Otto v článku [8] popísal tento priestor ako istý typ slabej nekonečne rozmernej Riemannovskej variety. Ku každému prvku $\nu \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ zostrojil dotykový priestor (bundle) ako podmnožinu váhového Hilbertovho priestoru $L^2(\nu; \mathbb{R}^d)$, tento priestor je zavedený na str. 10. S takouto úvahou potom bolo možné definovať subdiferenciál ako podmnožinu dotykového priestoru. S použitím subdiferenciálu potom vieme ukázať spojitosť medzi riešeniami niektorých parabolických parciálnych diferenciálnych rovníc (v našom prípade rovnica vedenia tepla) a gradientných tokov na priestore $\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$.

Budeme pracovať s konečným časom $T < +\infty$, avšak rovnaké výsledky sa dajú ukázať aj pre $T \to +\infty$. Najskôr zavedieme pojem absolútne spojitá krivka na metrickom priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$.

Definícia 4.1. Abstraktná funkcia $u\colon [0,T) \to \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ sa nazýva *lokálne absolútne spojitá* vzhľadom k metrike W_2, ak existuje $m \in L_{\mathrm{loc}}([0,T))$ tak, že platí

$$W_2(u(s), u(t)) \leq \int_s^t m(r)\,\mathrm{d}r \quad \text{pre každé } 0 \leq s \leq t < T.$$

Systém funkcií $\{u(t)\}_{t\in[0,T)}$ budeme nazývať (lokálne) *absolútne spojitou krivkou* na metrickom priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$.

Uvažujme rovnicu kontinuity

$$\frac{\partial \varrho}{\partial t} + \operatorname{div}(v\varrho) = 0, \tag{4.1}$$

kde $v\colon \mathbb{R}^d \times [0,T) \to \mathbb{R}^d$ je daná vektorová funkcia, ktorá predpisuje vektor rýchlosti hmoty v čase t a bode x. Riešenie tejto rovnice budeme chápať v zmysle distribúcií, t.j. riešením rovnice (4.1) nazývame funkciu $\varrho\colon \mathbb{R}^d \times [0,T) \to \mathbb{R}$ takú, že pre skoro všetky $t \in [0,T)$ je funkcia $\varrho(\cdot,t)$ z priestoru $\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ a spĺňa

$$\begin{aligned} -\int_{\mathbb{R}^d} \varphi(\cdot,0)\,\varrho(\cdot,0)\,\mathrm{d}x = \int_0^T \int_{\mathbb{R}^d} \varrho \frac{\partial \varphi}{\partial t}\,\mathrm{d}x\,\mathrm{d}t - \int_0^T \int_{\mathbb{R}^d} v \cdot \nabla\varphi\,\varrho\,\mathrm{d}x\,\mathrm{d}t \\ \forall \varphi \in C_c^\infty(\mathbb{R}^d \times [0,T)). \end{aligned} \tag{4.2}$$

Rovnica kontinuity popisuje v hydromechanike šírenie nestlačiteľnej kvapaliny. Keďže transformácia náhodnej veličiny zachováva objem (integrál cez celý priestor je vždy 1), je prirodzené očakávať, že každá absolútne spojitá krivka na priestore pravdepodobnostných mier je generovaná nejakým riešením rovnice (4.1). Pri uvažovaní miery v súvislosti s riešeniami rovnice kontinuity je integrálna podmienka (4.2) prirodzená, pretože miera sama je integrál hustoty, viď (3.1).

Definujme najskôr priestor, ktorého v je prvkom. Z dôsledku [2, Remark 1.22] Brenierovej vety [5, Theorem 2.5.10] plynie, že každý prvok μ (dostatočne malého) okolia hustoty $\nu \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ v metrike W_2 je tvaru $\mu = (\mathsf{Id} + \varepsilon\nabla\varphi)_\#\nu$, kde funkcia φ je z priestoru $C_c^\infty(\mathbb{R}^d)$ a $\varepsilon \in \mathbb{R}$, pričom symbolom Id značíme identitu na $\mathbb{R}^d$. Týmto komentárom sme chceli naznačiť dôvod, prečo priestor vektorových funkcií popisujúcich zmenu hustoty ν definujeme ako uzáver množiny gradientov hladkých funkcií, viď nasledujúca definícia.

Poznamenajme ešte, že pre každé $\nu \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ označíme $L^2(\nu;\mathbb{R}^d)$ Hilbertov priestor vektorových polí $w\colon \mathbb{R}^d \to \mathbb{R}^d$ takých, že $x \mapsto w(x)\cdot w(x)$ je lebesgueovsky integrovateľná na $\mathbb{R}^d$, pričom skalárny súčin v tomto priestore je definovaný vzťahom

$$\langle u, w\rangle_\nu := \int_{\mathbb{R}^d} u(x)\cdot w(x)\,\nu(x)\,\mathrm{d}x.$$

Definícia 4.2 (Dotykový priestor)**.** Nech je $\nu \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ hustota. Potom *dotykový priestor* $\mathrm{Tan}_\nu(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d))$ na $\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ v ν definujeme vzťahom

$$\mathrm{Tan}_\nu(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)) := \overline{\{\nabla\varphi\colon \varphi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\nu;\mathbb{R}^d)}.$$

Je možné dokázať, že $\mathrm{Tan}_\nu(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d))$ je podpriestor priestoru $L^2(\nu;\mathbb{R}^d)$. V nasledujúcej vete ukážeme vzťah medzi absolútne spojitými krivkami na priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$ a riešeniami rovnice kontinuity (4.1).

Veta 4.3 ([2, Theorem 2.29])**.** *Ak je $\{u(t)\}_{t\in[0,T)}$ absolútne spojitá krivka na metrickom priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$, potom funkcia $\varrho\colon \mathbb{R}^d \times [0,T) \to \mathbb{R}$ taká, že $\varrho(\cdot,t) = u(t)$ pre skoro všetky $t \in [0,T)$, je riešením rovnice (4.1) v zmysle distribúcií, v ktorej $v(\cdot,t) \in \mathrm{Tan}_{u(t)}(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d))$ pre skoro všetky $t \in [0,T)$.*

Opačne, keď ϱ je riešením rovnice (4.1) v zmysle distribúcií, potom je systém funkcií $\{\varrho(\cdot,t)\}_{t\in[0,T)}$ absolútne spojitá krivka na $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$.

Teraz pristúpime k definícii Fréchetovho subdiferenciálu na metrickom priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$.

Definícia 4.4 (Fréchetov subdiferenciál na priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$)**.** Nech je $\mathcal{G}\colon \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d) \to \mathbb{R}\cup\{+\infty\}$ funkcionál. *Fréchetovym subdiferenciálom* nazývame viachodnotový operátor $\partial\mathcal{G}\colon \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d) \to 2^{\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)}$ definovaný vzťahom

$$\partial\mathcal{G}[\nu] = \Big\{\xi \in \mathrm{Tan}_\nu(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d));\ \mathcal{G}[\mu] - \mathcal{G}[\nu] \geq \langle \xi, T_\nu^\mu - \mathsf{Id}\rangle_\nu\ \forall \mu \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)\Big\}$$

pre $\nu \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$, kde vektorové pole T_ν^μ spĺňa $\mu = (T_\nu^\mu)_\#\nu$.

Poznámka 4.5. Vyššie sme uviedli, že $\mathrm{Tan}_\nu(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)) \subseteq L_2(\nu;\mathbb{R}^d)$, pravá strana nerovnice v definícii 4.4 je teda definovaná vzťahom $\langle \xi, T_\nu^\mu - \mathsf{Id}\rangle_\nu = \int_{\mathbb{R}^d} \xi(x)\cdot (T_\nu^\mu(x) - x)\,\nu(x)\,\mathrm{d}x$ a vektorové pole T_ν^μ transformuje hustotu ν na hustotu μ.

Teraz zavedieme gradientný tok na metrickom priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$ tak, aby analogicky ako v priestore $L^2(\mathbb{R}^d)$, bol v ľubovoľnom bode každý dotykový vektor záporne vzatý prvok subdiferenciálu daného funkcionálu.

Definícia 4.6 (Charakterizácia gradientného toku pomocou subdiferenciálu, [1, Definition 11.1.1])**.** Povieme, že absolútne spojitá krivka $\{u(t)\}_{t\in[0,T)}$ na metrickom priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$ je *gradientný tok*, ak pre skoro každé $t \in [0,T)$ platí

$$-w \in \partial\mathcal{G}[u(t)] \quad \forall w \in \mathrm{Tan}_{u(t)}(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)).$$

Bez dôkazu uvedieme tvrdenie, ktoré využijeme pri určovaní subdiferenciálu. Toto tvrdenie sa nazýva reťazové pravidlo na $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$.

Tvrdenie 4.7 ([1, Propoposition 10.3.18])**.** *Nech $\nu \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ a $\varphi \in C_c^\infty(\mathbb{R}^d)$. Položme $\nu^\varepsilon = (\mathsf{Id} + \varepsilon\nabla\varphi)_\#\nu$ pre každé $\varepsilon \in \mathbb{R}$. Potom pre každé $w \in L^2(\nu;\mathbb{R}^d)$ platí*

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{F}[\nu^\varepsilon]\Big|_{\varepsilon=0} = \langle w, \nabla\varphi\rangle_\nu \quad \Leftrightarrow \quad w \in \partial\mathcal{F}[\nu].$$

Pristúpime k výsledku, kvôli ktorému sme teóriu gradientných tokov na priestore pravdepodobnostných mier s Wassersteinovou metrikou budovali. Uvažujeme funkcionál popisujúci tzv. vnútornú energiu (resp. záporne vzatú Boltzmannovu entropiu) daný vzťahom

$$\mathcal{F}_e[\nu] = \int_{\mathbb{R}^d} \nu \ln \nu \,\mathrm{d}x \quad \text{pre } \nu \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d).$$

V nasledujúcej vete určíme hodnoty subdiferenciálu funkcionálu $\mathcal{F}_e$ pre všetky prvky priestoru $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$.

Veta 4.8 (Subdiferenciál funkcionálu $\mathcal{F}_e$)**.** *Nech je $\nu \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ také, že $\nu \neq 0$ a gradient $\nabla\nu$ je spojitý na $\mathbb{R}^d$. Potom*

$$\partial\mathcal{F}_e[\nu] = \left\{\frac{\nabla\nu}{\nu}\right\}.$$

Myšlienka dôkazu. Položme $\nu^\varepsilon = (\mathsf{Id} + \varepsilon\nabla\varphi)_\#\nu$ pre každé $\varepsilon \in \mathbb{R}$. Z vety o substitúcii cez jakobián (viď veta 3.3) získame

$$\nu = \nu^\varepsilon(\mathsf{Id} + \varepsilon\nabla\varphi)\det\nabla(\mathsf{Id} + \varepsilon\nabla\varphi) \quad \Rightarrow \quad \nu^\varepsilon(\mathsf{Id} + \varepsilon\nabla\varphi) = \frac{\nu}{\det(I_d + \varepsilon\nabla^2\varphi)},$$

kde I_d je jednotková matica a $\nabla^2\varphi$ značí Hessovu maticu funkcie φ. Pretože pre maticu $A \in \mathbb{R}^{d\times d}$ platí

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\det(I_d + \varepsilon A) = \mathrm{tr}\,A + o(\varepsilon),$$

dostaneme

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\det(I_d + \varepsilon\nabla^2\varphi) = \Delta\varphi + o(\varepsilon),$$

keďže $\operatorname{tr}\nabla^2\varphi = \Delta\varphi$. Pre prehľadnosť položíme $u(z) = z\ln z$ pre $z > 0$. Potom máme

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{F}_e[\nu^\varepsilon] &= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int_{\mathbb{R}^d} u(\nu^\varepsilon)\,\mathrm{d}y \\
&= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int_{\mathbb{R}^d} u(\nu^\varepsilon(\mathsf{Id}+\varepsilon\varphi))\,\det(I_d+\varepsilon\nabla^2\varphi)\,\mathrm{d}x \\
&= \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int_{\mathbb{R}^d} u\Big(\frac{\nu}{\det(I_d+\varepsilon\nabla^2\varphi)}\Big)\,\det(I_d+\varepsilon\nabla^2\varphi)\,\mathrm{d}x \\
&= \int_{\mathbb{R}^d}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}u\Big(\frac{\nu}{\det(I_d+\varepsilon\nabla^2\varphi)}\Big)\,\det(I_d+\varepsilon\nabla^2\varphi)\,\mathrm{d}x.
\end{aligned}$$

Spočítame deriváciu vnútri integrálu nasledovne:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\,u\Big(\frac{\nu}{\det(I_d+\varepsilon\nabla^2\varphi)}\Big)\det(I_d+\varepsilon\nabla^2\varphi) \\
&\qquad = u'\Big(\frac{\nu}{\det(I_d+\varepsilon\nabla^2\varphi)}\Big)\frac{-\nu}{\det^2(I_d+\varepsilon\nabla^2\varphi)}\big(\Delta\varphi+o(\varepsilon)\big) \\
&\qquad\quad + u\Big(\frac{\nu}{\det(I_d+\varepsilon\nabla^2\varphi)}\Big)\big(\Delta\varphi+o(\varepsilon)\big).
\end{aligned}$$

Dosadením $\varepsilon = 0$ získame

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{F}_e[\nu^\varepsilon]\Big|_{\varepsilon=0} = \int_{\mathbb{R}^d}\Big(u'(\nu)(-\nu\Delta\varphi)+u(\nu)\,\Delta\varphi\Big)\,\mathrm{d}x$$

a použitím Greenovej vety dostaneme

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{F}_e[\nu^\varepsilon]\Big|_{\varepsilon=0} = \int_{\mathbb{R}^d}\nabla\big(u'(\nu)\nu-u(\nu)\big)\cdot\nabla\varphi\,\mathrm{d}x. \tag{4.3}$$

Keďže

$$\nabla(\nu u'(\nu)-u(\nu)) = u'(\nu)\,\nabla\nu+\nu\,\nabla u'(\nu)-u'(\nu)\,\nabla\nu = \nu\,\nabla u'(\nu),$$

z (4.3) získame

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{F}_e[\nu^\varepsilon]\Big|_{\varepsilon=0} = \int_{\mathbb{R}^d}\nabla u'(\nu)\cdot\nabla\varphi\,\nu\,\mathrm{d}x = \langle\nabla u'(\nu),\nabla\varphi\rangle_\nu\,.$$

S využitím tvrdenia 4.7 odtiaľ plynie $\nabla u'(\nu)\in\partial\mathcal{F}_e[\nu]$. Zrejme $u'(z) = \ln z + 1$ pre $z > 0$ a preto $\nabla u'(\nu) = \nabla\nu/\nu$, t.j. sme ukázali, že $\nabla\nu/\nu\in\partial\mathcal{F}_e[\nu]$. □

Dôsledok 4.9. *Nech je ϱ riešením problému (1.1) v slabom zmysle s počiatočnou podmienkou $\varrho_0\in\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$. Potom $\{\varrho(\cdot,t)\}_{t\in[0,T)}$ je gradientným tokom na $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d),W_2)$ po funkcionáli $\mathcal{F}_e$ daným vzťahom*

$$\mathcal{F}_e[\nu] = \int_{\mathbb{R}^d}\nu\ln\nu\,\mathrm{d}x \quad \textit{pre } \nu\in\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d).$$

Dôkaz. Z vety 4.8 plynie $\partial\mathcal{F}_e[\varrho(\cdot,t)] = \left\{\frac{\nabla\varrho(\cdot,t)}{\varrho(\cdot,t)}\right\}$ pre skoro každé $t \in [0,T)$. Dosadením $\frac{\nabla\varrho(\cdot,t)}{\varrho(\cdot,t)}$ za v do vzťahu (4.2), ktorý charakterizuje riešenie rovnice kontinuity v zmysle distribúcií, po vykrátení $\varrho(\cdot,t)$ dostaneme slabú formuláciu riešenia rovnice vedenia tepla (1.2). $\square$

Poznámka 4.10. Teória gradientných tokov na priestore $(\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d), W_2)$ je podrobne popísaná v knihe [1] a tiež [10]. Medzi klasické výsledky tejto teórie patrí, že riešenia parciálnej diferenciálnej rovnice

$$\frac{\partial\varrho}{\partial t} = \nabla\cdot(\varrho\nabla V) + \nabla\cdot(\varrho\nabla W * \varrho) + \Delta(\varrho^m); \quad m > \frac{1}{d} \tag{4.4}$$

generujú gradientné toky po funkcionáli popisujúceho voľnú energiu. Prvý sčítanec na pravej strane rovnice reprezentuje potenciálnu energiu (potenciál V reprezentuje napr. gravitačnú energiu), druhý sčítanec interakčnú energiu (sila, ktorou sa priťahujú častice hmoty medzi sebou) a posledný vnútornú energiu (difúziu a pre $m \neq 1$ tzv. nelineárnu difúziu). Podotknime, že kvôli konvolučnému členu a nelineárnej difúzii sa doteraz nepodarilo odvodiť rády konvergencie klasických numerických metód, napr. metódy konečných prvkov alebo konečných objemov pre túto rovnicu.

5. Diskretizácia gradientného toku

Teória popísaná v kapitolách 3 a 4 bola vytvorená na základe pozorovania nemeckého matematika Felixa Otta koncom 90. rokov minulého storočia. Pomocou časovej diskretizácie zistil, že riešenia rovnice vedenia tepla generujú gradientné toky na metrickom priestore pravdepodobnostných mier. Toto pozorovanie si popíšeme na diskretizácii gradientného toku na Euklidovom priestore $\mathbb{R}^d$ po hladkej konvexnej funkcii f a následne ho zovšeobecníme na priestor pravdepodobnostných mier. Uvažujme spätnú/implicitnú Eulerovu diskretizáciu systému obyčajných diferenciálnych rovníc v (1.3) s časovým krokom τ:

$$\begin{aligned}\frac{x_{k+1}-x_k}{\tau} = -\nabla f(x_{k+1}) \quad &\Rightarrow \quad \frac{x_{k+1}-x_k}{\tau} + \nabla f(x_{k+1}) = 0\\ &\Rightarrow \quad \nabla\left(\frac{\|x-x_k\|^2}{2\tau} + f(x)\right)\Big|_{x=x_{k+1}} = 0.\end{aligned}$$

Uvedomme si, že f je konvexná funkcia a výraz $\frac{\|x-x_k\|^2}{2\tau}$ je rýdzo konvexný. Z toho plynie, že výraz $\frac{\|x-x_k\|^2}{2\tau} + f(x)$ je rýdzo konvexný a v bode, kde je jeho gradient nulový, má tento výraz jediné minimum. Teda spätnú Eulerovu diskretizáciu počiatočnej úlohy (1.3) vieme zapísať ako

$$x_{k+1} := \arg\min_{x\in\mathbb{R}^d}\left\{\frac{\|x-x_k\|^2}{2\tau} + f(x)\right\} \quad k = 0,1,2,\ldots \tag{5.1}$$

Poznámka 5.1. Diskretizáciu vo vzťahu (5.1) môžeme písať v tvare $x_{k+1} := \mathrm{prox}_\tau^f(x_k)$, kde prox je tzv. proximálny operátor známy z teórie optimalizácie konvexných funkcií, viď [9].

Je známy výsledok numerických metód, že problém (5.1) konverguje v norme $\|\cdot\|$ k presnému riešeniu úlohy (1.3). Znovu si môžeme všimnúť, že na definovanie tohoto problému potrebujeme len metriku na priestore $\mathbb{R}^d$. Prirodzene tak môžeme navrhnúť časovú diskretizáciu problému gradientného toku na Hilbertovom priestore $L^2(\mathbb{R}^d)$ v definícii 2.2:

$$u_{k+1} := \arg\min_{u\in L^2(\mathbb{R}^d)} \left\{ \frac{\|u-u_k\|_{L^2}^2}{2\tau} + F[u] \right\} \quad k = 0, 1, 2, \dots \tag{5.2}$$

V knihe [4, Section 9.6, Theorem 2] je ukázané, že diskretizácia (5.2) konverguje v norme priestoru $L^2(\mathbb{R}^d)$ k presnému riešeniu.

Teraz prejdime k spomínanému spojeniu s priestorom $\mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ absolútne spojitých pravdepodobnostných mier určených hustotou ϱ. Jordan, Kindelehrer a Otto ukázali v článku [6], že časová diskretizácia (ktorej dnes hovoríme JKO schéma)

$$\varrho_{k+1} := \arg\min_{\varrho\in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R}^d)} \left\{ \frac{W_2^2(\varrho, \varrho_k)}{2\tau} + \int_{\mathbb{R}^d} \varrho \ln \varrho \, \mathrm{d}x \right\} \quad k = 0, 1, 2, \dots \tag{5.3}$$

konverguje k slabému riešeniu rovnice vedenia tepla $\frac{\partial \varrho}{\partial t} = \Delta \varrho$ vo Wassersteinovej metrike W_2. Hlavným argumentom pre význam tohoto tvrdenia je, že integrál $-\int \varrho \ln \varrho \, dx$ je Boltzmannova entropia a príroda sa skutočne pri zanedbaní ostatných vplyvov podľa fyzikálnych zákonov snaží túto entropiu čo najrýchlejšie navýšiť, t.j. čo najstrmšie klesať po funkcionáli vo vzťahu (5.3).

Keďže JKO schéma (5.3) napodobňuje spätnú/implicitnú diskretizáciu gradientného toku na Euklidovom, resp. Hilbertovom priestore, intuitívne by sa dalo očakávať, že by mala predstavovať diskretizáciu gradientného toku na priestore pravdepodobnostných mier s Wassersteinovou metrikou. Toto pozorovanie viedlo k podrobnejšiemu štúdiu jednak Wassersteinovej metriky [10] a jednak gradientných tokov. Tu je významná kniha [1], kde sa v kapitolách 1–4 opisujú toky na všeobecných metrických priestoroch a v kapitolách 5–12 na konkrétnom priestore pravdepodobnostných mier s Wassersteinovou metrikou.

6. Stochastické diferenciálne rovnice

V kapitole 4 sme ukázali, že riešenia rovnice vedenia tepla v zmysle distribúcií generujú absolútne spojité krivky na priestore pravdepodobnostných mier. Ukážeme, že riešeniam stochastických diferenciálnych rovníc odpovedajú riešenia niektorých parabolických parciálnych diferenciálnych rovníc. Budeme pracovať s pojmom normálne rozdelenie, takže si tento pojem zavedieme. Pre jednoduchosť a názornosť budeme v celej tejto kapitole pracovať v jednorozmernom prípade, t.j. uvažujeme $d = 1$.

Definícia 6.1 (Normálne rozdelenie, Gaussovská miera)**.** Povieme, že náhodná veličina X určená hustotou $\varrho \in \mathscr{P}_2^{\mathrm{ac}}(\mathbb{R})$, t.j. $X \sim \varrho$, má normálne rozdelenie (ϱ je Gaussovská miera) ak

$$\varrho(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{(x-\mu)^2}{2\sigma^2} \right) \quad \text{pre } x \in \mathbb{R},$$

kde $\mu \in \mathbb{R}$ a $\sigma > 0$. Píšeme

$$X \sim \mathsf{N}(\mu, \sigma^2) \qquad \text{alebo} \qquad \varrho = \mathsf{N}(\mu, \sigma^2).$$

Veta 6.2 (Momenty náhodnej veličiny)**.** *Ak je* $X \sim \mathsf{N}(\mu, \sigma^2)$*, potom*

$$\mathbb{E}[X] = \mu \qquad a \qquad \mathbb{E}[X^2] = \sigma^2 + \mu^2.$$

V tejto časti intuitívne popíšeme dôležité vlastnosti stochastických obyčajných diferenciálnych rovníc. Medzi klasické motivačné príklady patrí nasledujúca úloha (viď [7]). Ak v obyčajnej diferenciálnej rovnici

$$\frac{\mathrm{d}x}{\mathrm{d}t} = Ax$$

uvažujeme parameter A ako normálne rozdelenú náhodnú veličinu $A \sim \mathsf{N}(a, \sigma^2)$, dostaneme Itôovu stochastickú diferenciálnu rovnicu

$$\mathrm{d}X_t = a\, X_t\, \mathrm{d}t + \sigma\, X_t\, \mathrm{d}B_t\,,$$

kde B_t predstavuje Brownov pohyb (inak povedané Wienerov proces). Stochastická rovnica teda hovorí, že systém sa v každom bode správa približne podľa rovnice $\mathrm{d}X/\mathrm{d}t = a$ s odchýlkou popísanou šumom $\sigma\, \mathrm{d}B_t$. Poznamenajme, že tento zápis sa užíva preto, nakoľko Brownov pohyb nie je diferencovateľný (viď jeho definícia neskôr).

Príklad 6.3 (Stochastický gradientný tok)*.* Gradientný tok na Euklidovom priestore $\mathbb{R}$ generovaný počiatočnou úlohou (1.3) s $d = 1$, ktorý sa vychyľuje v každom bode o šum (t.j. sa pohybuje len približne v smere najväčšieho klesania funkcie) sa dá popísať stochastickou rovnicou

$$\mathrm{d}X_t = -\nabla f(X_t)\, \mathrm{d}t + \sqrt{2}\, \sigma\, \mathrm{d}B_t\,.$$

Ak je funkcia f rýdzo konvexná, potom existuje jediné $x_\star := \arg\min_{x\in\mathbb{R}}\{f(x)\}$, ktoré je zároveň stacionárnym riešením skalárnej obyčajnej diferenciálnej rovnice $\dot{x} = -\nabla f(x)$. V prípade stochastických diferenciálnych rovníc stacionárnym riešením nie je presne tento bod, ale rozdelenie pravdepodobnosti dané hustotou ϱ_∞ (tak, že $X_\infty \sim \varrho_\infty$), ktorej modus je práve hodnota $x_\star$. Je možné ukázať, že táto hustota má tvar

$$\varrho_\infty(x) = \frac{\exp(-\frac{1}{\sigma} f(x))}{\int_{\mathbb{R}} \exp(-\frac{1}{\sigma} f(x))\, \mathrm{d}x} \quad \text{pre } x \in \mathbb{R}.$$

Tu si môžeme všimnúť, že čím je hodnota σ väčšia, tým je rozptyl rozdelenia ϱ_∞ väčší. Toto je konzistentné s intuitívnym očakávaním, že čím zavádzame väčšiu neurčitosť do modelu, tým s väčšou nepresnosťou dosiahneme minimum funkcie f.

V prípade, že funkcia f je daná vzťahom $f(x) = \frac{x^2}{2}$ pre $x \in \mathbb{R}$, stacionárne rozdelenie ϱ_∞ je Gaussovská miera $\varrho_\infty = \mathsf{N}(0, 1)$.

Teraz definujeme Brownov pohyb, ktorý sa tiež nazýva Wienerov proces.

Definícia 6.4 (Brownov pohyb na $\mathbb{R}$)**.** Stochastický proces $\{B_t\}_{t\geq 0}$ sa nazýva Brownov pohyb (tiež Wienerov proces), ak spĺňa nasledujúce podmienky:

1. (počiatočná podmienka) $B_0 = 0$;

2. (nezávislosť prírastkov) Pre všetky hodnoty $0 < t_1 < \cdots < t_k$ sú náhodné premenné $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_k} - B_{t_{k-1}}$ vzájomne nezávislé;
3. (rozdelenie prírastkov) Pre všetky $0 \le s < t < +\infty$ platí
$$B_t - B_s \sim \mathsf{N}(0, t-s)\,;$$
4. (spojitosť) Skoro iste je zobrazenie $t \mapsto B_t$ spojité.

Poznámka 6.5. Píšeme
$$\int_s^t \mathrm{d}B_t = B_t - B_s \sim \mathsf{N}(0, t-s).$$
Keďže
$$\int_0^t \mathrm{d}B_t = B_t \sim \mathsf{N}(0,t), \quad \text{platí} \quad \mathbb{E}[B_t] = 0\,,\ \mathbb{E}[B_t^2] = t.$$

Nasledujúca lemma je základným prvkom stochastického počtu. Na prvý pohľad je zvláštna, pretože podľa nej v prvom diferenciáli transformovanej náhodnej veličiny vystupuje nenulová druhá priestorová derivácia transformačnej funkcie. Toto vychádza z vlastnosti Brownovho pohybu, druhý moment $\mathbb{E}[B_t^2]$ je totiž rádu t. Jej dôkaz spočíva v použití Taylorovho rozvoja a následnom využití spomenutej vlastnosti $\mathrm{d}B_t^2 \approx \mathrm{d}t$.

Lemma 6.6 (Itôova, [7, Theorem 4.1.2])**.** *Nech $a \in \mathbb{R}$, $b \in \mathbb{R}_0^+$ a X_t je riešenie Itôovej stochastickej diferenciálnej rovnice*
$$\mathrm{d}X_t = a\,\mathrm{d}t + b\,\mathrm{d}B_t$$
s jednorozmerným Brownovym pohybom B_t. Nech ďalej $f\colon \mathbb{R} \times [0,+\infty) \to \mathbb{R}$ je dostatočne hladká funkcia. Potom
$$Y_t = f(X_t, t)$$
je riešením stochastickej diferenciálnej rovnice
$$\mathrm{d}Y_t = \Big(\frac{\partial f}{\partial t} + a\frac{\partial f}{\partial x} + \frac{b^2}{2}\frac{\partial^2 f}{\partial x^2}\Big)\mathrm{d}t + b\,\frac{\partial f}{\partial x}\,\mathrm{d}B_t\,.$$

Teraz si ukážeme prepojenie Brownovho pohybu (jednoduchej SDR) a rovnice vedenia tepla (opäť pre jednorozmerný prípad, všetky výsledky sa dajú samozrejme získať aj pre vyššie dimenzie).

Veta 6.7. *Nech je X_t riešenie stochastickej diferenciálnej rovnice*
$$\mathrm{d}X_t = \sqrt{2}\,\mathrm{d}B_t, \tag{6.1}$$
kde B_t je jednorozmerný Brownov pohyb. Potom funkcia $\varrho\colon \mathbb{R}\times[0,T) \to \mathbb{R}$ taká, že $X_t \sim \varrho(\cdot, t)$ pre skoro všetky $t \in [0,T)$, je riešením rovnice vedenia tepla v slabom zmysle.

Dôkaz. Položme $Y_t = \varphi(X_t, t)$, kde $\varphi \in C_c^\infty(\mathbb{R} \times [0,T))$. Spojito dodefinujeme $\varphi(\cdot, T) = 0$, aplikujme Itôovu lemmu (v ktorej $a = 0$, $b = \sqrt{2}$) a zistíme, že Y_t je riešením stochastickej diferenciálnej rovnice
$$\mathrm{d}Y_t = \Big(\frac{\partial\varphi}{\partial t} + \frac{\partial^2\varphi}{\partial x^2}\Big)\,\mathrm{d}t + \sqrt{2}\,\frac{\partial\varphi}{\partial x}\,\mathrm{d}B_t\,.$$

Uvažujeme integrálnu formuláciu SDR (obe strany zintegrujeme na $[0,T]$):

$$\begin{aligned}
\int_0^T \mathrm{d}Y_t &= \int_0^T \Big(\frac{\partial\varphi}{\partial t} + \frac{\partial^2\varphi}{\partial x^2}\Big)\,\mathrm{d}t + \int_0^T \sqrt{2}\,\frac{\partial\varphi}{\partial x}\,\mathrm{d}B_t,\\
Y_T - Y_0 &= \int_0^T \Big(\frac{\partial\varphi}{\partial t} + \frac{\partial^2\varphi}{\partial x^2}\Big)\,\mathrm{d}t + \int_0^T \sqrt{2}\,\frac{\partial\varphi}{\partial x}\,\mathrm{d}B_t,\\
\varphi(X_T,T) - \varphi(X_0,0) &= \int_0^T \Big(\frac{\partial\varphi}{\partial t} + \frac{\partial^2\varphi}{\partial x^2}\Big)\,\mathrm{d}t + \int_0^T \sqrt{2}\,\frac{\partial\varphi}{\partial x}\,\mathrm{d}B_t.
\end{aligned}$$

Teraz spočítame na oboch stranách strednú hodnotu vzhľadom k X_t, vynásobíme teda jednotlivé členy ϱ a integrujeme cez $\mathbb{R}$. Vzhľadom k vlastnosti Brownovho pohybu bude stredná hodnota integrálu úplne vpravo nulová. Získame teda

$$-\int_{\mathbb{R}} \varphi(\cdot,0)\varrho(\cdot,0)\,\mathrm{d}x = \int_0^T \int_{\mathbb{R}} \frac{\partial\varphi}{\partial t}\varrho\,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_{\mathbb{R}} \frac{\partial^2\varphi}{\partial x^2}\,\varrho\,\mathrm{d}x\,\mathrm{d}t.$$

Odtiaľ integráciou per partes dostávame

$$-\int_{\mathbb{R}} \varphi(\cdot,0)\varrho(\cdot,0)\,\mathrm{d}x = \int_0^T \int_{\mathbb{R}} \frac{\partial\varphi}{\partial t}\varrho\,\mathrm{d}x\,\mathrm{d}t - \int_0^T \int_{\mathbb{R}} \frac{\partial\varphi}{\partial x}\,\frac{\partial\varrho}{\partial x}\,\mathrm{d}x\,\mathrm{d}t \qquad \forall\varphi \in C_c^\infty(\mathbb{R}\times[0,T)),$$

pretože funkcia φ bola ľubovoľná. Ukázali sme teda, že funkcia ϱ je riešením rovnice vedenia tepla v slabom zmysle (1.2). □

Dôsledok 6.8. *Nech je ϱ riešením rovnice vedenia tepla v slabom zmysle s počiatočnou podmienkou $\varrho_0 \sim \mathsf{N}(0,a_0)$, kde $a_0 \in \mathbb{R}^+$. Potom pre skoro každé $t > 0$ platí $\varphi(\cdot,t) \sim \mathsf{N}(0,a_0+2t)$.*

Dôkaz. Prevedieme (6.1) do integrálnej formulácie a využijeme vlastnosti Brownovho pohybu (viď definícia 6.4) a druhého momentu (resp. rozptylu) náhodnej veličiny. □

Poznámka 6.9. Parciálne diferenciálne rovnice tvaru (4.4) majú svoju stochastickú interpretáciu v tvare tzv. McKean-Vlasovej SDR

$$\mathrm{d}X_t = -\big(\nabla V(X_t) + \nabla W(X_t) * \varrho(\cdot,t)\big)\,\mathrm{d}t + \sqrt{2\varrho^{m-1}}\,\mathrm{d}B_t\,; \quad X_t \sim \varrho(\cdot,t).$$

7. Záver

V tomto článku sme pracovali s rôznymi interpretáciami rovnice vedenia tepla.

Najskôr sme definovali pojem gradientného toku na Hilbertovom priestore lebesgueovsky integrovateľných funkcií a ukázali sme, že riešenia rovnice vedenia tepla generujú gradientné toky po Dirichletovej energii.

Intuícia nemeckého matematika Felixa Otta viedla k tomu, že priestor pravdepodobnostných mier by mohol byť vhodným na popis riešenia rovnice vedenia tepla. Pozoroval, že riešenie tejto rovnice konverguje k minimalizácií funkcionálu záporne vzatej Boltzmannovej entropie, ktorý má fyzikálny význam pre túto rovnicu (teplo sa skutočne šíri tak, aby maximalizovalo Boltzmannovu entropiu). Toto viedlo na štúdium metrického priestoru mier, ktorého základné vlastnosti sme si popísali. Ďalej sme ukázali, že riešenia rovnice vedenia tepla generujú gradientné toky po funkcionáli zápornej Boltzmannovej entropie.

Pozorovanie, že riešenie parciálnej diferenciálnej rovnice môžeme rozumieť ako krivku na priestore náhodných veličín nás posunulo k úvahe, že by táto krivka mohla byť zároveň riešením stochastického procesu (stochastickej diferenciálnej rovnice). Ukázali sme, že riešenie Brownovho pohybu (Wienerovho procesu), konkrétne jeho hustota, je slabým riešením rovnice vedenia tepla.

Poďakovanie

Autor ďakuje kolegom za cenné pripomienky a návrhy na výrazné vylepšenie textu.

Literatúra

[1] L. Ambrosio, N. Gigli, G. Savaré: *Gradient flows: in metric spaces and in the space of probability measures*, Lectures in mathematics ETH Zürich, Basel: Birkhäuser, 2008.

[2] L. Ambrosio, N. Gigli: *User's guide to optimal transport*, Modelling and Optimisation of Flows on Networks, Lecture Notes in Mathematics, Berlin, Springer, 2013, 1–155.

[3] L. Evans: *Measure Theory and Fine Properties of Functions*, Revised Edition, Boca Raton: CRC Press, Taylor & Francis Group, 2015.

[4] L. Evans: *Partial differential equations*, Graduate studies in mathematics, Providence, Rhode Island: American Mathematical Society, 2022.

[5] A. Figalli, F. Glaudo: *An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows*, Berlin: European Mathematical Society, 2021.

[6] R. Jordan, D. Kinderlehrer, F. Otto: *The Variational Formulation of the Fokker–Planck Equation*, SIAM Journal on Mathematical Analysis, 1998, roč. 29, č. 1, 1–17.

[7] B. Øksendal: *Stochastic differential equations: an introduction with applications*, Heidelberg: Springer, 2013.

[8] F. Otto: *The Geometry of dissipative evolution equations: The porous medium equation*, Communications in Partial Differential Equations, 2001, roč. 26, č. 1–2, 101–174.

[9] N. Parikh: *Proximal Algorithms*, Foundations and Trends in Optimization, 2014, roč. 1, č. 3, s. 127–239.

[10] C. Villani: *Optimal transport: old and new*, Grundlehren der mathematischen Wissenschaften, A series of comprehensive studies in Mathematics, Berlin: Springer, 2009.

Matej Benko, Ústav matematiky, Fakulta strojního inženýrství, Vysoké učení technické v Brně, Technická 2, 616 69 Brno, Česká republika,

e-mail: `Matej.Benko@vutbr.cz`