

# Zeit-Diskretisierung von Gradientenflüssen

## 1 Generalisierte Gradientenflüsse

Wir betrachten folgenden Gradientenfluss:

$$\frac{dX}{dt} = -\text{grad } E(X), \quad X(0) = X^0 \quad (1.1)$$

$E : \mathbb{R}^k \rightarrow \mathbb{R}$  sei ein glattes, nach unten beschränktes Energie-Funktional, das außerdem  $E(X^0) < \infty$  erfülle.  $X(t) : [0, \infty) \rightarrow \mathbb{R}^k$  können wir z.B. als Position gegebener Partikel interpretieren. Wir wollen von einer eindeutigen Lösung ausgehen (z.B. via Picard-Lindelöf).

Statt dieses Problem explizit zu lösen, ist unser Ziel, den Gradientenfluss als Grenzwert eines zeitlich diskretisierten Problems darzustellen, sodass jeglicher expliziter Umgang mit Subdifferentialen und dem Gradientenoperator umgangen wird.

Variationelle Konstruktion: Der **approximierte Gradientenfluss**  $(\bar{X}_\tau)$  mit der euklidischen Distanz  $\|\cdot\|$  sei gegeben durch:

Sei die zeitliche Schrittweite  $\tau > 0$  gegeben,  $n \in \mathbb{N}_0$ , dann ist die Folge  $(X_\tau^n)_{n \geq 0}$ :

- i)  $X_\tau^0 := X_0$ ,
- ii)  $X_\tau^{n+1}$  ist ein Minimierer des Minimierungsproblems

$$\min_{X \in \mathbb{R}^k} \left[ E(X) + \frac{\|X_\tau^n - X\|^2}{2\tau} \right] \quad (1.2)$$

Sei  $\bar{X}_\tau : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^k$  als stückweise konstante Funktion mit Wert  $X_\tau(t) = X_\tau^n$  für  $t \in [n\tau, (n+1)\tau)$  definiert.

**Bemerkung 1.1.** 1. Wir setzen zunächst keine Eindeutigkeit des Minimums voraus;

2. Existenz des Minimierers muss durch geeignete Eigenschaften von  $E(X)$  gesichert werden, z.B. strikte Konvexität, untere Schranke.

Wir beginnen damit, den Zusammenhang zwischen dem approximated Gradientenfluss und dem Ausgangsproblem zu untersuchen.

**Lemma 1.2.**  $X_\tau^m, X_\tau^{m+1} \subset (X_\tau^n)_{n \geq 0}$  erfüllen für  $\omega \in \mathbb{R}^k$ :

$$-\text{grad } E(X_\tau^{m+1}) \cdot \omega = \frac{X_\tau^{m+1} - X_\tau^m}{\tau} \cdot \omega \quad (1.3)$$

*Beweis.* Dabei wird eine Strategie eingeführt, der auch - in einem etwas umfangreicheren Setting - der nächste Abschnitt über die Fokker-Planck-Gleichung folgt.

Wir befinden uns im  $\mathbb{R}^k$ . Sei  $\omega$  ein beliebiger (Tangenten-)Vektor auf  $X_\tau^{n+1}$ . Wir betrachten Perturbationen von  $X_\tau^{n+1}$  "entlang  $\omega$ ": Definiere für ein kleines  $\varepsilon > 0$  einen "Pfad"  $\tilde{X}_\varepsilon$  folgendermaßen:

$$\tilde{X}_0 = X_\tau^{n+1}, \quad \left. \frac{d\tilde{X}_\varepsilon}{d\varepsilon} \right|_{\varepsilon=0} = \omega \quad (1.4)$$

Essentiell für die folgenden Schritte ist, dass die Ableitung an der Stelle  $\varepsilon$  gleich Null ist. Der Einfachheit halber nehmen wir die lineare Form  $\tilde{X}_\varepsilon = X_\tau^{n+1} + \varepsilon\omega$  an.

(Anmerkung: Wenn wir allgemeinere Perturbationen betrachten wollen, verfahren wir ganz analog, indem wir nach Taylor

$$\tilde{X}_\varepsilon = X_0 + \left. \frac{d\tilde{X}_\varepsilon}{d\varepsilon} \right|_{\varepsilon=0} \cdot \varepsilon + O(\varepsilon^2), \quad (1.5)$$

$$E(\tilde{X}_\varepsilon) = E(X_\tau^{n+1}) + \varepsilon \langle \text{grad } E(X_\tau^{n+1}), \omega \rangle + O(\varepsilon^2) \quad (1.6)$$

verwenden.)

Da  $X_\tau^{n+1}$  Minimierer ist, gilt offensichtlich

$$E(X_\tau^{n+1}) + \frac{\|X_\tau^n - X_\tau^{n+1}\|^2}{2\tau} \leq E(\tilde{X}_\varepsilon) + \frac{\|X_\tau^n - \tilde{X}_\varepsilon\|^2}{2\tau} \quad \forall \varepsilon > 0 \quad (1.7)$$

$$\stackrel{1.}{\Leftrightarrow} \frac{E(X_\tau^{n+1}) - E(\tilde{X}_\varepsilon)}{\varepsilon} \leq \frac{\|X_\tau^n - \tilde{X}_\varepsilon\|^2 - \|X_\tau^n - X_\tau^{n+1}\|^2}{2\tau\varepsilon} \quad (1.8)$$

$$\Leftrightarrow -\frac{E(X_\tau^{n+1} + \varepsilon\omega) - E(X_\tau^{n+1})}{\varepsilon} \leq \frac{\|X_\tau^n - X_\tau^{n+1} - \varepsilon\omega\|^2 - \|X_\tau^n - X_\tau^{n+1}\|^2}{2\tau\varepsilon} \quad (1.9)$$

Wir schreiben nun vor beide Seiten  $\lim_{\varepsilon \rightarrow 0}$  und interpretieren diese Differenzenquotienten als Ableitungen nach  $\varepsilon$ . Wir erhalten (mit  $\frac{d}{d\varepsilon} \|X_\tau^n - X_\tau^{n+1} - \varepsilon\omega\|^2 = \frac{d}{d\varepsilon} (\dots)_1^2 + (\dots)_2^2 + \dots = 2(\dots)_1 \cdot (-\omega_1) + (\dots)$ ):

$$-\text{grad } E(X_\tau^{n+1}) \cdot \omega \leq \frac{\langle X_\tau^{n+1} - X_\tau^n, \omega \rangle}{\tau} \quad (1.10)$$

Gleichheit wird durch die Betrachtung von  $-\omega$  erreicht.  $\square$

Bevor wir  $\tau \rightarrow 0$  gehen lassen und betrachten, ob ein Grenzwert existiert und tatsächlich unsere Ausgangsgleichung 1.1 erfüllt, wollen wir zunächst einige Eigenschaften untersuchen.

## 1.1 Drei Ungleichungen

$X_\tau^{n+1}$  ist ein Minimierer des Funktionals  $X \mapsto E(X) + \frac{\|X_\tau^n - X\|^2}{2\tau}$ . Damit ist offensichtlich das Funktional an der Stelle  $X_\tau^{n+1}$  kleiner gleich dessen Wert an der Stelle  $X_\tau^n$  ( $\|X_\tau^n - X_\tau^n\| = 0$ ). Damit haben wir

$$E(X_\tau^{n+1}) + \frac{\|X_\tau^n - X_\tau^{n+1}\|^2}{2\tau} \leq E(X_\tau^n). \quad (1.11)$$

**Lemma 1.3** (Energy estimate).

$$\sup_{n \geq 0} E(X_\tau^n) \leq E(X^0) \quad (1.12)$$

*Beweis.* Folgt sofort aus (1.11), da der zweite Term nichtnegativ und somit die Folge der  $(E(X_\tau^n))$  für steigendes  $n$  monoton abnehmend ist.  $\square$

**Lemma 1.4.**

$$\frac{\tau}{2} \sum_{n \geq 0} \left( \frac{\|X_\tau^n - X_\tau^{n+1}\|}{\tau} \right)^2 \leq E(X^0) - \inf_{X \in \mathbb{R}^k} E(X) \quad (1.13)$$

*Interpretation als kinetische Energie!*

*Beweis.* Summation über (1.11), Teleskopsumme, Infimum da  $E(X_\tau^n)$  monoton fallend. Infimum entweder über  $X_\tau^n$  oder  $X \in \mathbb{R}^k$ , möglich wegen unterer Schranke an Potential.  $\square$

**Lemma 1.5** (Hölder 1/2-Abschätzung für  $\bar{X}_\tau$ ).

Für  $s < t$  gilt:

$$\|\bar{X}_\tau(s) - \bar{X}_\tau(t)\|^2 \stackrel{1)}{\leq} \left[ \frac{t-s}{\tau} + 1 \right] \sum_{\frac{s}{\tau}-1 \leq n \leq \frac{t}{\tau}-1} \|X_\tau^n - X_\tau^{n+1}\|^2 \stackrel{2)}{\leq} 2(E(X^0) - \inf_X E)[(t-s) + \tau]. \quad (1.14)$$

*Beweis.* 2) folgt wiederum aus Aufsummation von (1.11) und Monotonie.

1) Nach Definition:  $\bar{X}_\tau(s) = X_\tau^n$  für  $s \in [n\tau, (n+1)\tau)$ . Damit gilt für  $n$ :  $n = \lfloor \frac{s}{\tau} \rfloor$ . Analog sei  $\bar{X}_\tau(t) = X_\tau^m$ ,  $m = \lfloor \frac{t}{\tau} \rfloor$ .

$$\begin{aligned} \|\bar{X}_\tau(s) - \bar{X}_\tau(t)\| &= \|X_\tau^n - X_\tau^m\| \\ &\stackrel{\Delta-Ungl.}{\leq} \|X_\tau^n - X_\tau^{n+1}\| + \|X_\tau^{n+1} - X_\tau^{n+2}\| + \dots + \|X_\tau^{m-1} - X_\tau^m\| \\ &= \sum_{n \leq k \leq m-1} \|X_\tau^k - X_\tau^{k+1}\| = \sum_{\frac{s}{\tau}-1 \leq k \leq \frac{t}{\tau}-1} \|X_\tau^k - X_\tau^{k+1}\| \\ \|\bar{X}_\tau(s) - \bar{X}_\tau(t)\|^2 &\leq \left( \sum_{\frac{s}{\tau}-1 \leq k \leq \frac{t}{\tau}-1} \|X_\tau^k - X_\tau^{k+1}\| \right)^2 \\ &\stackrel{CS}{\leq} \left( \sum_{\frac{s}{\tau}-1 \leq k \leq \frac{t}{\tau}-1} \|X_\tau^k - X_\tau^{k+1}\|^2 \right) \underbrace{\left( \sum_{\frac{s}{\tau}-1 \leq k \leq \frac{t}{\tau}-1} 1 \right)}_{\frac{t}{\tau} - (\frac{s}{\tau} - 1)} \end{aligned}$$

□

**Bemerkung 1.6.** Wir haben hier nie die besonderen Eigenschaften der euklidischen Norm benutzt! Diese Ungleichungen gelten auch bei approximierten Gradientenflüssen jeder beliebigen anderen Norm.

Diese drei Abschätzungen stellen uns die relative Kompaktheit von  $(\bar{X}_\tau)$  für  $\tau \rightarrow 0$  sicher, und damit die Existenz einer konvergenten Teilfolge.

**Corollar 1.7.**

$$\bar{X}_\tau(t) \xrightarrow{\tau \rightarrow 0} X \in \mathcal{C}^{1/2}([0, \infty); \mathbb{R}^k) \text{ gleichmäßig bzgl. } t \geq 0,$$

und es gilt für alle  $0 \leq t_1 \leq t_2 < \infty$ :

$$-\int_{t_1}^{t_2} \text{grad } E(X(t)) \cdot \omega dt = [X(t_2) - X(t_1)] \cdot \omega \quad (1.15)$$

*Beweis.* Seien  $t_1, t_2 > 0$ . Wir beginnen mit (1.3) und summieren von  $n_1 = \lfloor \frac{t_1}{\tau} \rfloor$  bis  $n_2 = \lfloor \frac{t_2}{\tau} \rfloor - 1$ :

$$\sum_{n_1}^{n_2} \frac{1}{\tau} (X_\tau^{m+1} - X_\tau^m) \cdot \omega = \sum_{n_1}^{n_2} -\text{grad } E(X_\tau^{m+1}) \cdot \omega \quad (1.16)$$

Beachte für die rechte Seite, dass

$$-\text{grad } E(X^n) = \frac{1}{\tau} \int_{n\tau}^{n\tau+\tau}.$$

Damit haben wir - linke Seite Teleskopsumme -

$$\frac{1}{\tau} (X_\tau^{n_2+1} - X_\tau^{n_1}) \cdot \omega = \frac{1}{\tau} \int_{(n_1+1)\tau}^{(n_2+2)\tau} -\text{grad } E(\bar{X}_\tau(t)) \cdot \omega dt \quad (1.17)$$

$$\Leftrightarrow (\bar{X}_\tau(t_2) - \bar{X}_\tau(t_1)) \cdot \omega = \int_{t_1}^{t_2} -\text{grad } E(\bar{X}_\tau(t)) \cdot \omega dt + O(\tau), \quad (1.18)$$

wobei  $O(\tau)$  daraus folgt, dass wir ein Integral über  $< \tau$  eine (E glatt, endlich..) beschränkte Funktion bilden.

Die vorherigen Ungleichungen sichern uns die Existenz einer konvergenten (Teil-)Folge von  $\bar{X}_\tau$  für  $\tau \rightarrow 0$ .  $-\text{grad } E$  ist glatt und konvergiert damit ebenfalls gegen  $-\text{grad } E(X)$ .  $O(\tau)$  konvergiert natürlich gegen Null. Wir bekommen damit eine schwache Formulierung unseres ursprünglichen Gradientenflusses (1.1)!

$$\int_{t_1}^{t_2} \text{grad } E(X(t)) \cdot \omega dt = [X(t_2) - X(t_1)] \cdot \omega$$

□

## 2 Die lineare Fokker-Planck-Gleichung, Monge-Kantorovich und der approximierte Gradientenfluss

(Notation in diesem Abschnitt:  $\rho = \rho(x, t)$ )

Diese Strategie wurde zum ersten Mal am Beispiel der linearen Fokker-Planck-Gleichung angewandt.

**Definition 2.1.** Die **lineare Fokker-Planck-Gleichung (FP)** mit glattem Potential  $V : \mathbb{R}^d \rightarrow [0, \infty)$  und Anfangswert  $\rho_0 \in \mathcal{P}_{ac,2}$ , absolut stetige Wahrscheinlichkeitsdichte mit zweitem endlichen Moment auf  $\mathbb{R}^d$ , ist gegeben durch:

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \Delta \rho + \nabla \cdot (\rho \nabla V), \\ \rho(0) &= \rho_0. \end{aligned} \quad (2.1)$$

Alle Lösungen  $\rho(x, t)$  sind ebenfalls abs. stetige Wahrscheinlichkeitsdichten auf  $\mathbb{R}^d$  für fast alle festen  $t$ , d.h.  $\rho(x, t) \geq 0$  für fast alle  $(x, t) \in \mathbb{R}^d \times (0, \infty)$  und  $\int_{\mathbb{R}^d} \rho(x, t) dx = 1$  für fast alle  $t \in (0, \infty)$ . Eine Wahrscheinlichkeitsdichte  $\rho \in \mathcal{P}_{ac,2}$  erfüllt die (LFP) im schwachen Sinne, falls für alle Testfunktionen  $\psi$  gilt (Herleitung wie üblich, Testfunktion und aufintegrieren):

$$\int_{\mathbb{R}^d} \int_T (\rho \Delta \psi + \rho \nabla V \cdot \nabla \psi - \rho \psi_t) dt dx = \int_{\mathbb{R}^d} \rho_0 \psi(x, 0) dx \quad (2.2)$$

Die Fokker-Planck Gleichung beschreibt die zeitabhängige Entwicklung einer Wahrscheinlichkeitsdichte für einen stochastischen Prozess. Sie wird verwendet, um die Wahrscheinlichkeitsdichte für Ort oder Geschwindigkeit eines Partikels zu beschreiben, das verschiedenen Kräften ausgesetzt ist.

**Bemerkung 2.2.** Wenn  $V$  geeignete Wachstumsbedingungen erfüllt, existiert eine eindeutige stationäre Lösung zur (LFP), die sog. Gibbsche Verteilung.

Auch wenn die explizite Lösung der Gleichung bereits bekannt war, liefert die Anwendung des hier zu zeigenden Approximationsverfahren interessante Einblicke in die Methode, die auch für andere Diffgleichungen und in numerischer Implementierung Anwendung findet.

Wir konstruieren ein zeitlich diskretisiertes, iteratives variationelles Verfahren (d.h. die Lösung minimiert ein bestimmtes konvexes Funktional), dessen Lösungen gegen die Lösung der linearen Fokker-Planck-Gleichung konvergieren.

**Definition 2.3.** Für eine Wahrscheinlichkeitsdichte  $\rho \in P_{ac}(\mathbb{R}^d)$  und ein Potential  $V \in \mathcal{L}^1(\mathbb{R}^d)$  sei das **Freie Energie Funktional**  $H(\rho)$  gegeben durch:

$$H(\rho) := \int_{\mathbb{R}^d} \rho \log \rho dx + \int_{\mathbb{R}^d} \rho V dx \quad (2.3)$$

**Bemerkung 2.4.** • Das Freie Energie Funktional ist wohldefiniert für Lösungen  $\rho(x, t)$  der (LFP), falls  $H(\rho_0) < \infty$ .

•  $H(\rho(x, t))$  ist für Lösungen des (LFP) bzgl. der Zeit monoton abnehmend.

Nach der Definition des letzten Vortrages betrachten wir die Summe aus interner und potentieller Energie.

Wir wollen die Beziehung zwischen Fokker-Planck und dem freien Energie-Funktional zeigen, d.h. dass die Dynamik als Gradientenfluss interpretiert werden kann für die Frei Energie mit der Wassersteinmetrik, d.h. die Lösung der Gleichung folgt der Richtung des steilsten Abstieges des Funktionals. Die für dieses Setting benötigte Metrik wird die bereits eingeführte Wassersteinmetrik sein.

Otto:(FP) entspricht  $\frac{\partial \rho}{\partial t} = -\text{grad}_{W_2} H(\rho)$

Wiederholung:

**Definition 2.5.** Seien  $f, g \in \mathcal{P}(\mathbb{R}^d)$  zwei Wahrscheinlichkeitsmaße (Notation: bzw. ihre assoziierten Wahrscheinlichkeitsdichten.)

$$W_2(f, g) := \inf_{\pi \in \Pi(f, g)} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\pi(x, y) \right\}^{\frac{1}{2}}. \quad (2.4)$$

$\Pi(f, g)$ : Menge aller multivariaten Wahrscheinlichkeitsmaße auf  $\mathbb{R}^d \times \mathbb{R}^d$  mit Randverteilungen  $f$  und  $g$ .

Da die Wassersteinmetrik zusammen mit dem Raum aller Wahrscheinlichkeitsmaße aber keine Riemann'sche Mannigfaltigkeit ist, und damit der Gradient zunächst nicht wohldefiniert ist, müssen wir analog zum ersten Teil des Vortrags zeitlich diskretisieren.

**Theorem 2.6.** Sei  $V \in \mathcal{C}^2(\mathbb{R}^d)$  ein Potential, sodass  $V(x) = O(|x|^2)$  für  $x \rightarrow \infty$ . Sei  $\rho_0 \in P_{ac,2}(\mathbb{R}^d)$  derart, dass für das Freie Energie Funktional gilt  $H(\rho_0) < \infty$ . Sei  $\tau > 0$ .

Sei außerdem  $\bar{\rho}_\tau(x, t)$  der approximierte Gradientenfluss für  $H(\rho)$  bzgl. der quadratischen Wassersteinmetrik. Dann gilt:

$$\bar{\rho}_\tau(x, t) \rightarrow \rho(x, t) \text{ in } L^1(\mathbb{R}^d) \text{ für } \tau \rightarrow 0,$$

wobei  $\rho \in C^\infty(\mathbb{R}^d \times (0, \infty))$  die eindeutige Lösung der linearen Fokker-Planck-Gleichung mit Anfangswert  $\rho_0$  ist. D.h. der approximierte Gradientenfluss konvergiert schwach gegen die Lösung der Fokker-Planck-Gleichung.

Der approximierte Gradientenfluss  $\bar{\rho}_\tau : (0, \infty) \times \mathbb{R}^d \rightarrow [0, \infty)$  wird hierbei analog zu Abschnitt 1 konstruiert:

$$\bar{\rho}_\tau(t) := \rho_\tau^n \text{ für } t \in [n\tau, (n+1)\tau)$$

für die Folge  $(\rho_\tau^n)_n$ ,  $n \in \mathbb{N}_0$ :

$$\rho_\tau^0 := \rho^0; \quad (2.5)$$

$$\rho_\tau^{n+1} = \operatorname{argmin}_{\rho \in \mathcal{P}_2(\mathbb{R}^d)} \left\{ H(\rho) + \frac{W_2(\rho_\tau^n, \rho)^2}{2\tau} \right\} \quad (2.6)$$

Das heißt eben genau: Die Fokker-Planck-Gleichung ist der Gradientenfluss für das Freie-Energie Funktional mit der Wassersteindistanz.

*Beweis.* 1. Da  $H$  strikt konvex im üblichen Sinne ist, ebenso wie die Wassersteinmetrik, ist der Minimierer in (2.6) eindeutig. Auch möglich: nach dem letzten Vortrag ist für  $V$  konvex das Funktional displacement konvex.

2. Analog zu (1.12) erhält man eine obere Grenze für unser Funktional:  $\sup_{t \geq 0} H(\bar{\rho}_\tau(t)) = \sup_n H(\rho_\tau^n) \leq H(\rho_0)$ . Dies garantiert die Straffheit der Familie  $(\bar{\rho}_\tau)_{\tau > 0}$  in der  $x$ -Variable sowie schwache  $L^1$ -Kompaktheit. Die Hölder  $1/2$  Ungleichung (1.14) hinsichtlich der Wassersteindistanz sichert gleichgradige Stetigkeit bzgl.  $t$  von  $(\bar{\rho}_\tau)_\tau$ .

$$(W_2(\bar{\rho}_\tau(s) - \bar{\rho}_\tau(t)) < \varepsilon \forall |s - t| < \delta; \forall \varepsilon \text{ setze } \delta = \frac{\varepsilon^2}{2(E(\rho_0) - \inf E)} - \tau)$$

$\Rightarrow$  All dies ist hinreichend, damit (betrachte falls nötig Teilfolge)  $(\bar{\rho}_\tau)_\tau$  zu einer Funktion  $\rho : \mathbb{R}_{\geq 0} \mapsto P_{ac,2}(\mathbb{R}^d)$  konvergiert, stetig falls  $P_{ac,2}(\mathbb{R}^d)$  mit der schwachen  $L^1$ -Topologie verknüpft ist.

3. Erfüllt nun dieser Grenzwert den Satz, d.h. die (FK)?

Wir betrachten eine geeignete Variation von  $\rho_\tau^{n+1}$ . Ein wenig komplexer als in Abschnitt 1 (dort war es  $X_\tau^{n+1} + \varepsilon\omega$ ), um unserer Wassersteindistanz gerecht zu werden, jedoch mit dem gleichen

Ziel.

Sei  $\xi$  ein glattes Vektorfeld mit kompaktem Träger, und  $T_\varepsilon$  eine Familie von Kurven verknüpft mit dem Vektorfeld  $\text{Id} + \varepsilon\xi$ . Dann definiere folgende Perturbation:

$$\tilde{\rho}_\varepsilon = T_\varepsilon \# \rho_\tau^{n+1}$$

(Pushforward:  $(\tilde{\rho}_\varepsilon \mathcal{L}^d)(B) = (\rho_\tau^{n+1} \mathcal{L}^d)(T_\varepsilon^{-1}(B))$ )

Für  $\varepsilon$  klein genug ist  $T_\varepsilon$  ein  $C^1$ -Diffeomorphismus und invertierbar ( $\det(\nabla T_\varepsilon) > 0$ ).

4. Es gilt aufgrund der Invertierbarkeit von  $T_\varepsilon$  mithilfe von Substitution und McCann-Theorem (Vortrag Alex):

$$H(\tilde{\rho}_\varepsilon) = \int \tilde{\rho}_\varepsilon \log \tilde{\rho}_\varepsilon + \int \tilde{\rho}_\varepsilon V \quad (2.7)$$

$$= \int \rho_\tau^{n+1} \log \frac{\rho_\tau^{n+1}}{\det(I_k \varepsilon \nabla \xi)} + \int \rho_\tau^{n+1}(x) V(x + \varepsilon \xi(x)) dx. \quad (2.8)$$

5. Aus der Energy Estimate und der Definition des Minimierungsproblems folgt andererseits, dass  $\rho_\tau^n, \rho_\tau^{n+1}$  absolut stetig (bzgl. Lebesgue-maß) sind. Damit (Brenier, ebenfalls Vortrag Alex) existiert eine bzgl. der Wassersteinmetrik optimale Abbildung  $\nabla \phi$  sodass  $\nabla \phi \# \rho_\tau^n = \rho_\tau^{n+1}$ , also gilt

$$W_2(\rho_\tau^n, \rho_\tau^{n+1})^2 = \int \rho_\tau^n(x) |x - \nabla \phi(x)|^2 dx. \quad (2.9)$$

Dann können wir  $\tilde{\rho}_\varepsilon$  auch in Abhängigkeit von  $\rho_\tau^n$  statt  $\rho_\tau^{n+1}$  schreiben:  $\tilde{\rho}_\varepsilon = [(\text{Id} + \varepsilon \xi) \circ \nabla \phi] \# \rho_\tau^n$ . Wir verwenden nun nach der Definition der Wassersteinmetrik analog zum letzten Schritt die bestimmte Abbildung  $[(\text{Id} + \varepsilon \xi) \circ \nabla \phi]$  zwischen  $\tilde{\rho}_\varepsilon$  und  $\rho_\tau^n$ , diesmal haben wir allerdings keinen Anspruch darauf, dass dies tatsächlich die optimale wäre, und erlangen so folgende Ungleichung:

$$W_2(\rho_\tau^n, \tilde{\rho}_\varepsilon)^2 \leq \int |x - \underbrace{\nabla \phi(x) - \varepsilon \xi \circ \nabla \phi(x)}_{[(\text{Id} + \varepsilon \xi) \circ \nabla \phi](x)}|^2 \rho_\tau^n(x) dx$$

Damit kommen wir auf

$$\begin{aligned} 0 &\leq \frac{W_2(\rho_\tau^n, \tilde{\rho}_\varepsilon)^2}{2\tau} + H(\tilde{\rho}_\varepsilon) - \frac{W_2(\rho_\tau^n, \rho_\tau^{n+1})^2}{2\tau} - H(\rho_\tau^{n+1}) \\ &\stackrel{\log(\frac{a}{b}) = \log a - \log b}{\leq} \int \rho_\tau^n(x) \left( \frac{|x - \nabla \phi(x) - \varepsilon \xi \circ \nabla \phi(x)|^2 - |x - \nabla \phi(x)|^2}{2\tau} \right) dx \\ &\quad + \int \rho_\tau^{n+1}(x) [V(x + \varepsilon \xi(x)) - V(x)] dx - \int \rho_\tau^{n+1}(x) \log(\det(I_k + \varepsilon \nabla \xi(x))) dx \end{aligned} \quad (2.10)$$

wobei die Nichtnegativität aus der Minimierungseigenschaft von  $\rho_\tau^{n+1}$  folgt.

Nun dividieren wir durch  $\varepsilon$  und lassen dieses gegen  $0^+$  gehen, analog zu Abschnitt 1, Beweis des 1.Lemmas. Grenzwerte und Integrale dürfen wir vertauschen, da unsere Funktionen beschränkt bzw. Wahrscheinlichkeitsmaße endlichen zweiten Moments sind (deswegen auch die Wachstumsbedingungen!!) Die sich ergebenden “Differenzenquotienten“ formulieren wir als Ableitungen an

der Stelle  $\varepsilon = 0$ . Damit wird (2.10) zu:

$$\begin{aligned}
0 &\leq \lim_{\varepsilon \rightarrow 0^+} \left( \frac{1}{2\tau} \int \left( \frac{|x - \nabla\phi(x) - \varepsilon\xi \circ \nabla\phi(x)|^2 - |x - \nabla\phi(x)|^2}{\varepsilon} \right) \rho_\tau^n(x) dx \right. \\
&\quad \left. + \int \frac{1}{\varepsilon} [V(x + \varepsilon\xi(x)) - V(x)] \rho_\tau^{n+1}(x) dx - \int \frac{\log(\det(I_k + \varepsilon\nabla\xi(x))) - \log(\det I_k)}{\varepsilon} \rho_\tau^{n+1}(x) dx \right) \\
&= \frac{1}{2\tau} \int \frac{\partial}{\partial\varepsilon} |x - \nabla\phi - \varepsilon\xi \circ \nabla\phi(x)|^2 \Big|_{\varepsilon=0} \rho_\tau^n(x) dx + \int \frac{\partial}{\partial\varepsilon} [V(x + \varepsilon\xi(x))] \Big|_{\varepsilon=0} \rho_\tau^{n+1}(x) dx \\
&\quad - \int \frac{\partial}{\partial\varepsilon} [\log(\underbrace{\det(I_k + \varepsilon\nabla\xi(x))}_{=1 + \text{tr } \varepsilon\nabla\xi(x) + O(\varepsilon^2(\nabla\xi(x))^2}))] \Big|_{\varepsilon=0} \rho_\tau^{n+1}(x) dx \\
&= \frac{1}{\tau} \int \rho_\tau^n(x) \langle \nabla\phi(x) - x, \xi \circ \nabla\phi(x) \rangle dx + \int \rho_\tau^{n+1}(x) \langle \nabla V(x), \xi(x) \rangle dx - \int \rho_\tau^{n+1}(x) 1 \cdot \underbrace{(\text{tr } \nabla\xi)(x)}_{\nabla \cdot \xi} dx
\end{aligned}$$

Da dies genauso für  $-\xi$  gilt, bekommen wir Gleichheit. Umformuliert:

$$\frac{1}{\tau} \int \rho_\tau^n(x) \langle \nabla\phi(x) - x, \xi \circ \nabla\phi(x) \rangle dx = \int \rho_\tau^{n+1}(x) [(\nabla \cdot \xi)(x) - \langle \nabla V(x), \xi(x) \rangle] dx \quad (2.11)$$

6. Jetzt sei  $\xi$  derart, dass für ein  $\psi \in \mathcal{C}^\infty(\mathbb{R}^n)$  gilt  $\xi = \nabla\psi$ . Es gilt (Taylor):

$$\psi(\nabla\phi(x)) - \psi(x) = \langle \nabla\phi(x) - x, \nabla\psi \circ \nabla\phi(x) \rangle + O(|x - \nabla\phi(x)|^2) \quad (2.12)$$

und damit kann die linke Seite von (2.11) ersetzt werden durch

$$\begin{aligned}
&\frac{1}{\tau} \left( \int \rho_\tau^n(x) \psi \circ \nabla\phi(x) dx - \int \rho_\tau^n(x) \psi(x) dx \right) + O\left(\frac{1}{\tau} \int \rho_\tau^n(x) |x - \nabla\phi(x)|^2 dx\right) \\
&\quad \stackrel{\nabla\phi \# \rho_\tau^n = \rho_\tau^{n+1}}{=} \frac{1}{\tau} \left( \int \rho_\tau^{n+1} \psi - \int \rho_\tau^n \psi \right) + O\left(\frac{W_2(\rho_\tau^n, \rho_\tau^{n+1})^2}{\tau}\right)
\end{aligned}$$

7. Nun setzt man dies in (2.11) ein und es wird aufsummiert von  $n_1 = \lfloor t_1/\tau \rfloor$  bis zu  $n_2 = \lfloor t_2/\tau \rfloor - 1$  für  $t_1, t_2 \in [0, \infty)$  beliebige Zeiten (Grenzen: siehe Abschnitt 1, Beweis Corollar):

$$\frac{1}{\tau} \left( \int \bar{\rho}_\tau(t_2) \psi dx - \int \bar{\rho}_\tau(t_1) \psi dx + \underbrace{O\left(\sum_{n=n_1}^{n_2} W_2(\rho_\tau^n, \rho_\tau^{n+1})^2\right)}_{(*)} \right) \quad (2.13)$$

$$= \sum_{n_1}^{n_2} \int \rho_\tau^{n+1}(x) [(\nabla \cdot \nabla\psi)(x) - \langle \nabla V(x), \nabla\psi(x) \rangle] dx \quad (2.14)$$

$$= \frac{1}{\tau} \int_{t_1}^{t_2} \int \bar{\rho}_\tau(t) (\Delta\psi - \nabla V \cdot \nabla\psi) dx dt + O(\tau) \quad (2.15)$$

Wobei  $O(\tau)$  dem gerundeten Summationsgrenzen  $n_1, n_2$  geschuldet ist: Der Fehlerterm ist  $\leq \tau \int \bar{\rho}_\tau (\Delta\psi - \nabla V \cdot \nabla\psi) \leq \tau \|\bar{\rho}_\tau\|_1 \|\Delta\psi - \nabla V \cdot \nabla\psi\|_\infty = O(\tau)$  mithilfe der Hölderungleichung, wobei ja  $\Delta\psi$  und  $\nabla V \cdot \nabla\psi$  beschränkte Funktionen sind (kompakter Support) und  $\bar{\rho}_\tau(t) = \bar{\rho}_\tau(t, \cdot)$  Wahrscheinlichkeitsmaß auf  $\mathbb{R}^k$  ist. (ähnliche Argumente, um im Vorangegangenen Grenzwerte und Integrale vertauschen zu dürfen)

8. Die Kinetische Energie Abschätzung bzgl.  $W_2$  liefert uns

$$(*) = O(2\tau(H(\rho_0) - \inf H)) = O(\tau).$$

Damit gilt schließlich für alle festen  $t_1, t_2 \in [0, \infty)$ :

$$\int \bar{\rho}_\tau(t_2) \psi - \int \bar{\rho}_\tau(t_1) \psi + O(\tau) = \int_{t_1}^{t_2} \int \bar{\rho}_\tau(t) (\Delta\psi - \nabla V \cdot \nabla\psi) dt + O(\tau) \quad (2.16)$$



9. Grenzwert  $\tau \rightarrow 0$  liefert:

$$\int \rho(t_2)\psi - \int \rho(t_1)\psi = \int_{t_1}^{t_2} \int \rho(t)(\Delta\psi - \nabla V \cdot \nabla\psi)dt \quad (2.17)$$

Setze  $t_1 = 0$ : Der Grenzwert der Approximation erfüllt nach diesem Beweis die schwache Formulierung der linearen Fokker-Planck-Gleichung mit Potential  $V$ . Da die Lösung eindeutig ist, und bekanntermaßen eine klassische Lösung existiert, haben wir die (FP) gelöst.

□