Prof. Dr. Matthes
Optimaler Transport - Geometrie und Anwendungen

Manuela Lambacher
09.12.2015

Zeit-Diskretisierung von Gradientenflüssen

1 Generalisierte Gradientenflüsse

Wir betrachten folgenden Gradientenfluss:

$$\frac{dX}{dt} = -\operatorname{grad} E(X), \qquad X(0) = X^0 \tag{1.1}$$

$E : \mathbb{R}^k \to \mathbb{R}$ sei ein glattes, nach unten beschränktes Energie-Funktional, das außerdem $E(X^0) < \infty$ erfülle. $X(t) : [0, \infty) \to \mathbb{R}^k$ 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_τ^{n+1} ist ein Minimierer des Minimerungsproblems

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

Sei $\bar{X}_\tau : \mathbb{R}_{\geq 0} \to \mathbb{R}^k$ 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 approximierten 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 $\omega \in \mathbb{R}^k$:*

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

Beweis. Dabei wird eine Strategie eingeführt, der auch - in einem etwas umfangreicherem Setting - der nächste Abschnitt über die Fokker-Planck-Gleichung folgt.
Wir befinden uns im $\mathbb{R}^k$. Sei ω ein beliebiger (Tangenten-)Vektor auf X_τ^{n+1}. Wir betrachten Perturbationen von X_τ^{n+1} "entlang ω": Definiere für ein kleines $\varepsilon > 0$ einen "Pfad" $\tilde{X}_\varepsilon$ folgendermaßen:

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

Essentiell für die folgenden Schritte ist, dass die Ableitung an der Stelle ε 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

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

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

verwenden.)

Da X_τ^{n+1} Minimierer ist, gilt offensichtlich

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

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

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

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

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

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

Bevor wir $\tau \to 0$ 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_τ^{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_τ^{n+1} kleiner gleich dessen Wert an der Stelle X_τ^n ($\|X_\tau^n - X_\tau^n\| = 0$). Damit haben wir

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

Lemma 1.3 (Energy estimate)**.**

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

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. □

Lemma 1.4.

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

Interpretation als kinetische Energie!

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

Lemma 1.5 (Hölder 1/2-Abschätzung für X_τ).
Für $s < t$ gilt:

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

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

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

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

□

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 $\tau \to 0$ sicher, und damit die Existenz einer konvergenten Teilfolge.

Corollar 1.7.

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

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

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

Beweis. Wir beginnen mit (1.3) und summieren:

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

□

2 Die lineare Fokker-Plank-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-Plank-Gleichung angewandt.

Definition 2.1. *Die **lineare Fokker-Planck-Gleichung** (LFP) mit glattem Potential $V(x) : \mathbb{R}^k \to [0,\infty)$ und Anfangswert ρ_0, Wahrscheinlichkeitsdichte auf $\mathbb{R}^k$, ist gegeben durch:*

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

Alle Lösungen $\rho(x,t)$ sind ebenfalls Wahrscheinlichkeitsdichten auf $\mathbb{R}^k$ für fast alle festen t, d.h. $\rho(x,t) \geq 0$ für fast alle $(x,t) \in \mathbb{R}^k \times (0,\infty)$ und $\int_{\mathbb{R}^k} \rho(x,t)dx = 1$ für fast alle $t \in (0,\infty)$ Eine Wahrscheinlichkeitsdichte ρ erfüllt die (LFP) im schwachen Sinne, falls für alle Testfunktionen ζ gilt:

$$\int_{\mathbb{R}^k} \int_T (\rho\Delta\zeta + \rho\nabla V \cdot \nabla\zeta - \rho\zeta_t) dtdx = \int \rho_0 \zeta(x,0) dx \tag{2.2}$$

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).*

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 $\rho \in P(\mathbb{R}^k)$ und ein hinreichend glattes Potential V sei das **Freie Energie Funktional** $E(\rho)$ gegeben durch:*

$$E(\rho) := \int_{\mathbb{R}^k} \rho \log \rho dx + \int_{\mathbb{R}^k} \rho V dx \tag{2.3}$$

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

Wiederholung:

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

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

$\Pi(f,g)$: Menge aller multivariaten Wahrscheinlichkeitsmaße auf $\mathbb{R}^k \times \mathbb{R}^k$ 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 $V(x)$ ein glattes Potential, das folgende Wachstumsbedingung erfüllt:*
$V(x) = O(|x|^2)$ für $x \to \infty$.
Sei ρ_0 eine Wahrscheinlichkeitsdichte, sodass das Freie Energie Funktional $E(\rho)$ (2.3) an der Stelle ρ_0 endlich ist. Sei $\tau > 0$.
Dann gilt:
Der approximierte Gradientenfluss ρ_τ für $E(\rho)$ bzgl. der quadratischen Wassersteindistanz konvergiert für $\tau \to 0$ für alle $t \in (0, \infty)$ schwach in $L^1(\mathbb{R}^k)$ gegen die eindeutige Lösung $\rho(x,t) \in C^\infty((\mathbb{R}^k \times (0,\infty))$ der linearen Fokker-Planck-Gleichung (2.1) mit Anfangswert ρ_0.

Der approximierte Gradientenfluss (bzw. die Interpolation) $\rho_\tau : (0,\infty) \times \mathbb{R}^k \to [0,\infty)$ wird hierbei analog zu Def **??** *konstruiert:*

$$\rho_\tau(t) := \rho_\tau^n \quad \text{für } t \in [n\tau, (n+1)\tau) \tag{2.5}$$

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

$$\rho_\tau^0 = \rho^0;$$

Gegeben ρ_τ^n, definiere den nächsten Schritt als Minimierer von

$$\rho \mapsto E(\rho) + \frac{W_2(\rho_\tau^n, \rho)^2}{2\tau}, \tag{2.6}$$

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

Beweis.

1. Da E 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: $\sup_{t \geq 0} E(\rho_\tau(t)) = \sup_n E(\rho_\tau^n) \leq E(\rho_0)$ Dies garantiert die Straffheit der Familie $(\rho_\tau)_{\tau > 0}$ 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 $(\rho_\tau)_\tau$. ($W_2(\rho_\tau(s) - \rho_\tau(t)) < \varepsilon \forall |s-t| < \delta$; $\forall \varepsilon$ setze $\delta = \frac{\varepsilon^2}{2(E(\rho_0) - \inf E)} - \tau$)

   $\Rightarrow$ All dies ist hinreichend, damit(betrachte falls nötig Teilfolge)$(\rho_\tau)_\tau$ zu einer Funktion $\rho : \mathbb{R}_+ \mapsto P_{ac}(\mathbb{R}^k)$ konvergiert, stetig falls $P_{ac}(\mathbb{R}^k)$ mit der schwachen L^1-Topologie verknüpft ist.
3. Erfüllt nun dieser Grenzwert den Satz, d.h. die (LFK)?
   Wir betrachten eine geeignete Variation von ρ_τ^{n+1}. Ein wenig komplexer als in Abschnitt 1, um unserer Wassersteindistanz gerecht zu werden, jedoch mit dem gleichen Ziel.

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

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

(Pushforward: $(\tilde{\rho}_\varepsilon \mathcal{L}^k)(B) = (\rho_\tau^{n+1}\mathcal{L}^k)(T_\varepsilon^{-1}(B)))$
Für ε klein genug ist T_ε ein C^1-Diffeomorphismus) und invertierbar $(\det(\nabla T_\varepsilon) > 0$.

4. Es gilt aufgrund der Invertierbarkeit von T_ε mithilfe von Substitution und McCann-Theorem (Vortrag Alex):

$$\begin{aligned} E(\tilde{\rho}_\varepsilon) &= \int \tilde{\rho}_\varepsilon \log \tilde{\rho}_\varepsilon + \int \tilde{\rho}_\epsilon V && (2.7)\\ &= \int \rho_\tau^{n+1} \log \frac{\rho_\tau^{n+1}}{\det(I_k \varepsilon \nabla \xi)} + \int \rho_\tau^{n+1}(x) V(x + \varepsilon\xi(x))dx. && (2.8) \end{aligned}$$

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

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

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

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

Damit kommen wir auf

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

wobei die Nichtnegativität aus der Minimierungseigenschaft von ρ_τ^{n+1} folgt.
Nun dividieren wir durch ε und lassen dieses gegen 0^+ gehen, analog zu Abschnitt 1. Die sich ergebenden "Differenzenquotienten" formulieren wir als Ableitungen an der Stelle $\varepsilon = 0$. Damit wird (2.10) zu:

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

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

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

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

$$\zeta(\nabla\phi(x)) - \zeta(x) = \langle \nabla\phi(x) - x, \nabla\zeta\circ\nabla\phi(x)\rangle + O(|x - \nabla\phi(x)|^2 \tag{2.12}$$

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

$$\frac{1}{\tau}\left(\int \rho_\tau^n(x)\zeta\circ\nabla\phi(x)dx - \int \rho_\tau^n(x)\zeta(x)dx\right) + O\left(\frac{1}{\tau}\int \rho_\tau^n(x)|x-\nabla\phi(x)|^2 dx\right)$$
$$\overset{\nabla\phi\#\rho_\tau^n=\rho_\tau^{n+1}}{=} \frac{1}{\tau}\left(\int \rho_\tau^{n+1}\zeta - \int \rho_\tau^n\zeta\right) + O\left(\frac{W_2(\rho_\tau^n,\rho_\tau^{n+1})^2}{\tau}\right)$$

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 Beweis der Energy Gradient Estimate):

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

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

$$= \frac{1}{\tau}\int_{t_1}^{t_2}\int \rho_\tau(t)(\Delta\zeta - \nabla V\cdot\nabla\zeta)dxdt + O(\tau) \tag{2.15}$$

Wobei $O(\tau)$ dem gerundeten Summationsgrenzen n_1, n_2 geschuldet ist: Der Fehlerterm ist $\leq 2\tau\int\rho_\tau(\Delta\zeta - \nabla V\nabla\zeta) \leq 2\tau\|\rho_\tau\|_1\|\Delta\zeta - \nabla V\nabla\zeta\|_\infty = O(\tau)$ mithilfe der Hölderungleichung, wobei ja $\Delta\zeta$ und $\nabla V\cdot\nabla\zeta$ beschränkte Funktionen sind(kompakter Support) und $\rho_\tau(t) = \rho_\tau(t,\cdot)$ Wahrscheinlichkeitsmaß auf $\mathbb{R}^n$ ist. (ähnliche Argumente, um im Vorangegangenen Grenzwerte und Integrale vertauschen zu dürfen)

8. Die Total Square Estimate bzgl. W_2 liefert uns

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

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

$$\int \rho_\tau(t_2)\zeta - \int\rho_\tau(t_1)\zeta + O(\tau) = \int_{t_1}^{t_2}\int \rho_\tau(t)(\Delta\zeta - \nabla V\cdot\nabla\zeta)dt + O(\tau) \tag{2.16}$$

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

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

Setze $t_1 = 0$: Der Grenzwert der Approximation erfüllt nach diesem Beweis die schwache Formulierng der linearen Fokker-Planck-Gleichung mit Potential V. Da die Lösung eindeutig ist, haben wir die (LFP) gelöst.

□