

# Note "THE BACK-AND-FORTH METHOD FOR WASSERSTEIN GRADIENT FLOWS"

坂井幸人

2023 年 6 月 9 日

## 概要

ワッサーシュタイン勾配流を効率的に計算する方法を提案。アプローチは、最適輸送問題を解くために Jacobs と Léger が導入した往復法 (BFM) の一般化に基づいています [Numer. Math. 146 (2020) 513-544.]. JKO スキームの双対問題を解くことにより、勾配流を進化させる。一般的に、双対問題は原始問題よりも扱いやすい。これにより、特異な非凸エネルギーを含む多くの内部エネルギーに対して、大規模な勾配流シミュレーションを効率的に実行することができる。

## 1 INTRODUCTION

この研究では、以下のような形式の放物型方程式の進化シミュレーションに興味があります。

$$\begin{aligned}\partial_t \rho - \nabla \cdot (\delta \nabla \phi) &= 0, \\ \phi &= \delta U(\rho).\end{aligned}\tag{1}$$

方程式 (1) はしばしばダルシーの法則または一般化された多孔質媒体方程式と呼ばれ、内部エネルギー関数  $U$  によって生成された圧力勾配  $\nabla \phi$  に沿って流れる質量密度  $\rho$  の進化を記述します。このクラスの方程式は、流体流、熱伝導、拡散 (律速) 凝集、人流など、さまざまな物理現象をモデル化します。一般的に、**これらの方程式は剛性があり非線形であり、数値的に解くのは困難です**。例えば、

$$U(\rho) = \frac{1}{m-1} \int \rho^m (m > 1)$$

の重要な特殊な場合では、方程式 (1) は熱方程式の非線形バージョンである多孔質媒体方程式 (PME)

$$\partial_t \rho - \delta(\rho^m) = 0$$

となる。

$U$  が微分不可能または凸でない場合、これらの方程式のシミュレーションはさらに困難。したがって、この論文では、多種多様な内部エネルギー  $U$  の式 (1) を効率的かつ正確にシミュレートするための手法を設計することを目標としています。

私たちの Darcy の法則のシミュレーション手法は、方程式 (1) を Wasserstein 距離に関する勾配流として解釈するという優れたアプローチに基づいています [19, 25]。この解釈は、JKO スキームとして知られる離散

時間近似法を作成するために使用することができます [19]。このスキームは、次の反復によって近似解を構築します。

$$\min_{\rho \in \mathcal{P}} U(\rho) + \frac{1}{2\tau} W_2^2(\rho, \mu) \quad (2)$$

$$\rho^{(n+1)} := \arg \min_{\rho} U(\rho) + \frac{1}{2\tau} W_2^2(\rho, \rho^{(n)})$$

ここで、 $\tau$  はスキーム内の時間ステップを表し、 $W_2(\cdot, \cdot)$  は最適輸送理論の 2-Wasserstein 距離です [27] (最適輸送と 2-Wasserstein 距離の簡単な概要についてはセクション 2.1 を参照)。[スキームの変分的構造により](#)、反復解は無条件でエネルギー安定性を持ち、時間ステップ  $\tau$  を任意の空間離散化から独立して選択することができます。さらに、JKO スキームは連続方程式の比較型や収縮型の原理など、多くの望ましい特性を保持しています [1, 10, 20]。

JKO スキームの多くの有利な特性を考慮すると、問題 (2) の最小化問題の計算に多くの研究が注がれてきました。例えば、[2-4, 6-9, 22, 26] などが挙げられます。この問題に関する研究が多く行われているにも関わらず、高い解像度で JKO スキームを効率的に解くことは依然として課題です。[問題 \(2\) を解く上での主な困難は、Wasserstein 距離項の扱いです。実際、密度  \$\rho\$  に関する Wasserstein 距離の変動を与える簡単な公式は存在しません。そのため、\(2\) を解くためのほぼすべての方法は、2 つの固定された密度間の Wasserstein 距離を計算するアルゴリズムの適応である。](#)

この論文では、[21] で紹介された back-and-forth method (BFM) を適応して、問題 (2) を解決します。BFM は、2 つの固定された密度間の最適輸送写像を計算するための最先端のアルゴリズムです。[BFM は、モンジュの最適輸送問題を直接解くのではなく、関連するカントロビッチの双対問題を解くことによって最適写像を見つけます。このアプローチを基に、直接問題 \(2\) を解く代わりに、その双対問題の解を計算します。双対問題は凹最大化問題であり、次の時刻ステップの圧力変数  \$\phi^{\(n+1\)}\$  を生成します。最適密度変数は、圧力との双対関係  \$\phi^{\(n+1\)} = \delta U\(\rho^{\(n+1\)}\)\$  を介して簡単に回復することができます。](#)

双対問題を解くことによる利点はいくつかあります。圧力変数  $\phi$  は密度変数  $\rho$  よりも正則性が高いです。最悪の場合でも、圧力の勾配は[自乗可積分](#)である必要があります。その結果、圧力は離散的な近似スキームに適しています。さらに、双対汎関数の微分を計算するための明示的な式があるので、双対問題を解くために勾配上昇法を適用することができます (原始問題に対応する勾配降下法ははるかに困難です)。最後に、密度の非圧縮性など、 $U$  が厳しい制約を表現している場合 (例: 密度の非圧縮性)、双対問題は制約のない形で表現されるため、双対アプローチは非常に便利です。

[\(2\) の双対問題と BFM の特殊な勾配上昇構造を活用することで、内部エネルギー  \$U\$  の広範なクラスに対して JKO スキームを迅速かつ正確に解くことができます。我々は、アルゴリズムが各ステップで双対問題の値を増加させることを示しています。特に、この解析は  \$U\$  の Hessian が特異である場合や、計算グリッドのサイズに依存しない場合でも成立します。その結果、従来よりもはるかに大規模なスケールで式 \(1\) をシミュレートすることができ、障害物を持つ非圧縮性のある群衆モデルや集合拡散方程式 \(aggregation-diffusion equations\) のような難解なケースを簡単に扱うことができます。](#)

## 1.1 Overall approach

Wasserstein 勾配流における back-and-forth 法は、JKO スキームに関連する双対問題を解くことに基づいています。この分析の出発点は、Kantorovich の最適輸送の双対形式です。2つの測度  $\mu$  と  $\nu$  が与えられた場合、2-Wasserstein 距離の双対形式は次のようになります。

$$\frac{1}{2\tau} W_2^2(\mu, \nu) = \sup_{(\psi(x), \phi(y)) \in C} \left\{ \int \psi(x) d\mu(x) + \int \phi(y) d\nu(y) \right\}, \quad (3)$$

ここで、以下の制約条件を満たす範囲で最大化されます。

$$C := \{(\phi, \psi) \in C(\Omega) \times C(\Omega) : \psi(x) + \phi(y) \leq \frac{1}{2\tau} |x - y|^2\}.$$

最適輸送の双対形式を用いると、問題 (2) を次のように書き直すことができます。

$$\begin{aligned} \min_{\rho \in \mathcal{P}} U(\rho) + \frac{1}{2\tau} W_2^2(\rho, \mu) &= \min_{\rho \in \mathcal{P}} \left( U(\rho) + \sup_{(\varphi, \psi) \in C} \left( \int \varphi d\rho + \int \psi d\mu \right) \right) \\ &= \min_{\rho \in \mathcal{P}} \sup_{(\varphi, \psi) \in C} \left( U(\rho) + \int \varphi d\rho + \int \psi d\mu \right) \end{aligned}$$

$U$  が凸である場合、 $\inf$  と  $\sup$  を入れ替えることで、(2) と同等の双対問題を得ることができます。  
[YS](P2.detail)

$$\begin{aligned} \min_{\rho \in \mathcal{P}} U(\rho) + \frac{1}{2\tau} W_2^2(\rho, \mu) &= \sup_{(\varphi, \psi) \in C} \min_{\rho \in \mathcal{P}} \left( U(\rho) + \int \varphi d\rho + \int \psi d\mu \right), \\ &= \sup_{(\varphi, \psi) \in C} \left( \min_{\rho \in \mathcal{P}} \left( U(\rho) + \int \varphi d\rho \right) + \int \psi d\mu \right), \\ &= \sup_{(\varphi, \psi) \in C} \left( \int \psi d\mu - U^*(-\varphi) \right). \end{aligned}$$

$$\sup_{(\phi, \psi) \in C} \int \psi(x) d\rho^{(n)}(x) - U^*(-\phi), \quad (4)$$

ここで、 $U^*$  は  $U$  の (凸共役) を表し、次のように定義されます。

$$\begin{aligned} U^*(\phi) &:= \sup_{\rho \in \mathcal{P}} \left( \int \phi d\rho - U(\rho) \right) \\ U^*(-\phi) &= -\inf_{\rho} \left( \int \phi(y) d\rho(y) + U(\rho) \right). \end{aligned}$$

この双対問題を解くことで、問題 (2) の双対変数  $\phi$  を得ることができる。また、Legendre 変換 (凸共役) を用いて双対変数から密度変数を容易に復元することもできる。

問題 (4) は、 $\mathcal{C}$  によって表現される制約のために困難に見える。しかし、問題を再定式化する非常に便利な方法があります。 $\rho^{(n)}$  が非負測度であるため、できるだけ大きな値を持つように  $\psi$  を選ぶことが好ましい。

$\phi$  を固定すれば、対応する  $\psi$  の最大可能な選択肢は次のようになる。

$$\phi^c(x) := \inf_{y \in \Omega} \frac{1}{2\tau} |x - y|^2 - \phi(y). \quad (5)$$

また、 $\mu \geq 0$  より、 $\psi$  を固定すると、 $\phi$  の最大の選択肢は次のようになる。

$$\psi^c(y) := \inf_x \left( \frac{1}{2\tau} |x - y|^2 - \psi(x) \right) \quad (6)$$

式 (5) と (6) はそれぞれ backward-c-transform と forward-c-transform として知られています。これらの変換は最適輸送において重要な役割を果たし、私たちの手法には不可欠です。重要な点として、これらの変換を使用して制約  $\mathcal{C}$  と  $\phi$  または  $\psi$  のいずれかを排除することができます。

注意しておきますが、与えられた  $(\varphi, \psi) \in \mathcal{C}$  に対して、 $\psi^c \geq \varphi$  が成り立ちます。なぜなら、 $\mathcal{C} := \{(\phi, \psi) \in C(\Omega) \times C(\Omega) : \psi(x) + \phi(y) \leq \frac{1}{2\tau} |x - y|^2\}$  であり、 $\varphi(x) \leq \frac{1}{2\tau} |x - y|^2 - \psi(y)$  なので、右辺で  $\inf$  をとっているため、左辺の  $\varphi(x)$  の中で  $\sup$  が  $\psi^c$  になるためである。

また、 $\rho \geq 0$  の場合、 $-U^*(-\varphi)$  は  $\varphi$  に関して増加する関数です。よって、 $-U^*(\varphi) \leq -U^*(\psi)^c$  である。したがって、以下ようになります。

$$\sup_{(\varphi, \psi) \in \mathcal{C}} \left( \int \psi d\mu - U^*(-\varphi) \right) \leq \sup_{\psi} \left( \int \psi d\mu - U^*(-\psi^c) \right)$$

また、 $(\varphi, \psi) \in \mathcal{C} \implies (\psi^c, \psi) \in \mathcal{C}$  であるので、

$$\sup_{(\varphi, \psi) \in \mathcal{C}} \left( \int \psi d\mu - U^*(-\varphi) \right) \geq \sup_{\psi} \left( \int \psi d\mu - U^*(-\psi^c) \right)$$

が成立する。よって、

$$\sup_{(\varphi, \psi) \in \mathcal{C}} \left( \int \psi d\mu - U^*(-\varphi) \right) = \sup_{\psi} \left( \int \psi d\mu - U^*(-\psi^c) \right) \quad (7)$$

同様に、 $\mu \geq 0$  であるため、

$$\sup_{(\varphi, \psi) \in \mathcal{C}} \left( \int \psi d\mu - U^*(-\varphi) \right) = \sup_{\varphi} \left( \int \varphi^c d\mu - U^*(-\varphi) \right) \quad (8)$$

となります。

c-変換 (5)(6) を使用して制約  $\mathcal{C}$  と  $\phi$  または  $\psi$  のいずれかを排除することによって、問題 (4) は次の 2 つの制約のない汎関数の最大化として等価になる：

$$J(\phi) := \int_{\Omega} \phi^c(x) d\rho^{(n)}(x) - U^*(-\phi) \quad (9)$$

$$I(\psi) := \int_{\Omega} \psi(x) d\rho^{(n)}(x) - U^*(-\psi^c) \quad (10)$$

すなわち、

$$\sup_{(\phi, \psi) \in \mathcal{C}} \int \psi(x) d\rho^{(n)}(x) - U^*(-\phi) = \sup J(\phi) = \sup I(\psi).$$

加えて、もし  $\phi_*$  が  $J$  の最大化関数であり、 $\psi_*$  が  $I$  の最大化関数であるならば、

$$\phi_*^c = \psi_*, \quad \psi_*^c = \phi_*$$

の関係が成り立ち、 $(\phi_*, \psi_*)$  は (4) の最大化関数となります。 $I$  と  $J$  の再定式化は、最大化関数を見つける作業を実際に簡素化します。正規の離散グリッド上では、 $c$ -transform は非常に効率的に計算できます [21, 23]。その結果、(4) を直接扱うよりも、 $I$  と  $J$  を最大化する方がはるかに取り扱いやすくなります。

私たちは、[21] で紹介された BFM アルゴリズムを基にして、最大化関数  $\phi_*$  と  $\psi_*$  を見つける。元の BFM は、 $U^*$  が線形関数という特別な場合に最大化関数を効率的に見つけるための手法を提供している。 $I$  または  $J$  に焦点を当てるよりも早く、BFM は両方の関数を同時に最大化します。この手法は、 $\phi$ -空間での  $J$  の勾配上昇更新と  $\psi$ -空間での  $I$  の勾配上昇更新を交互に行うことで進行します（そのため、「back-and-forth」の名前があります）。勾配ステップの間には、一方の空間（ $\phi$ -空間または  $\psi$ -空間）の情報を他方に伝達するために、前方/後方  $c$ -transform を適用します。[21] で指摘されているように、back-and-forth アプローチの利点は、最適解のペア  $(\phi_*, \psi_*)$  の特定の特徴が、片方の空間よりも他方の空間でより簡単に構築できることです。その結果、back-and-forth 法は、 $\phi$ -空間のみまたは  $\psi$ -空間のみで操作する通常の勾配上昇法よりもはるかに迅速に収束します。

Wasserstein gradient flow の場合に BFM を一般化するためには、 $U^*$  が非線形の場合に (9) と (10) の勾配上昇ステップの安定性を保証する必要があります。実際には、多くの重要なケースでは、 $U^*$  の Hessian には特異成分が存在する可能性があります。この困難を克服するために、適切に重み付けされた Sobolev 空間で勾配上昇ステップを行います。Sobolev 制御により、境界積分を全空間上の積分に変換することができ、 $U^*$  の特異性を抑えることができます（詳細はセクション 3.2 を参照）。この連続解析の結果、離散化スキームは格子サイズに依存しない収束率を持つことになります。back-and-forth 法は、Algorithm 1 にまとめられています。ここで、 $H$  は前述の重み付き Sobolev 空間です。

一度双対問題を解決した後、元の問題 (2) の解を回復することができます。 $U$  が凸であれば、最適な双対変数  $\phi_*$  は  $\rho^{(n+1)}$  との双対関係

$$\rho^{(n+1)} = \delta U^*(\phi_*)$$

を介して関連付けられます（セクション 2.2 の定理 2.14 を参照）。 $U$  が凸でない場合、(2) と双対問題との間の関係はより不確かになります。幸いなことに、凸性分割スキームを使用することで、この困難を回避することができます [12]。実際に、 $U = U_1 + U_0$  と書けるようにすると、 $U_1$  は凸であり、 $U_0$  は凹であるとしします。その場合、JKO スキーム (2) を次の修正されたスキームに置き換えることができます。

$$\rho^{(n+1)} = \operatorname{argmin}_{\rho} U_1(\rho) + U_0(\rho^{(n)}) + (\delta U_0(\rho^{(n)}), \rho - \rho^{(n)}) + \frac{1}{2\tau} W_2^2(\rho, \rho^{(n)}) \quad (11)$$

凸性分割は、完全暗黙のスキームのエネルギー安定性を保持することがよく知られています。重要なのは、(11) のエネルギー項  $U_1(\rho) + U_0(\rho^{(n)}) + (\delta U_0(\rho^{(n)}), \rho - \rho^{(n)})$  は変数  $\rho$  に対して凸関数であるため、双対アプローチを適用できることです。すべてを考慮すると、私たちの方法は  $U$  が凸でない場合や不規則な場合でも、PDE (1) を非常に迅速にシミュレートする手法を提供します。

## 2 BACKGROUND

### 2.1 The $c$ -transform and optimal transport

ここから、 $\Omega$  上の連続関数の空間を  $C(\Omega)$  で表す。

**Definition 2.1.**  $\phi \in C(\Omega)$  の *backward  $c$ -transform* は以下のように表す:

$$\phi^c(x) := \inf_{y \in \Omega} \left( \frac{1}{2\tau} |x - y|^2 - \phi(y) \right). \quad (12)$$

また、 $\psi \in C(\Omega)$  の *forward  $c$ -transform* は以下のように表す:

$$\psi^c(y) := \inf_x \left( \frac{1}{2\tau} |x - y|^2 - \psi(x) \right). \quad (13)$$

**Definition 2.2.**  $\phi$  が  $c$ -concave であるとは、 $\phi = \psi^c$  となる  $\psi \in C(\Omega)$  が存在することである。また、 $(\phi, \psi) \in \mathcal{C}$  が  $c$ -conjugate とは、 $\phi = \psi^c$  かつ  $\psi = \phi^c$  であることをいう。

**Lemma 2.1.**  $\phi, \psi \in C(\Omega)$  のとき、 $\forall x \in \Omega$  に対し、 $\phi^{cc} \geq \phi$  が成り立つ。また、 $\phi^{cc} = \phi$  の必要十分条件は  $\phi$  が  $c$ -concave の時である。特に、 $\phi^{ccc} = \phi^c$  が成立する。[JL](Lemma 1(i))

*Proof.*

$$\begin{aligned} \phi^c(x) &= \inf_{y \in \Omega} \left( \frac{1}{2\tau} |x - y|^2 - \phi(y) \right) \\ \phi^{cc}(y) &= \inf_{x \in \Omega} \left( \frac{1}{2\tau} |x - y|^2 - \phi^c(x) \right) \\ &= \inf_{x \in \Omega} \left( \frac{1}{2\tau} |x - y|^2 - \inf_{z \in \Omega} \left( \frac{1}{2\tau} |x - z|^2 - \phi(z) \right) \right) \\ &\geq \inf_{x \in \Omega} \left( \frac{1}{2\tau} |x - y|^2 - \left( \frac{1}{2\tau} |x - y|^2 - \phi(y) \right) \right) \\ &= \phi(y). \end{aligned}$$

次に  $\phi^{cc} = \phi$  の必要十分条件は  $\phi$  が  $c$ -concave であることを示す。

( $\Leftarrow$ )

$\phi$  :  $c$ -concave すなわち、 $\exists \psi$  s.t.  $\phi = \psi^c$  のとき、 $\phi^{cc} = \phi$  を示す。

$\exists \psi$  s.t.  $\phi = \psi^c$  と仮定する。 $\psi := \phi^c$  とおくと、

$$\phi = \psi^c = (\phi^c)^c = \phi^{cc}$$

( $\Rightarrow$ )

$\phi^{cc} = \phi$  のとき、 $\phi : c\text{-concave}$  すなわち、 $\exists \psi \text{ s.t. } \phi = \psi^c$  を示す。

$\phi^{cc} = \phi$  と仮定すると、

$$\phi = \phi^{cc} = (\phi^c)^c$$

よって、 $\phi = (\phi^c)^c$  となる  $\phi^c$  が存在する。したがって、 $\phi$  は  $c\text{-concave}$  である。

最後に、 $\phi^{ccc} = \phi$  を示す。まず、 $\phi^{ccc} \geq \phi^c$  を示す。

$$\begin{aligned} \phi^{ccc}(x) &= \inf_{w \in \Omega} \left( \frac{1}{2\tau} |x - w|^2 - \phi^{cc}(w) \right) \\ &= \inf_{w \in \Omega} \left( \frac{1}{2\tau} |x - w|^2 - \inf_{z \in \Omega} \left( \frac{1}{2\tau} |w - z|^2 - \phi^c(z) \right) \right) \\ &= \inf_{w \in \Omega} \left( \frac{1}{2\tau} |x - w|^2 + \sup_{z \in \Omega} \left( -\frac{1}{2\tau} |w - z|^2 + \phi^c(z) \right) \right) \\ &\geq \inf_{w \in \Omega} \left( \frac{1}{2\tau} |x - w|^2 + \left( -\frac{1}{2\tau} |w - x|^2 + \phi^c(x) \right) \right) \\ &= \phi^c(x) \end{aligned}$$

次に、 $\phi^{ccc} \leq \phi^c$  を示す。

$$\begin{aligned} \phi^{ccc}(x) &= \inf_{w \in \Omega} \left( \frac{1}{2\tau} |x - w|^2 - \phi^{cc}(w) \right) \\ &\leq \inf_{w \in \Omega} \left( \frac{1}{2\tau} |x - w|^2 - \phi(w) \right) \\ &= \phi^c(x) \end{aligned}$$

よって、 $\phi^{ccc} = \phi^c$  が示された。

$\phi^{cc} = \phi$  の必要十分条件は  $\phi$  が  $c\text{-concave}$  であることを示す。(別解)

( $\Leftarrow$ )

$\phi : c\text{-concave}$  すなわち、 $\exists \psi \text{ s.t. } \phi = \psi^c$  のとき、 $\phi^{cc} = \phi$  を示す。

$\phi = \psi^c$  と仮定すると、

$$\phi^{cc} = (\psi^{cc})^c = \psi^{ccc} = \psi^c = \phi$$

( $\Rightarrow$ )  $\phi^{cc} = \phi$  のとき、 $\phi : c\text{-concave}$  すなわち、 $\exists \psi \text{ s.t. } \phi = \psi^c$  を示す。

$\phi^{cc} = \phi$  と仮定すると、

$$\phi = \phi^{cc} = (\phi^c)^c$$

よって、 $\phi = (\phi^c)^c$  となる  $\phi^c$  が存在する。したがって、 $\phi$  は  $c\text{-concave}$  である。  $\square$

**Proposition 2.1.**  $\phi \in C(\Omega)$  が  $c\text{-concave}$  のとき、以下の写像は *well-defined* かつ *almost everywhere* で *unique* である。

$$T_\phi(x) := \operatorname{argmin}_{y \in \Omega} \left( \frac{1}{2\tau} |x - y|^2 - \phi(y) \right) \quad (14)$$

すなわち、 $\phi^c$  はただ一つの最少点を持ち、 $\phi^c$  が最小値を取る座標は *transport map*  $T_\phi(x)$  で表される。さらに、 $u \in C(\Omega)$  であるとき、ほとんど全ての (*almost every*)  $x, y \in \Omega$  に対し、以下のような  $c\text{-transform}$  の撰

動公式が成り立つ。

$$\lim_{\varepsilon \rightarrow 0} \frac{(\phi + \varepsilon u)^c(x) - \phi^c(x)}{\varepsilon} = -u(T_\phi(x)) \quad (15)$$

最後に、以下も成り立つ。

$$T_\phi(x) = x - \tau \nabla \phi^c(x),$$

$$T_\psi(y) = y - \tau \nabla \psi^c(y),$$

また、 $T_\phi(T_\psi(y)) = y$ ,  $T_\psi(T_\phi(x)) = x$  が *almost everywhere* で成り立つ。

*Proof.* まず、写像  $T_\phi(x)$  が well-defined であることを示します。つまり、任意の  $x \in \Omega$  に対して  $\operatorname{argmin}_{y \in \Omega} (\frac{1}{2\tau}|x - y|^2 - \phi(y))$  が存在することを示す。関数  $\frac{1}{2\tau}|x - y|^2 - \phi(y)$  は  $\Omega$  上の連続関数であり、 $\Omega$  がコンパクト (閉集合  $\rightarrow$  コンパクト) であることから、この関数は最小値を持つことが保証される。(コンパクト空間の連続関数は  $\max$  と  $\min$  を持つ。) したがって、 $\operatorname{argmin}_{y \in \Omega} (\frac{1}{2\tau}|x - y|^2 - \phi(y))$  は存在する。

次に、 $T_\phi(x)$  が *almost everywhere* で一意であることを示す。つまり、ほとんど全ての  $x \in \Omega$  に対して、 $\operatorname{argmin}_{y \in \Omega} (\frac{1}{2\tau}|x - y|^2 - \phi(y))$  がただ一つの要素を持つことを示す。 $\phi$  が  $c$ -concave であることから、 $\operatorname{argmin}_{y \in \Omega} (\frac{1}{2\tau}|x - y|^2 - \phi(y))$  は凸集合となります。(Proof) また、凸集合上の任意の点は最小値を持つ点として一意に定まります。したがって、ほとんど全ての  $x \in \Omega$  に対して  $\operatorname{argmin}_{y \in \Omega} (\frac{1}{2\tau}|x - y|^2 - \phi(y))$  がただ一つの要素を持つことが示されます。

さらに、 $u \in C(\Omega)$  の場合、ほとんど全ての  $x, y \in \Omega$  に対して以下の摂動公式が成り立つことを示します。

$$\lim_{\varepsilon \rightarrow 0} \frac{(\phi + \varepsilon u)^c(x) - \phi^c(x)}{\varepsilon} = -u(T_\phi(x))$$

最後に、 $T_\phi(x) = x - \tau \nabla \phi(x)$  が成り立つことを示します。

関数  $\frac{1}{2\tau}|x - y|^2 - \phi(y)$  の最小値を実現する  $y$  は、 $\frac{\partial}{\partial y} (\frac{1}{2\tau}|x - y|^2 - \phi(y)) = 0$  で求められる。なぜなら、 $T_\phi$  は凸関数であるためである。まず、 $\nabla (\frac{1}{2\tau}|x - y|^2 - \phi(y))$  を求める。

$$\begin{aligned} \nabla (\frac{1}{2\tau}|x - y|^2 - \phi(y)) &= \frac{\partial}{\partial x} (\frac{1}{2\tau}|x - y|^2 - \phi(y)) + \frac{\partial}{\partial y} (\frac{1}{2\tau}|x - y|^2 - \phi(y)) \\ &= (x - y) + (-(x - y) - \nabla \phi(y)) \\ &= -\nabla \phi(y) \end{aligned}$$

$T_\phi$  の定義から、 $y = T_\phi$  のとき、 $\frac{1}{2\tau}|x - y|^2 - \phi(y)$  が  $\inf$  をとる。よって、 $\phi^c(y) = \inf_{y \in \Omega} (\frac{1}{2\tau}|x - y|^2 - \phi(y))$  と  $\phi(y)$  の関係は以下のとおりである。

$$\nabla \phi^c(x) = -\nabla \phi(T_\phi(x))$$

よって、

$$\begin{aligned} &\frac{\partial}{\partial y} (\frac{1}{2\tau}|x - y|^2 - \phi(y)) = 0 \\ \iff &-(x - y) - \nabla \phi(y) = 0 \\ \iff &y = x + \nabla \phi(y) \\ \iff &T_\phi(x) = x + \nabla \phi(T_\phi(x)) \\ \iff &T_\phi(x) = x - \nabla \phi^c(x) \end{aligned}$$



$$\begin{aligned} T_\psi(T_\phi(x)) &= T_\psi(x - \nabla\phi^c(x)) \\ &= y - \tau\nabla\psi(y) \end{aligned}$$

□

$T_\phi(T_\psi(y)) = y$  を証明します。

$T_\psi(y) = \operatorname{argmin}_{z \in \Omega} \left( \frac{1}{2\tau} |y - z|^2 - \psi(z) \right)$  と定義されています。

$T_\phi(T_\psi(y))$  を考えると、これは以下の最小化問題の解です：

$$T_\phi(T_\psi(y)) = \operatorname{argmin}_{x \in \Omega} \left( \frac{1}{2\tau} |T_\psi(y) - x|^2 - \phi(x) \right)$$

この最小化問題の解  $x$  について考えましょう。最適解  $x = T_\phi(T_\psi(y))$  と仮定します。

最適解  $x = T_\phi(T_\psi(y))$  に対して、以下の条件が成り立ちます：

$$\frac{1}{\tau} (T_\psi(y) - x) - \nabla\phi(x) = 0$$

この式を変形すると、

$$T_\psi(y) - x = \tau\nabla\phi(x)$$

両辺に  $-1$  をかけて整理すると、

$$x - T_\psi(y) = -\tau\nabla\phi(x)$$

ここで、 $\nabla\phi(x) = -\nabla\phi^c(x)$  を利用します（先程の議論で示しました）。

したがって、上記の式は次のように書き換えられます：

$$x - T_\psi(y) = \tau\nabla\phi^c(x)$$

両辺に  $-\tau$  をかけて整理すると、

$$T_\psi(y) - x = -\tau\nabla\phi^c(x)$$

これは  $T_\psi(y)$  と  $x$  の関係を表しています。右辺は  $c$ -transform の勾配であり、左辺は  $T_\psi(y)$  と  $x$  の差を表しています。

よって、 $T_\phi(T_\psi(y)) = y$  が成り立ちます。

同様に、 $T_\psi(T_\phi(x)) = x$  も証明することができます。

**Proposition 2.2.** もし、 $\mu$  と  $\nu$  が  $\Omega$  上の非負密度関数で、質量が等しい場合、つまり、

$$\int_{\Omega} \mu(x) dx = \int_{\Omega} \nu(y) dy$$

であるならば、次の式が成り立ちます。

$$\frac{1}{2\tau} W_2^2(\mu, \nu) = \sup_{\phi \in C(\Omega)} \left( \int_{\Omega} \phi^c(x) \mu(x) dx + \int_{\Omega} \phi(y) \nu(y) dy \right)$$

$$\frac{1}{2\tau} W_2^2(\mu, \nu) = \sup_{\psi \in C(\Omega)} \left( \int_{\Omega} \psi(x) \mu(x) dx + \int_{\Omega} \psi^c(y) \nu(y) dy \right)$$

この結果により、最適輸送写像の存在と一意性が保証されます。

*Proof.* 注意しておきますが、与えられた  $(\varphi, \psi) \in \mathcal{C}$  に対して、 $\phi^c \geq \psi$  が成り立ちます。ただし、

$$\phi^c(x) := \inf_y \left( \frac{1}{2\tau} |x - y|^2 - \phi(y) \right)$$

は  $\phi$  の  $c$ -変換 ( $c$ -transform) です。なぜなら、 $\mathcal{C} := \{(\phi, \psi) \in C(\Omega) \times C(\Omega) : \psi(x) + \phi(y) \leq \frac{1}{2\tau} |x - y|^2\}$ . であり、 $\psi(x) \leq \frac{1}{2\tau} |x - y|^2 - \phi(y)$  したがって、以下のようになります。

$$\begin{aligned} \frac{1}{2\tau} W_2^2(\mu, \nu) &= \sup_{(\psi(x), \phi(y)) \in \mathcal{C}} \left( \int \psi(x) d\mu(x) + \int \phi(y) d\nu(y) \right), \\ &\leq \sup_{\phi \in C(\Omega)} \left( \int_{\Omega} \phi^c(x) \mu(x) dx + \int_{\Omega} \phi(y) \nu(y) dy \right) \end{aligned}$$

また、 $(\varphi, \psi) \in \mathcal{C} \implies (\psi^c, \psi) \in \mathcal{C}$  であるので、

$$\sup_{(\psi(x), \phi(y)) \in \mathcal{C}} \left( \int \psi(x) d\mu(x) + \int \phi(y) d\nu(y) \right) \geq \sup_{\phi \in C(\Omega)} \left( \int_{\Omega} \phi^c(x) \mu(x) dx + \int_{\Omega} \phi(y) \nu(y) dy \right)$$

が成立する。よって、

$$\sup_{(\psi(x), \phi(y)) \in \mathcal{C}} \left( \int \psi(x) d\mu(x) + \int \phi(y) d\nu(y) \right) = \sup_{\phi \in C(\Omega)} \left( \int_{\Omega} \phi^c(x) \mu(x) dx + \int_{\Omega} \phi(y) \nu(y) dy \right)$$

□

**Theorem 2.1.** もし、 $\mu$  と  $\nu$  が  $\Omega$  上の非負密度関数で、質量が等しい場合、つまり、

$$\int_{\Omega} \mu(x) dx = \int_{\Omega} \nu(y) dy$$

であるならば、次の条件を満たす  $c$  共役のペア  $(\phi_*, \psi_*)$  が存在する:

$$\begin{aligned} \phi_* &\in \operatorname{argmax}_{\phi \in C(\Omega)} \left\{ \int_{\Omega} \phi^c(x) \mu(x) dx - \int_{\Omega} \phi(y) \nu(y) dy \right\} \\ \psi_* &\in \operatorname{argmax}_{\psi \in C(\Omega)} \left\{ \int_{\Omega} \psi(x) \mu(x) dx - \int_{\Omega} \psi^c(y) \nu(y) dy \right\} \end{aligned}$$

さらに、 $T_{\phi_*}$  は  $\mu$  を  $\nu$  に送る唯一の最適輸送写像であり、 $T_{\psi_*}$  は  $\nu$  を  $\mu$  に送る唯一の最適輸送写像です。つまり、 $T_{\phi_*} \mu = \nu$  および  $T_{\psi_*} \nu = \mu$  が成り立ちます。

また、2-Wasserstein 距離  $W_2^2(\mu, \nu)$  と関数  $\phi_*, \psi_*$  との関係は次のようになります。

$$\frac{1}{2\tau} W_2^2(\mu, \nu) = \int_{\Omega} \psi_*(x) \mu(x) dx - \int_{\Omega} \phi_*(y) \nu(y) dy.$$

## 2.2 Convex duality

一般化最適輸送 (GOT) 問題:

$$\rho_* = \operatorname{argmax}_{\rho \in L^1(\Omega)} \left( U(\rho) + \frac{1}{2\tau} W_2^2(\rho, \mu) \right), \quad (16)$$

ここで、 $\mu \in L^1(\Omega)$  は与えられた非負密度です。

**Assumption 2.1.** 内部エネルギー  $U$  は、proper で、凸かつ、下半連続な汎関数  $U : L^1(\Omega) \rightarrow \mathbb{R} \cup \{+\infty\}$  によって与えられます。ただし、 $\rho$  が正の測度を持つ集合上で負になる場合は  $U(\rho) = \infty$  とします。

**Assumption 2.2.** 超線形の成長 (superlinear growth) を持つ  $s : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$  が存在し、以下の条件を満たす:

$$U(\rho) \geq \int_{\Omega} s(\rho(y)) dy$$

**Remark 2.1.** 仮定 2.1 は、密度  $\rho$  が非負である必要があることを表しています。つまり、 $\rho$  は負の値を取ることはできません。これにより、非負の密度に対して関数  $U(\rho)$  が適切に定義されることが保証されます。一方、仮定 2.2 は、任意の  $B \in \mathbb{R}$  に対して、集合  $\{\rho \in L^1(\Omega) : U(\rho) < B\}$  が弱収束位相において弱コンパクトであることを保証します。

**Remark 2.2.** 凸性の要件を除けば、仮定 2.1 と仮定 2.2 は Wasserstein 勾配流の文脈では非常に自然なものです。ただし、セクション 3.3 では非凸な  $U$  も考慮します。

**Definition 2.3.** 関数  $U : L^1(\Omega) \rightarrow \mathbb{R}$  の凸共役  $U^* : L^\infty(\Omega) \rightarrow \mathbb{R}$  は次のように定義されます:

$$\begin{aligned} U^*(\phi) &:= \sup_{\rho \in L^1(\Omega)} \left\{ \int_{\Omega} \phi(y) \rho(y) dy - U(\rho) \right\}, \\ &= \sup_{\rho \in L^1(\Omega)} \left\{ \int_{\Omega} \phi d\rho - U(\rho) \right\}. \end{aligned}$$

仮定 2.1 のおかげで、 $U^*$  は重要な単調性を持ちます。

**Lemma 2.2.**  $U^*$  は単調増加である、つまり、 $\phi_0, \phi_1 : \Omega \rightarrow \mathbb{R}$  が  $\phi_0 \leq \phi_1$  となるすべての点で成り立つ場合、次の不等式が成り立ちます。

$$U^*(\phi_0) \leq U^*(\phi_1).$$

*Proof.* 仮定 2.1 により、内部エネルギーは非負の密度に対してのみ有限です。したがって、

$$U^*(\phi) = \sup_{\rho \geq 0} \left\{ \int_{\Omega} \phi(y) \rho(y) dy - U(\rho) \right\}.$$

ある  $\rho \in L^1(\Omega)$  を取る、ただし  $\rho(y) \geq 0$  a.e. であるとしします。すると、次の不等式が成り立ちます。

$$\int_{\Omega} \phi_0(y) \rho(y) dy - U(\rho) \leq \int_{\Omega} \phi_1(y) \rho(y) dy - U(\rho).$$

$\rho \geq 0$  に対して上記の不等式の上限を取ることで、証明が完了する。  $\square$

**Proposition 2.3.** 与えられた非負密度  $\mu \in L^1(\Omega)$  に対して、汎関数  $I$  と  $J$  は以下のように定義される (9), (10):

$$J(\phi) := \int_{\Omega} \phi^c(x) d\rho^{(n)}(x) - U^*(-\phi) \quad (17)$$

$$I(\psi) := \int_{\Omega} \psi(x) d\rho^{(n)}(x) - U^*(-\psi^c) \quad (18)$$

これらの汎関数は *proper* であり、弱上半連続であり、凹であり、さらに  $\sup_{\varphi \in C(\Omega)} J(\varphi) = \sup_{\psi \in C(\Omega)} I(\psi)$  を満たします。さらに、 $\varphi, \psi$  が  $c$ -凹である場合、 $J$  と  $I$  は以下のような一次変分を持ちます:

$$\delta J(\phi) = \delta U^*(-\phi) - T_{\phi\#}\mu,$$

$$\delta I(\psi) = \mu - T_{\psi\#}\delta U^*(-\psi^c).$$

*Proof.* Lemma 2.2 の証明の論理に従って、次のように表現することができます。

$$U^*(\psi^c) = \sup_{\rho \geq 0} \left( \int_{\Omega} \psi^c(y) \rho(y) dy - U(\rho) \right).$$

$$U^*(-\psi^c) = \sup_{\rho \geq 0} \left( \int_{\Omega} -\psi^c(y) \rho(y) dy - U(\rho) \right).$$

次に、 $\mathcal{M}(\Omega \times \Omega)$  を  $\Omega \times \Omega$  上の非負測度の集合とし、与えられた密度  $\rho \geq 0$  に対して次のように定義します。

$$\Pi(\rho) = \left\{ \pi \in \mathcal{M}(\Omega \times \Omega) : \iint_{\Omega \times \Omega} f(y) d\pi(x, y) = \int_{\Omega} f(y) \rho(y) dy, \forall f \in C(\Omega) \right\}.$$

c-変換の定義を用いると、式は次のように書けます。

$$\begin{aligned} \int_{\Omega} -\psi^c(y) \rho(y) dy &= - \int_{\Omega} \inf_{x \in \Omega} \left( \frac{1}{2\tau} |x - y|^2 - \psi(x) \right) \rho(y) dy, \\ &= \int_{\Omega} \sup_{x \in \Omega} \left( \psi(x) - \frac{1}{2\tau} |x - y|^2 \right) \rho(y) dy, \\ &= \sup_{\pi \in \Pi(\rho)} \iint_{\Omega \times \Omega} \left( \psi(x) - \frac{1}{2\tau} |x - y|^2 \right) d\pi(x, y). \end{aligned}$$

したがって、以下ようになります。

$$\begin{aligned} -U^*(-\psi^c) &= - \sup_{\rho \geq 0} \left( \int_{\Omega} -\psi^c(y) \rho(y) dy - U(\rho) \right), \\ &= - \sup_{\rho \geq 0} \left( \sup_{\pi \in \Pi(\rho)} \left( \iint_{\Omega \times \Omega} \left( \psi(x) - \frac{1}{2\tau} |x - y|^2 \right) d\pi(x, y) \right) - U(\rho) \right), \\ &= \inf_{\rho \geq 0} \left( - \sup_{\pi \in \Pi(\rho)} \left( \iint_{\Omega \times \Omega} \left( \psi(x) - \frac{1}{2\tau} |x - y|^2 \right) d\pi(x, y) \right) + U(\rho) \right), \\ &= \inf_{\rho \geq 0} \left( \inf_{\pi \in \Pi(\rho)} \left( - \iint_{\Omega \times \Omega} \left( \psi(x) - \frac{1}{2\tau} |x - y|^2 \right) d\pi(x, y) \right) + U(\rho) \right), \\ &= \inf_{\rho \geq 0} \inf_{\pi \in \Pi(\rho)} U(\rho) - \iint_{\Omega \times \Omega} \left( \psi(x) - \frac{1}{2\tau} |x - y|^2 \right) d\pi(x, y) \end{aligned}$$

これにより、 $I$  は  $\psi$  に関する一連の線形汎関数の下限として表現できることが明らかになりました。したがって、 $I$  は proper であり、凹であり、弱上半連続です。同様の議論が  $J$  にも適用されます。 $U^*$  が単調増加であることから、Lemma 2.2 は任意の  $\varphi, \psi \in C(\Omega)$  に対して次の不等式が成り立つことを意味します。

$$J(\varphi) \leq I(\varphi^c), \quad I(\psi) \leq J(\psi^c)$$

したがって、次の関係が成り立ちます。

$$\sup I(\psi) = \sup J(\varphi).$$

$\varphi$  と  $\psi$  がそれぞれ  $c$ -凸性/凹性を持つ場合、一次変分の式は Proposition 2.1 から直接導かれます。  $\square$

この subsection の最後に、 $I$  と  $J$  の最大化値から (16) の解を復元する方法を以下に示す。

**Theorem 2.2.**  $\mu \in L^1(\Omega)$  であり、 $U$  が Assumptions 2.1 と 2.2 を満たし、さらに  $\delta U(\mu)$  が定数関数でない場合、次の条件を満たす一意の密度  $\rho_*$  と  $c$ -共役な関数のペア  $(\phi_*, \psi_*)$  が存在します。

$$\rho_* = \operatorname{argmin}_{\rho \in L^1(\Omega)} U(\rho) + \frac{1}{2\tau} W_2(\rho, \mu), \quad \phi_* \in \operatorname{argmax}_{\phi \in C(\Omega)} J(\phi), \quad \psi_* \in \operatorname{argmax}_{\psi \in C(\Omega)} I(\psi),$$

$$U(\rho_*) + \frac{1}{2\tau} W_2^2(\rho_*, \mu) = J(\phi_*) = I(\psi_*),$$

$$\rho_* \in \delta U^*(\phi_*), \quad \phi_* \in \delta U(\rho_*), \quad \rho_* = T_{\phi_* \#} \mu.$$

**Remark 2.3.** 注意すべきは、もし  $\delta U(\mu)$  が定数関数である場合、 $\mu = \operatorname{argmin}_{\rho \in L^1(\Omega)} U(\rho) + \frac{1}{2\tau} W_2^2(\rho, \mu)$  となります。したがって、排除されたケースは自明な場合です。

## 2.3 Concave gradient ascent

JKO スキームを  $I$  と  $J$  の双対関数に関連付ける方法を見てきたので、 $I$  と  $J$  の最大化方法を開発する必要があります。そのために、このサブセクションでは、古典的な制約のない勾配上昇法について説明します。まず、勾配の概念を思い出す。これには、実ヒルベルト空間  $\mathcal{H}$  の内積  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  とノルム  $\|\cdot\|_{\mathcal{H}}$  の構造が必要です。

**Definition 2.4.** 点  $\varphi \in \mathcal{H}$  に対して、有界線型写像  $\delta F(\varphi) : \mathcal{H} \rightarrow \mathbb{R}$  が  $F$  の第 1 変分（フレシェ微分）であるとは、

$$\lim_{\|h\|_{\mathcal{H}} \rightarrow 0} \frac{\|F(\varphi + h) - F(\varphi) - \delta F(\varphi)(h)\|_{\mathcal{H}}}{\|h\|_{\mathcal{H}}} = 0$$

が成り立つことである。

**Definition 2.5.** 写像  $\nabla_{\mathcal{H}} F : \mathcal{H} \rightarrow \mathcal{H}$  が  $\mathcal{H}$ -勾配（または単に勾配、 $\mathcal{H}$  についての曖昧さがない場合）とは、

$$\langle \nabla_{\mathcal{H}} F(\varphi), h \rangle_{\mathcal{H}} = \delta F(\varphi)(h)$$

をすべての  $(\varphi, h) \in \mathcal{H} \times \mathcal{H}$  に対して満たすものを指す。

上記の等式は、勾配がヒルベルト空間の内積と密接に関連していることを強調しています。一方、第 1 変分の概念は、任意のノルム付きベクトル空間上で定義することができますが、勾配の概念は内積を必要とします。

### 2.3.1 Gradient ascent method

与えられた  $\mathcal{H}$  上の凹関数  $J$  に対し、勾配上昇法の反復式

$$\phi_{k+1} = \phi_k + \nabla_{\mathcal{H}} J(\phi_k). \tag{19}$$

を考える、勾配上昇法（式 (19)）は、次の変分形式でも同等に書くことができます。

$$\phi_{k+1} = \underset{\phi}{\operatorname{argmax}} J(\phi_k) + \delta J(\phi_k)(\phi - \phi_k) - \frac{1}{2} \|\phi - \phi_k\|_H^2. \quad (20)$$

ここで、式 (19) と (20) には通常、勾配方向でどれだけ進むかを制御するステップサイズパラメータが含まれています。後で明らかになる理由（式 (2.11) とその後の議論を参照）から、パラメータをノルム  $\|\cdot\|_H$  自体に組み込むほうがよいことがわかる。

スキームの収束性を得るために、効率的な収束率で

$$J(\phi_k) \xrightarrow{k \rightarrow \infty} \sup_{\phi} J(\phi),$$

を達成するためには、適切なノルム  $\|\cdot\|_H$  を選ぶことが重要です。ノルムが弱すぎると、アルゴリズムは不安定になり収束しない可能性があります。一方、ノルムが強すぎると、各ステップでほとんど変化が起らず、アルゴリズムは収束が遅くなります。最適化の基盤の 1 つである以下の定理は、これらの競合する考慮事項をバランス良く取る方法を説明しています。

**Theorem 2.3.**  $J: \mathcal{H} \rightarrow \mathbb{R}$  を 2 階フレシェ微分可能な凹関数であり、最大化子  $\phi^*$  を持つとする。以下の条件がすべての  $\phi, h \in \mathcal{H}$  に対して成り立つとき、 $J$  は「1-smooth」と呼ばれる：

$$-\delta^2 J(\phi)(h, h) \leq \|h\|_{\mathcal{H}}^2. \quad (21)$$

$J$  が 1-smooth のとき、勾配上昇法

$$\phi_{k+1} = \phi_k + \nabla_{\mathcal{H}} J(\phi_k)$$

を初期点  $\phi_0$  から開始すると、次の上昇特性 (22) を満たし、また収束率 (23) を持つ：

$$J(\phi_{k+1}) \geq J(\phi_k) + \frac{1}{2} \|\nabla_{\mathcal{H}} J(\phi_k)\|_{\mathcal{H}}^2, \quad (22)$$

$$J(\phi^*) - J(\phi_k) \leq \frac{\|\phi^* - \phi_0\|_{\mathcal{H}}^2}{2k}. \quad (23)$$

定理 2.3 から、ノルム  $\|\cdot\|_H$  を弱めるか強めるかという競合する要素が再び現れます。より強いノルムは式 (21) を満たしやすくなりますが、より弱いノルムは収束率 (23) を改善します。これらの考慮事項をまとめると、式 (21) を満たす限りできるだけ弱いノルムを選ぶのが最適であることが分かります。

### 2.3.2 Sobolev norm

$\Omega$  を  $\mathbb{R}^d$  の開有界凸部分集合とする。勾配上昇法は、Sobolev 空間  $H^1(\Omega)$  に基づいたノルム  $H$  を使用します。  $\Theta_1 > 0$  and  $\Theta_2 > 0$  を定数として、次のようにノルムを定義します。

$$\|h\|_H^2 = \int_{\Omega} \Theta_2 |\nabla h(x)|^2 + \Theta_1 |h(x)|^2 dx. \quad (24)$$

$\Theta_1$  や  $\Theta_2$  の具体的な値は、最大化される関数によって異なる（例えば、セクション 3 の定理 3.3 を参照してください）。多くの場合、 $\Theta_1$  と  $\Theta_2$  を大きく異なる値に設定することが最適です。そのため、これらのパラメータを単一のステップサイズの値に縮小しないほうがよい。次の補題では、この内積に関する勾配の計算方法が示されています。

**Lemma 2.3.**  $F = F(\varphi)$  がフレシェ微分可能な汎関数であり、任意の  $\phi$  に対して、任意の点  $h$  で評価される第 1 変分  $\delta F(\phi)$  が関数  $f_\phi$  に対する積分として表されるとします。つまり、

$$\delta F(\phi)(h) = \int_{\Omega} h(x) f_\phi(x) dx$$

とします。式 (24) で  $\|\cdot\|_H$  を定義します。すると、 $F$  の  $H$ -勾配は次のように表されます。

$$\nabla_H F(\phi) = (\Theta_1 \text{Id} - \Theta_2 \Delta)^{-1} f_\phi$$

ここで、 $\text{Id}$  は単位作用素、 $\Delta$  はラプラス作用素であり、ゼロノイマン境界条件を伴って考えます。

*Proof.* 固定した  $\phi$  に対して、楕円型方程式

$$\begin{cases} (\Theta_1 \text{Id} - \Theta_2 \Delta)g = f_\phi & \text{in } \Omega \\ n \cdot \nabla g = 0 & \text{on } \partial\Omega \end{cases} \quad (25)$$

の一意な解を考えます。ただし  $n(x)$  は  $\Omega$  の境界上での単位法線ベクトル。すると、以下の等式の連鎖が得られます：

$$\begin{aligned} \delta F(\phi)(h) &= \int_{\Omega} h(x) f_\phi(x) dx \\ &= \int_{\Omega} h(x) (\Theta_1 \text{Id} - \Theta_2 \Delta)g(x) dx \\ &= \int_{\Omega} h(x) \Theta_1 dx - \int_{\Omega} h(x) \Theta_2 \Delta g(x) dx \\ &= \int_{\Omega} \Theta_1 h(x) g(x) dx - \int_{\Omega} \Theta_2 h(x) \Delta g(x) dx \\ &= \int_{\Omega} \Theta_1 h(x) g(x) dx - [h(x) \Theta_2 \nabla g(x) \cdot n]_{\partial\Omega} + \int_{\Omega} \Theta_2 \nabla h(x) \cdot \nabla g(x) dx \quad (\text{部分積分} \cdot \text{ガウスの発散定理}) \\ &= \int_{\Omega} \Theta_1 h(x) g(x) dx + \int_{\Omega} \Theta_2 \nabla h(x) \cdot \nabla g(x) dx \\ &= \langle h, g \rangle_H. \end{aligned}$$

これにより、 $g$  が  $F$  の  $H$ -勾配であることが示されます。

□

上記の結果は次のように言い換えることができます： $F$  の  $H$ -勾配は、 $\delta F$  を逆演算子  $(\Theta_1 \text{Id} - \Theta_2 \Delta)^{-1}$  で「事前調整」することによって得られます。

つまり、 $F$  の  $H$ -勾配を計算する場合、 $\delta F$  を直接適用する代わりに、まず逆演算子  $(\Theta_1 \text{Id} - \Theta_2 \Delta)^{-1}$  を  $\delta F$  に適用します。この逆演算子を適用するプロセスは、「事前調整」と呼ばれ、収束特性を改善し、 $H$ -勾配の計算を効率化するために、 $\delta F$  を修正する役割を果たします。

### 3 THE BACK-AND-FORTH METHOD

私たちの目標は、興味深いエネルギー関数  $U$  の大きなクラスに対して、JKO スキームを効率的に解くアルゴリズムを開発することです。第 3.1 節では、 $U$  が  $\rho$  に関して凸である場合を考えます。この場合、JKO ス

キームには等価な双対問題があり、[21] からの back-and-forth メソッドの改良版を使用して解決します。第 3.2 節では、以下の形式の凸エネルギーに対して、アルゴリズムが適切に重み付けられた  $H^1$  空間で勾配安定性を持つことを示します。

$$U(\rho) = \int_{\Omega} u_m(\rho(x)) + V(x)\rho(x) dx$$

ここで、 $V : \Omega \rightarrow [0, +\infty]$  は固定された関数であり、

$$u_m(\rho) = \begin{cases} \frac{\gamma}{m-1}(\rho^m - \rho) & \text{if } \rho \geq 0 \\ +\infty & \text{otherwise} \end{cases} \quad (26)$$

となります。ただし、定数  $\gamma > 0$  と  $m > 1$  が存在します。また、 $m \rightarrow 1$  および  $m \rightarrow \infty$  の 2 つの極限の場合も考慮します。私たちの解析はより一般的な汎関数に拡張することも可能ですが、わかりやすさのために上記の特殊なケースに焦点を当てています。凸エネルギー関数  $U$  のメソッドを開発した後、第 3.3 節では非凸な  $U$  に対してアルゴリズムを一般化する方法を示します。

### 3.1 The back-and-forth method for convex $U$

JKO スキームを反復するためには、任意の非負密度  $\mu \in L^1(\Omega)$  に対して、以下の一般化最適輸送 (GOT) 問題を解くことが必要です:

$$\rho_* = \operatorname{argmin}_{\rho \in L^1(\Omega)} U(\rho) + \frac{1}{2\tau} W_2(\rho, \mu), \quad (27)$$

ここで、 $U$  は凸である場合には、一般化最適輸送問題は双対性を持つことが Section 2.2 で示されました (Thm. 2.2)。具体的には、 $U$  が凸であるとき、一般化最適輸送問題は次のような対双対関係にあります。

$$\inf_{\rho \in L^1(\Omega)} U(\rho) + \frac{1}{2\tau} W_2^2(\rho, \mu) = \sup_{\phi} J(\phi) = \sup_{\psi} I(\psi)$$

$I$  と  $J$  は次のように定義されます。

$$J(\phi) = \int_{\Omega} \phi^c(x) \mu(x) dx - U^*(\phi) \quad (28)$$

$$I(\psi) = \int_{\Omega} \psi(x) \mu(x) dx - U^*(\psi^c) \quad (29)$$

さらに、問題 (27) の最小化解  $\rho_*$  は、最大化解  $\phi_*, \psi_*$  と関連しており、次の関係が成り立ちます。

$$\rho_* = T_{\phi_* \#} \mu, \quad \rho_* \in \delta U^*(\phi_*), \quad \phi_*^c = \psi_* \quad (30)$$

$I$  と  $J$  は制約のない凹関数汎関数です (Prop. 2.3 を参照)。したがって、いずれかの関数の最大化解を標準的な勾配上昇法で見つけることができます。一方、単に  $I$  または  $J$  だけを使用することは問題の対称性を破壊します。したがって、関数のいずれかにだけ焦点を当てるのではなく、back-and-forth メソッドでは  $I$  と  $J$  の交互の勾配上昇ステップが行われます。 $I$  と  $J$  は異なる変数を使用していますが、 $c$ -変換を使用することで  $\psi$  と  $\phi$  の間を切り替えることができます。[21] で指摘されているように、 $I$  と  $J$  による交互のステップは、標準的な勾配上昇法を超えた方法の収束速度を大幅に加速させます。

今、問題 (27) の双対最大化子  $(\phi_*, \psi_*)$  を見つけるためのアプローチを紹介します。この方法は、アルゴリズム 2 に概要が示されており、以下の 2 つの主要なアイデアに基づいています:



### 3.2 $H^1$ gradient ascent

### 3.3 Back-and-forth for non-convex $U$

## 4 NUMERICAL IMPREMENTATION AND EXPERIMENTS

### 4.1 Implementation details

### 4.2 Experiments

#### 4.2.1 Accuracy: Barenblatt solutions

#### 4.2.2 Slow diffusion with drifts and obstacles

#### 4.2.3 Non-convex $U$ (aggregation-diffusion)

#### 4.2.4 Incompressible projections and flows

## 5 Appendix

### 自乗可積分

自乗可積分関数 (square-integrable function) とは、実数値または複素数値可測関数で絶対値の自乗の積分が有限であるものである。すなわち

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < +\infty$$

ならば、 $f$  は実数直線  $(-\infty, +\infty)$  上で自乗可積分である。場合によっては積分区間が  $[0, 1]$  のように有界区間のこともある。

### Legendre 変換 (凸共役)

関数  $\xi$  を  $\omega$  上で定義された凸で微分可能な関数としたとき、その Legendre 変換 (凸共役)  $\xi^*$  は以下のように定義できる。

$$\xi^*(p) := \sup_x p \cdot x - \xi(x)$$

ここで、双対汎関数  $J$  は  $J(\varphi) = \int \phi \nu + \int \phi^c \mu$  で定義していたことを思い出す。これは凹な汎関数であり、第一項  $\int \phi \nu$  は  $\phi$  に対して線形であるため、 $J$  の凸性には影響しません。そのため、 $F$  を次のように定義します。

$$F(\phi) = - \int \phi^c \mu, \quad (31)$$

これは、本質的には線形項のみが異なるため、 $J$  と同じ凸性を持つ凸関数です。例えば、任意のポテンシャル  $\phi_1$  と  $\phi_2$  に対して、次の式が成立することを直接確認できます。

$$J(\phi_2|\phi_1) = -F(\phi_2|\phi_1). \quad (32)$$

最後に、 $F$  の凸共役は次のように定義されます。

$$F^*(\rho) := \int_{\omega} \phi \rho - F(\phi) \quad (33)$$

## Proof

*Proof of*  $T_\phi(x) = \operatorname{argmin}_{y \in \Omega} \left( \frac{1}{2\tau} |x - y|^2 - \phi(y) \right)$  は凸集合.  $\phi$  が  $c$ -concave であると仮定します。

関数  $f(y) = \frac{1}{2\tau} |x - y|^2 - \phi(y)$  を考えます。この関数は凸関数です。なぜなら、次のような理由からです：

- (i) 二次関数  $\frac{1}{2\tau} |x - y|^2$  は凸関数です。二次関数はそのヘッセ行列が半正定値であるため、凸性を持ちます。
- (ii)  $\phi$  が  $c$ -凹関数より、 $-\phi(y)$  は下に凸です。 $\phi$  が  $c$ -concave であるため、 $\phi$  の下に凸な接平面が存在し、その接平面よりも下に凸な部分を持ちます。
- (iii) 凸関数の足し算は凸関数。

したがって、 $f(y) = \frac{1}{2\tau} |x - y|^2 - \phi(y)$  は凸関数です。

関数  $f(y)$  の最小値を実現する点の集合は凸集合です。なぜなら、凸関数の最小値を実現する点の集合は常に凸集合であり、 $f(y)$  は凸関数であるためです。

したがって、 $\operatorname{argmin}_{y \in \Omega} \left( \frac{1}{2\tau} |x - y|^2 - \phi(y) \right)$  は凸集合です。 □