arXiv:2603.01977v1 [math.AP] 2 Mar 2026

核均值差异的 Wasserstein 梯度流的定量收敛

Lénaïc Chizat、Maria Colombo、Roberto Colombo、Xavier Fernández-Real

摘要

我们研究核均值差异（KMD）（亦称最大均值差异（MMD））泛函的 Wasserstein 梯度流的定量收敛。我们的研究框架尤其涵盖无限宽度和连续时间极限下浅层神经网络的训练动力学，以及平均场和过阻尼极限下具有成对 Riesz 核相互作用的相互作用粒子系统。我们的主要分析针对如下模型情形：KMD 泛函由平方 Sobolev距离 $\mathscr{E}_s^\nu(\mu) = \frac{1}{2}\|\mu - \nu\|_{H^{-s}}^2$ 给出，其中 $s \geq 1$，且 ν是定义在 d 维环面上的固定概率测度。首先，受 $2d$-Euler-方程的 Yudovich理论启发，我们在自然的弱正则性类中证明了存在性和唯一性。接着，我们证明，当 $s = 1$ 时，该梯度流在最小假设下以指数速率全局收敛；而当 $s > 1$ 时，我们证明其以多项式速率局部收敛，该速率显式取决于 s 以及 μ和 ν的 Sobolev正则性。这些速率在能量层面和更高正则性类中均成立，并且对于 ν 均匀而言是紧的。随后，我们考察具有 ReLU激活函数的浅层神经网络的总体损失的梯度流；该梯度流可表示为球面 $\mathbb{S}^d$上非负测度空间中的 Wasserstein–Fisher–Rao梯度流。利用其与 $s = (d+3)/2$的 Sobolev能量情形之间的对应关系，我们为该动力学推导出显式的多项式局部收敛速率。除 $s = 1$这一特殊情形外，在所有这些设置中，即使是非定量收敛，此前也仍是一个未解决的问题。我们还在维数 $d = 1$ 下使用偏微分方程和粒子方法开展了数值实验，以说明我们的分析结果。

MSC： 49Q22, 68T07, 49J45, 35Q68.
关键词： Wasserstein梯度流、最大均值差异、Riesz 核、ReLU 神经网络、定量收敛。

目录

1 引言 **2**
1.1 适定性结果 .4 1.2 收敛结果 .5 1.3 无限宽浅层神经网络的定量收敛 .6 1.4 数值示例 .8 1.5 证明思路 .9

2 适定性理论 **11**
2.1 预备知识 .11 2.2 弱类中的局部适定性 .15 2.3 正则性传播 .20

3 定量收敛结果 **26**
3.1 情形 $s = 1$ .26 3.2 情形 $s > 1$. .31

4 连续浅层神经网络的收敛 324.1 化归为 Wasserstein–Fisher–Rao 流.334.2 球面上的分析与反余弦核算子 .344.3 Wasserstein–Fisher–Rao 动力学的适定性 .384.4 Wasserstein–Fisher–Rao 动力学的定量收敛 .42

附录 **46**
A.1 Riesz 核估计 .47A.2 环面上的 Kato–Ponce 交换子估计 d维 .50

1 引言

令 $\mathscr{P}(M)$ 表示 d 维光滑流形 M 上概率测度的集合。给定目标$\nu \in \mathscr{P}(M)$，我们研究从某个 $\bar{\mu} \in \mathscr{P}(M)$ 出发的核均值差异泛函（也称为最大均值差异泛函[1]）的 Wasserstein 梯度流的适定性和长时间行为。这些泛函具有如下形式

$$\mathscr{E}^{\nu}(\mu) := \frac{1}{2}\int_M \int_M K(x,y)d(\mu-\nu)(x)d(\mu-\nu)(y), \tag{1.1}$$

其中 $K\colon M \times M \to \mathbb{R}$ 是一个对称且条件正定（但不一定连续）的核。其动力学可表示为如下形式的主动标量连续性方程柯西问题的解

$$\partial_t \mu_t = \operatorname{div}\left(\mu_t \nabla \mathcal{K}\left(\mu_t - \nu\right)\right) \qquad \text{in } (0,T)\times M, \qquad\qquad \mu_0 = \bar{\mu}, \tag{1.2}$$

其中 $\mathcal{K}$ 是由下式给出的半正定算子

$$\mathcal{K}(\eta)(x) := \int_M K(x,y)d\eta(y).$$

方程（1.2）可解释为定律为 μ_t的带正电粒子过阻尼系统的演化，该系统与固定的带负电背景 ν 通过势 K相互作用。

动机。研究核均值差异型（或最大均值差异型）差异度的兴趣来自机器学习与统计学。例如，（1.2）描述了浅层（即单隐藏层）神经网络训练动力学的平均场（无限宽度）极限，其中 μ_t 表示参数的演化分布，目标泛函 $\mathscr{E}^{\nu}$ 与总体损失一致，而核 K 依赖于激活函数和输入数据分布 [MMN18, RVE22, SS20, CB18]（另见 [FRF22]）。此类动力学也在生成建模的背景下受到启发，其目标是找到一个映射，将易于采样的源密度输运到目标密度（参见例如 [AKSG19, AHS23, GdBG25, CMG+25]）。在此背景下，（1.2）可解释为训练动力学的简化模型，即通过基于梯度的优化最小化目标 $\theta \mapsto \mathscr{E}^{\nu}((f_\theta)_{\#}\mu_0)$，其中 f_θ 是参数为 θ 的神经网络，而 μ_0 是参考概率测度 [UNS+18]。

定量收敛。尽管二次目标函数（1.1）表面上十分简单，但除了特殊情形之外，人们对其 Wasserstein 梯度流的定性和定量收敛性质仍知之甚少。关键障碍在于几何结构：虽然 $\mathscr{E}^{\nu}$ 对于测度的线性结构是凸的（这得益于核的正性），但通常不在 $(\mathscr{P}(M)\ W_2)$上沿测地线凸，因此标准的压缩

[1]MMD 这一术语由 [GBR+06] 提出，指一种一般的积分概率度量（IPM）。在此，我们遵循 [Wai19], 的做法，采用 KMD 这一名称，以表明我们使用的是基于核的积分概率度量。在机器学习中，人们通常假设该核生成一个再生核希尔伯特空间（RKHS）（例如在（1.3）的 Riesz 情形下），但我们的分析并不需要这一假设。

以及在沿测地线（均匀）凸情形下适用于 Wasserstein 梯度流的定量收敛机制 [AGS08] 并不适用。

已有研究（见上文参考文献）致力于为长时间行为提供保证。例如， [CB18] 证明，若 $\bar{\mu}$ 的支撑满足“全方向性”条件，且 Wasserstein–Fisher–Rao 梯度流（见第 1.3 节）收敛，则其收敛到一个全局极小化子；但这一结果不适用于 Wasserstein 梯度流。即使对于 Wasserstein–Fisher–Rao 流，这也只是部分保证，因为它并未证明收敛，更遑论提供有关收敛速率的定量信息。此外， [AKSG19] 证明，若沿流的过程中 $\mu_t - \nu$满足有界性条件，则 Wasserstein 梯度流会以速率 $O(1/t)$全局收敛。然而，鉴于紧密相关情形中的反例 [SS18]，尚不清楚这一假设何时成立或失效，甚至一般情况下是否可以期待全局收敛。总而言之，核均值差异的 Wasserstein（或 Wasserstein–Fisher–Rao）梯度流的长时间收敛保证仍是一个悬而未决的问题，即便是在局部且非定量的层面上也是如此。

一个模型案例：环面上的 Riesz 核。 我们聚焦于环面上的 Riesz 核这一 d-环面模型案例。不过需要强调的是，我们的方法具有更广泛的适用性，正如第 1.3 节对 ReLU 浅层神经网络的分析所示（其中 $M = \mathbb{S}^d$ 且 K 是另一种核）。

令 M 为 d 维环面 $\mathbb{T}^d \cong \mathbb{R}^d/\mathbb{Z}^d$，并令 $\mathcal{K}$ 为某个幂次 $s \geq 1$的逆拉普拉斯算子，

$$\mathcal{K}(\eta) := (-\Delta)^{-s}\eta, \qquad \int_{\mathbb{T}^d} \eta = 0.$$

相应的 能量 $\mathscr{E}_s^{\nu}(\mu)$，(1.1)是 H^{-s}-齐次 Sobolev 差异度，定义于 μ 与 ν之间，

$$\mathscr{E}_s^{\nu}(\mu) = \frac{1}{2}\|\mu - \nu\|_{H^{-s}}^2 = \frac{1}{2}\int_{\mathbb{T}^d}\int_{\mathbb{T}^d} K_s(x-y)d(\mu-\nu)(x)d(\mu-\nu)(y), \tag{1.3}$$

其中 K_s： $\mathbb{T}^d \to \mathbb{R}$ 是 Riesz 核，满足 $(-\Delta)^s K_s = \delta_0 - 1$ （在分布意义下）。在此框架中，演化（1.2）可写为

$$\begin{cases} \partial_t \mu + \operatorname{div}(\mu v) &= 0 & \text{in } (0,T)\times\mathbb{T}^d, \\ v_t &= -\nabla K_s * (\mu_t - \nu) & \forall t \in (0,T), \\ \mu_0 &= \bar{\mu}. \end{cases} \tag{1.4}$$

（关于（1.4）的弱解的精确定义，参见定理 1.4。）

如上所述，Riesz核相互作用已引起理论界和应用领域的关注。随着 $s \geq 1$ 变化，短程相互作用 $K_s(x-y) \sim_{c_{d,s}} |x-y|^{2s-d}$ 呈现出不同的渐近行为，从而形成一系列正则性区间，如图 1 所示（详见定理 A.1），进而决定了方程（1.4）的解所表现出的多种定性行为。我们至少可以指出 s三个特别值得关注的取值:

1) ($s = 1$)。此时，(1.3) 恰好是带符号电荷分布 $\mu - \nu$ 的库仑相互作用能。由于其在物理学上的重要意义，数学文献对库仑相互作用进行了广泛研究（参见 [Ser24], [BV25]及其中的参考文献）。2) $\left(s = \frac{d}{2} + \frac{1}{2}\right)$。此时，核 $K(x-y)$ 在短程相互作用下表现为负距离 $-|x-y|$。负距离核 [SR13]及其 Wasserstein 梯度流，例如在 [HHA+24、HWAH24、HGBS24]，中均有研究，并应用于机器学习和生成建模。3) $\left(s = \frac{d}{2} + \frac{3}{2}\right)$。如第 1.3 节中详细说明的，这种情形与理解无限宽浅层 ReLU 神经网络上的梯度流有关。

本文分析方程（1.4），对所有 $s \geq 1$，探讨自然函数类中的适定性问题，以及收敛到极小化子的定量收敛问题。

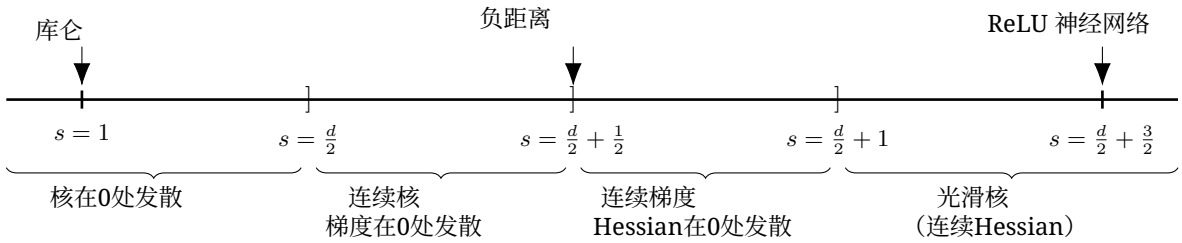


图 1：环面上的 Riesz 核正则性区间 K_s 以及一些特别值得关注的取值（箭头）；括号方向表示端点是否包含。负距离和 ReLU 神经网络情形严格来说并非环面上的 Riesz 核，但它们在对角线上的正则性与所示指数的正则性一致。对于 $d=1$，当 $s\geq 1$ 时，核始终不奇异（此时库仑情形与负距离情形重合）。

1.1 适定性结果

下面的收敛结果建立在方程（1.4）稳健的适定性理论之上。因此，我们首先证明自然函数类中的存在性、唯一性、稳定性和正则性传播。这些结果本身也具有重要意义，并且对于定量长时间分析至关重要。

对每个 $s\geq 1$，我们确定一个自然的弱解类，其指导原则是：该类中的解应生成一个（拟）利普希茨向量场。我们用 $\mathcal{M}(\mathbb{T}^d)$ 表示 $\mathbb{T}^d$ 中有限（带符号）测度的空间，并用 $L^{p,1}(\mathbb{T}^d)$ 表示洛伦兹空间（见下文（2.5）；或者，对某个 $q>p$，也可以用 L^q 替换$L^{p,1}$ ）。我们始终将绝对连续测度与其关于勒贝格测度的密度等同。最后，我们令

$$\mathscr{X}_s(\mathbb{T}^d):=\begin{cases} L^\infty(\mathbb{T}^d) & \text{if } s=1,\\ L^{p,1}(\mathbb{T}^d) \quad \text{for } p=\frac{d}{2s-2} & \text{if } s\in\left(1,\frac{d}{2}+1\right),\\ \mathcal{M}(\mathbb{T}^d) & \text{if } s\geq\frac{d}{2}+1. \end{cases} \tag{1.5}$$

命题1.1（局部适定性）。设 $s\geq 1$，并设 $\mathscr{X}_s(\mathbb{T}^d)$ 由（1.5）给出。则对每个$\bar{\mu}$，$\nu\in\mathscr{P}\cap\mathscr{X}_s(\mathbb{T}^d)$ 存在极大存在时间 $T>0$ 以及唯一极大解 $\mu\in L^\infty_{\mathrm{loc}}([0,T);\ \mathscr{X}_s(\mathbb{T}^d))$，即方程（1.4）在定理 2.3意义下的解。此外：

- 如果 $s\geq\frac{d}{2}+1$，那么 $T=\infty$。
- 如果 $s\in\left[1,\frac{d}{2}+1\right),$，则当且仅当 上极限$\limsup_{t\to T^-}\|\mu_t\|_{L^p}=+\infty$ 时有 $T<\infty$（1，∞]。 ——

最后， 解保持 *Hölder* 正则性和索博列夫正则性（参见定理 2.10 和定理 2.14 ）.

该结果受 Yudovich理论中关于 L^∞-解的 $2d$-Euler 方程涡量形式[Yud63], 的启发，不过我们的证明采用了近期在 [MP94] 和 [Loe06]中发展的方法。另请参见[BLL12、BLR11]，其中讨论了不含目标测度的相关模型的适定性。尽管我们证明的结果针对 $M=\mathbb{T}^d$ 和 $K=K_s$ Riesz核，但相同的方法也适用于更一般的 d维流形（例如 $M=\mathbb{S}^d$ 或 $M=\mathbb{R}^d$）以及其他具有可比渐近行为的核。

对于 $s>d/2+1$，K_s是半凸的，因此根据一般理论[AGS08]（另见 [CLM13]），可得方程（1.4）在 $\mathcal{M}$ 上的全局适定性[AGS08]（另见 [CLM13]）。在端点 $s=d/2+1$，K_s并非半凸，但我们仍可利用其诱导速度场的对数Lipschitz正则性，得到方程（1.2）在 $\mathcal{M}$上的全局适定性。

我们在第1.1 节证明定理2，并在该节得到其他性质，例如，解在 Wasserstein 度量下关于初始测度与目标测度变化的定量稳定性$\bar{\mu}$ 以及 ν。此外，与上述形式化辨识一致，我们在定理2.5中指出，由定理 μ（1.4）给出的解1.1 是 Riesz偏差能量$\mathscr{E}_s^\nu$的 Wasserstein 梯度流（见定理2.1）。事实上，$\mu\in\mathrm{Lip}_{\mathrm{loc}}\left([0,T);\ \left(\mathscr{P}(\mathbb{T}^d)\quad W_2\right)\right)$ 并且能量耗散恒等式成立：

$$\frac{d}{dt}\mathscr{E}_s^\nu(\mu_t)=-\int_{\mathbb{T}^d}|\nabla K_s*(\mu_t-\nu)|^2 d\mu_t \qquad \forall t\in(0,T). \tag{1.6}$$

1.2 收敛结果

接下来陈述主要结果：对于1.2，定理 $s=1$；对于1.4，定理 $s>1$，讨论方程（1.4）的解收敛到目标测度的问题。这样的划分反映了动力学在定性上的变化：端点情形$s=1$ 具有额外的结构（例如最大值原理），而 $s>1$ 这一范围在技术上要求更高，也是本文的核心情形。

定理 1.2（全局收敛到目标：$s=1$）。设 $s=1$，$\bar{\mu}$，$\nu\in\mathscr{P}\cap L^\infty(\mathbb{T}^d)$ 并设 $\mu\in L^\infty_{\mathrm{loc}}([0,T)$；$L^\infty(\mathbb{T}^d)$ 为定理（1.4）给出的极大解，见定理 *1.1*。那么，$T=\infty$ 且有：

$$\min\{\inf\bar{\mu},\inf\nu\}\le\inf\mu_t\le\sup\mu_t\le\max\{\sup\bar{\mu},\sup\nu\}\qquad\forall t\in[0,\infty);\tag{1.7}$$

$$\mu_t\ \textit{converges weakly-}*\ \textit{in}\ L^\infty(\mathbb{T}^d)\ \textit{to}\ \nu\ \textit{as}\ t\to\infty.\tag{1.8}$$

*Su*此外，假设 $\bar{\mu},\nu\ge\alpha>0$ 在 $\mathbb{T}^d$ 中几乎处处成立。则有以下结论：

i）（能量的指数弱收敛及 W_2）。有：

$$\alpha^{1/2}W_2(\mu_t,\nu)\le\|\mu_t-\nu\|_{\dot{H}^{-1}}\le\|\bar{\mu}-\nu\|_{\dot{H}^{-1}}e^{-\alpha t}\qquad\forall t\in[0,\infty).\tag{1.9}$$

ii）（*Uniform* 收敛）。假设 $\nu\in C(\mathbb{T}^d)$ 具有 *Dini* 连续模[2]。那么

$$\|\mu_t-\nu\|_{L^\infty}\to0\qquad\textit{as}\ t\to\infty.\tag{1.10}$$

此外，只要 ν 是 *Hölder* 连续的，一致收敛就是指数收敛。

iii)（光滑收敛）。设 $\gamma>d/2$ 并假设 $\bar{\mu}\in H^\gamma(\mathbb{T}^d)$，$\nu\in H^{\gamma+1}(\mathbb{T}^d)$（见第 *2.1.1* 节）。那么，存在一个常数 $C>0$，它仅取决于 d、α、γ、$\|\bar{\mu}\|_{H^\gamma}$和 $\|\nu\|_{H^{\gamma+1}}$，使得

$$\|\mu_t-\nu\|_{\dot{H}^\gamma}\le\left(\|\bar{\mu}-\nu\|^2_{\dot{H}^\gamma}+C(1-e^{-t/C})\right)^{1/2}e^{-\alpha t+C(1-e^{-t/C})}\qquad\forall t\in[0,\infty).$$

注 1.3（下界的尖锐性）。我们强调，对初始测度 $\bar{\mu}$ 的下界并非实现向目标 ν指数收敛的必要条件，正如第3.1.3 节所述（见定理3.3）。事实上，若 $\nu\ge\alpha>0$，则 $\bar{\mu}$ 为零的任意区域都会随时间推移而以指数速度“填满”；我们称之为孔洞的指数填充（见定理 3.2）。相比之下，第1.4 节中的数值实验表明，要确保指数收敛，ν 必须有正下界。■

在 [BV25] 中，作者发起了对（1.4）的研究，针对 $s=1$证明了（1.7），并在初始密度和目标密度具有 Hölder 连续性这一更强假设下，得出了相应的指数收敛（1.9）。

本文的结果通过一种独立且自洽完整的方法获得，并且既强化又完善了上述结论：

i）首先，我们在有界密度这一更一般的情形下开展研究（这是我们预期能够保证适定性的最一般空间），并首先建立最大值原理（1.7），从而在相应的正下界假设下推导出指数收敛。

ii）我们进一步证明了解向极小化子的无条件全局定性收敛（1.8），在存在正下界时推导出定量收敛速率，并在数据具有适当正则性的条件下，将解的收敛扩展到更高正则性拓扑。特别地，Dini连续目标蕴含一致收敛，而对于不连续目标，则无法期待一致收敛。

iii) 最后，我们还能够证明能量的指数收敛，而无需对初始测度施加任何下界（见定理 1.3），如第 1.4节中的数值模拟所示。

[2]连续模是一个严格递增、连续且凹的函数 $\omega:[0,\infty)\to[0,\infty)$ 满足 $\omega(0)=0$。我们称 $f:\mathbb{T}^d\to\mathbb{R}$ 具有连续模 ω，如果 $|f(x)-f(y)|\le\omega(|x-y|)$ 对所有 x，$y\in T^d$。我们称 ω 是 Dini 连续模，如果满足 $\int_0^1\frac{\omega(r}{r}dr<\infty$.)

接下来，我们陈述情形 $s > 1$ 下的定量局部收敛结果。我们认为这是本文的主要结果。

定理 *1.4*（局部收敛至目标： $s > 1$）。设 $s > 1$, $\gamma > d/2$, $\bar{\mu} \in \mathscr{P} \cap H^{\gamma}(\mathbb{T}^d)$, $\nu \in \mathscr{P} \cap H^{\gamma+s}(\mathbb{T}^d)$, 满足 $\bar{\mu}, \nu \geq \alpha > 0$, 并令 $\mu \in L^{\infty}_{\mathrm{loc}}([0,T); H^{\gamma})$ 为定理 *1.1* 给出的方程 *(1.4)* 的唯一极大解。存在常数 C, $\delta > 0$，仅依赖于 $d, s, \gamma, \alpha, \|\bar{\mu}\|_{H^{\gamma}}$ 和 $\|\nu\|_{H^{\gamma+s}}$，使得若 $\|\bar{\mu} - \nu\|_{\dot{H}^{-s}} \leq \delta$，则

$$T = \infty, \qquad \|\mu_t - \nu\|_{\dot{H}^{-s}} \leq \|\bar{\mu} - \nu\|_{\dot{H}^{-s}} (1 + t/C)^{-\frac{\gamma+s}{2(s-1)}} \qquad \text{and} \qquad \|\mu_t\|_{H^{\gamma}} \leq C \qquad \forall t \in [0, \infty). \quad (1.11)$$

特别地，由索博列夫插值可得，

$$\|\mu_t - \nu\|_{\dot{H}^{\gamma'}} \leq C^{\frac{\gamma'+s}{\gamma+s}} \|\bar{\mu} - \nu\|_{\dot{H}^{-s}}^{\frac{\gamma-\gamma'}{\gamma+s}} (1 + t/C)^{-\frac{\gamma-\gamma'}{2(s-1)}} \qquad \forall t \in [0, \infty), \quad \forall \gamma' \in [-s, \gamma].$$

备注1.5（关于局部性假设）。在 [SS18] 中，作者构造了具有 ReLU 激活函数的离散浅层神经网络的总体平方损失的严格局部极小值示例。这类系统与我们的情形密切相关， $s = \frac{d+3}{2}$ 正如第1.3节所解释的。这表明，局部性假设 $\|\bar{\mu} - \nu\|_{\dot{H}^{-s}} \leq \delta$ 可能是 $s > 1$情形下获得简洁定量收敛结果所必需的。另一方面，需要注意的是，对于库仑相互作用（$s = 1$），则不存在这类离散局部极小值示例，因为根据静电学中的厄恩肖定理，库仑势下不存在点电荷的稳定静止构型。■

注记 1.6（多项式衰减率的尖锐性）。 在对初始数据于 H^{γ} 中的控制下，指数 $2^{\frac{\gamma+s}{s-1}}()$ 在来自（1.11）的能量衰减中是尖锐的。为说明这一点，我们取 $\nu = 1$ 并考察关于 σ_t 的线性化方程：$= \mu_t - 1$:

$$\partial_t \sigma_t = -(-\Delta)^{1-s} \sigma_t.$$

将 σ_t 在Four中展开 ier并求解其系数$\hat{}\sigma_k(t)$，我们得到$\hat{}\sigma_k'(t) = -(2\pi|k|)^{2-2s}\hat{}\sigma_k(t)$，由此给出

$$\hat{\sigma}_k(t) = \hat{\sigma}_k(0) \exp\left(-(2\pi|k|)^{2-2s} t\right) \qquad \forall t \geq 0, \quad \forall k \in \mathbb{Z}^d \setminus \{0\}. \quad (1.12)$$

现在，对任意整数 $n \geq 1$，我们考虑初始数据 $\sigma_0^n(x) := \sqrt{2}(2\pi n)^{-\gamma} \cos(2\pi n x_1)$，对此我们有 $\|\sigma_0^n\|_{\dot{H}^{\gamma}} = 1$。由（1.12）我们得到

$$\|\sigma_t^n\|_{\dot{H}^{-s}} = \sqrt{2}(2\pi n)^{-\gamma-s} \exp\left(-(2\pi n)^{2-2s} t\right) \gtrsim t^{-\frac{\gamma+s}{2(s-1)}} \qquad \text{for } t = n^{\frac{1}{2s-2}}.$$

这表明，在所有具有单位 H^{γ}-范数的初始数据中，我们不能期望 H^{-s}-范数在时刻 t 小于 $c_{s,\gamma} t^{-2^{\frac{\gamma+s}{s-1}}()}$，从而证明了定理 1.4 中衰减率的最优性。■

因此，对于 $s > 1$ 我们在自然的小差异假设下，于一系列黎曼型情形中得到多项式松弛速率（关于 H^{-s}，并通过插值推广至中间 Sobolev 范数）。这些情形尤其包括 $s = \frac{d}{2} + \frac{1}{2}$ 时的负距离（“能量距离”）核，该核广泛应用于统计学，以及近期用于成像和生成建模的基于流的方法 [SR13, ĤHA⁺24, HWAH24]。在维度 $d \geq 2$ 下，负距离核的全局时间收敛问题超出了测地凸性技术的适用范围 [BV25, DSB⁺25]（这是因为泛函缺乏一致测地凸性）。我们的分析揭示了一种稳健机制——通过高阶能量估计传播的局部 Łojasiewicz 不等式；此外，该机制还可推广至其他核和几何。我们在下文第 1.3 节中通过处理球面上的反余弦/ReLU 核来说明这一点。关于定理 1.5 中的收敛结果背后的证明思路，参见第 1.2 节和 1.4节。

1.3 无限宽浅层神经网络的定量收敛

考虑一个无限宽的 ReLU 神经网络，即一个函数 $f_{\mu} : \mathbb{R}^{d+1} \to \mathbb{R}$，它通过概率测度 $\mu \in \mathscr{P}(\mathbb{R}^{d+2})$ 参数化，其表达式为

$$f_{\mu}(x) := \int_{\mathbb{R}^{d+2}} \Phi(w, x) d\mu(w), \qquad w = (a, b) \in \mathbb{R} \times \mathbb{R}^{d+1}, \quad \Phi(w, x) = a(b \cdot x)_+. \quad (1.13)$$

注意， f_μ 是关于 x 的正 1 次齐次函数，因此我们不妨将其输入限制在单位球面 $\mathbb{S}^d$ 上。对于给定的 $\nu \in \mathcal{P}(\mathbb{R}^{d+2})$，我们考虑均方能量

$$\mathscr{E}^{\nu}_{\mathbb{S}^d}(\mu) := \frac{1}{2}\int_{\mathbb{S}^d} |f_\mu - f_\nu|^2 dx = \frac{1}{2}\int_{\mathbb{R}^{d+2}}\int_{\mathbb{R}^{d+2}} \hat{K}(w,w')d(\mu-\nu)(w)d(\mu-\nu)(w'),$$
$$\hat{K}(w,w') = \int_{\mathbb{S}^d} aa'(b\cdot x)_+(b'\cdot x)_+ dx, \qquad w=(a,b), w'=(a',b') \in \mathbb{R}\times\mathbb{R}^{d+1}.$$

在初始化 化 $\bar{\mu} \in \mathscr{P}(\mathbb{R}^{d+2})$ 下， $\mathscr{E}^{\nu}_{\mathbb{S}^d}$ 的 Wasserstein 梯度流由以下方程给出

$$\begin{cases} \partial_t \mu_t + \operatorname{div}(\mu_t v_t) &= 0 & \text{in } (0,T)\times\mathbb{R}^{d+2}, \\ v_t &= -\nabla \hat{\mathcal{K}}(\mu_t - \nu) & \forall t \in (0,T), \\ \mu_0 &= \bar{\mu}, \end{cases} \tag{1.14}$$

其中

$$\hat{\mathcal{K}}(\eta)(w) := \int_{\mathbb{R}^{d+2}} \hat{K}(w,w')d\eta(w') \qquad \forall \eta \in \mathcal{M}(\mathbb{R}^{d+2}).$$

该动力学描述了浅层 ReLU 神经网络在总体平方损失下通过（随机）梯度下降进行训练时参数的演化，其初始权重独立采样自$\bar{\mu}$；输入数据在球面上均匀分布，贝叶斯预测函数为 f_ν。这对应于小学习率和无限宽度极限（关于这一联系的详情，参见 [MMM19、Woj20、CB20] ）。

利用 $\Phi(w,\ x)$ 的结构（尤其是分别关于 a 和 b的一次齐次性），我们还可以将演化过程中的输出进一步化简为

$$f_{\mu_t}(x) = \int_{\mathbb{S}^d} (b\cdot x)_+ d\mu_t(b)$$

对于 $\mathbb{S}^d$上的某个（有符号）偶测度，为简便起见，我们仍将其记作 μ_t。此外，除去仅取决于 $f_{\bar{\mu}}$ 和 f_ν（初始输出和目标输出）正则性的一个加性常数后，我们还可以进一步假设，在训练动力学过程中 μ_t 是一个非负测度；该测度并非守恒，其表达式为

$$\begin{cases} \partial_t \mu_t + \operatorname{div}_{\mathbb{S}^d}(\mu_t v_t) &= -4\mathcal{K}(\mu_t-\nu)\mu_t & \text{in } (0,T)\times\mathbb{S}^d, \\ v_t &= -\nabla_{\mathbb{S}^d}\mathcal{K}(\mu_t-\nu) & \forall t\in(0,T), \\ \mu_0 &= \bar{\mu}, \end{cases} \tag{1.15}$$

其中我们记

$$\mathcal{K}(\eta)(x) = \int_{\mathbb{S}^d} K(x,y)d\eta(y) \quad \forall \eta \in \mathcal{M}(\mathbb{S}^d), \qquad K(x,y) = \int_{\mathbb{S}^d} (x\cdot\xi)_+(y\cdot\xi)_+ d\xi, \quad x,y\in\mathbb{S}^d. \tag{1.16}$$

这样一来，我们将网络的表达能力略微限制为偶函数且足够正则的输出。我们请读者参阅第4.1节，其中我们给出关于这一约化的进一步细节。

在下文中，我们因此考虑数据 $\bar{\mu},\ \nu \in \mathcal{M}_+(\mathbb{S}^d)$ [3], 在 $\mathbb{S}^d$ 上为偶函数，并研究解 $t\mapsto\mu_t \in \mathcal{M}_+(\mathbb{S}^d)$ 所对应的方程（1.15）。我们证明，在偶函数非负测度类中存在唯一全局解，其关于时间弱-$*$ 连续，并且它传播数据的赫尔德正则性与索博列夫正则性（见定理 4.5 与定理 4.7）。这里相对于（1.4）的适定性理论，需要一种略有不同的方法，因为我们必须处理额外的非守恒项 $-4\mathcal{K}(\mu_t-\nu)\mu_t$在右端。

式（1.15）的动力学可解释为一个Wasserstein–Fisher–Rao梯度流（例如参见 [GM17、LMS23、Chi22]），其能量 $\mathscr{E}^{\nu}_{\mathbb{S}^d}$： $\mathcal{M}_+(\mathbb{S}^d) \to [0,\infty)$ 由下式给出

$$\mathscr{E}^{\nu}_{\mathbb{S}^d}(\mu) = \frac{1}{2}\int_{\mathbb{S}^d}\int_{\mathbb{S}^d} K(x,y)d(\mu-\nu)(x)d(\mu-\nu)(y) = \frac{1}{2}\int_{\mathbb{S}^d}\mathcal{K}(\mu-\nu)(x)d(\mu-\nu)(x), \tag{1.17}$$

其对应的耗散恒等式为

$$\frac{d}{dt}\mathscr{E}^{\nu}_{\mathbb{S}^d}(\mu_t) = -\int_{\mathbb{S}^d}\left(|\nabla_{\mathbb{S}^d}\mathcal{K}(\mu_t-\nu)|^2 + 4|\mathcal{K}(\mu_t-\nu)|^2\right)d\mu_t \qquad \forall t\in(0,\infty). \tag{1.18}$$

[3] $\mathbb{S}^d$上的（非负）有限测度空间 .

分析算子的谱行为 $\mathcal{K}$ （见第4.2节），我们证明

$$\mathscr{E}^{\nu}_{\mathbb{S}^d}(\mu) \approx_d (\mu(\mathbb{S}^d) - \nu(\mathbb{S}^d))^2 + \|\mu - \nu\|^2_{\dot{H}^{-s}(\mathbb{S}^d)} \qquad \forall \mu, \nu \in \mathcal{M}_+(\mathbb{S}^d) \text{ even}, \qquad s = \frac{d+3}{2}. \tag{1.19}$$

接着，将定理 1.4中的论证调整到这一情形，我们得到以下局部多项式收敛保证（见第 4.4节）。尽管当 ν 是稀疏测度 [AS21、Chi22、LMZ20、ZGJ21、ZG24], 时，已有一些关于平均场极限下浅层神经网络的局部保证，但据我们所知，我们的结果是首个适用于 ν 具有密度且属于真正无限维空间的情形的收敛结果。

定理 *1.7*（神经网络的收敛）。设 $d \in \mathbb{N}$ 为奇函数，且 $s := \frac{d+3}{2} \in \mathbb{N}$。设 $\gamma \in \mathbb{N}$、$\gamma \geq s-1$ 和 $\bar{\mu} \in H^{\gamma}(\mathbb{S}^d)$、$\nu \in H^{\gamma+1}(\mathbb{S}^d)$ 在 $\mathbb{S}^d$ 中为偶函数，并且满足 $\bar{\mu}$、$\nu \geq \alpha > 0$。存在常数 C、$\delta > 0$，它们仅依赖于 d、γ、α、$\|\bar{\mu}\|_{H^{\gamma}}$ 和 $\|\nu\|_{H^{\gamma+1}}$，使得若 $\mathscr{E}^{\nu}_{\mathbb{S}^d}(\bar{\mu}) \leq \delta$，则 *(1.15)* 的解 μ 满足

$$\mathscr{E}^{\nu}_{\mathbb{S}^d}(\mu_t) \leq \mathscr{E}^{\nu}_{\mathbb{S}^d}(\bar{\mu})(1 + t/C)^{-\frac{\gamma+s}{s-1}} \qquad \textit{and} \qquad \|\mu_t\|_{H^{\gamma}} \leq C \qquad \forall t \in [0, \infty). \tag{1.20}$$

有必要作以下几点说明：

注 1.8。 若不限制测度为非负，神经网络便能表示任意正则性足够高的目标函数（无需像下文定理 S 中那样添加常数1.10 ）；但此时我们得到的将是一个双物种演化系统（见第 4.2 节中的式 4.1）。在这种情况下，要得到收敛定理会因种种原因而更具挑战性。例如，对于一对 (μ_t^+, μ_t^-)，其无限时间下的预期极限存在内在的歧义，因为每一对 $(\hat{}\nu^+, \hat{}\nu^-)$ 只要满足 $\hat{}\nu^+ - \hat{}\nu^- = \nu^+ - \nu^-$，就能达到该系统的最小能量。此外，由于两个项之间存在非局部相互作用，通过能量估计来保证更高的索博列夫正则性也更加困难 μ_t^+与 μ_t^-。 ■

注 1.9。 我们假设维数 d 为奇数，仅仅是为了方便处理整数阶导数；但我们认为这一假设在实际中并不重要。 ■

为明确起见，我们直接用目标函数的正则性来陈述定理1.7 的假设 f_{ν}。如前所述，将讨论限制在非负测度上的动力学，意味着我们只能将所有正则性足够好的偶函数表示到一个常数的精度 S。

推论 1.10。 令 $d \in \mathbb{N}$ 为奇函数，$s = \frac{d+3}{2} \in \mathbb{N}$，且 $\beta \in \mathbb{N}$ 满足 $\beta \geq 2s$。令 $g \in H^{\beta}(\mathbb{S}^d)$ 为偶函数，并令 $M > 0$。存在仅依赖于 d、β 和 C 的 $K > 0$，以及依赖于 d、β、M 和 $\|g\|_{H^{\beta}}$ 的 $\delta > 0$，使得以下结论成立。

令 $S = K(\|g\|_{H^{\beta}} + 1)$ 并考虑梯度流 μ_t，其定义于 (1.15)，初始条件为偶初始化 $\bar{\mu} \in \mathcal{M}_+(\mathbb{S}^d)$，目标为 $f = f_{\nu} = g + S$。若 $\|\bar{\mu}\|_{H^{\beta-s-1}} \leq M$，且 $\|f_{\bar{\mu}} - f\|^2_{L^2} \leq \delta$，则

$$\|f_{\mu_t} - f\|^2_{L^2} \leq \|f_{\bar{\mu}} - f\|^2_{L^2}(1 + t/C)^{-\frac{\beta-1}{s-1}} \qquad \textit{and} \qquad \|f_{\mu_t}\|_{H^{\beta-1}} \leq C \qquad \forall t \in [0, \infty).$$

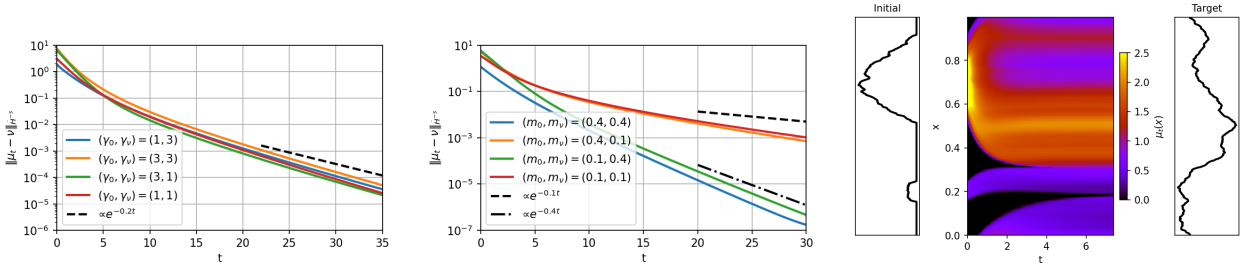
1.4 数值示例

我们在三种设置下进行数值实验： $d = 1$ （A）具有 $s = 1$的黎曼相互作用，（B）具有 $s = 2$的黎曼相互作用，以及（C）浅层 ReLU 神经网络。这三种设置分别展示了定理1.2、定理1.4 和定理1.7。[4]。

在设置（A）和（B）中，我们采用带迎风格式和可变时间步长的有限体积离散。给定离散速度场 v_t（通过在傅里叶域中应用相应的乘子计算得到），该格式在相邻单元之间构造全局中性的质量转移，并与 v_t 一致，从而确保质量守恒和非负性得到精确保持。在设置（C）中，我们采用反余弦核相互作用的粒子离散，这等价于对总体损失进行梯度下降，并对目标 f_{ν} 作逼近。

我们按如下方式生成具有所需索博列夫正则性的随机密度 ν 和 $\bar{\mu}$。首先，通过使用中心正态权重对其傅里叶系数进行采样，构造一个未归一化的密度，并采用适当的方差衰减，以控制满足 $\nu \in H^{\gamma}$的最大 γ 的取值。然后对该

[4]用于复现实验的代码可在以下网址找到：https://github.com/lchizat/2026-WGF-KMD .



（a）源与目标正则性的变化　　（b）源与目标最小值的变化　　（c）孔洞的指数填充

图2：情形 $s=1$ 与 $d=1$ 采用迎风格式的有限体积离散进行积分。（a）对于 $s=1$，$\bar{\mu}$ 与 ν 的正则性不影响指数收敛率。此处 γ_0（相应地 γ_ν）是使得 $\bar{\mu}\in H^{\gamma_0}$（相应地 $\nu\in H^{\gamma_\nu}$）成立的最大标量，且各密度以0.2为下界，因此我们的理论速率为 $O(e^{-0.2t})$。（b）正如理论所预测，收敛速率以 ν的最小密度为上界，且与 $\bar{\mu}$的最小密度无关。此处 m_0（相应地 m_ν）表示$\bar{\mu}$（相应地 ν）的最小值。在本实验中，（γ_0，γ_ν）=（1，1）。（c）当 ν具有正下界时，$\bar{\mu}$ 的零密度区域（黑色）以指数速度迅速收缩，正如理论所预测（此处（γ_0，γ_ν）=（1，1））；参见定理1.3和定理 3.2。

函数进行平移和缩放，得到最小值为 0 的概率密度。最后，我们将其与均匀分布混合，以控制该密度的最小值。在设置 (C) 中，当 ν 需要离散化时，我们使用由等间距分位数处的相同狄拉克质量构成的离散测度。

设定 (A) 的结果如图 2 所示。所有观察结果均与定理 1.2 的结论一致。特别是，收敛率界（以虚线表示）看起来相当紧。在 ν 存在 0 密度区域而 $\bar{\mu}$ 不存在（未显示）的情形下，我们观察到 $\min\mu_t=\Theta(1/t)$，并且能量的指数收敛率丧失，这表明 ν 上的下界假设对于指数收敛是必要的（见定理1.3）。

设定 (B) 的结果如图 3 所示。尽管我们的保证是局部的，但我们在所有实验中都观察到了全局收敛，这表明反例可能难以找到。关于 $\dot{H}^{-s}$ 距离 $O\left(t^{-\frac{\min(\gamma_0,\gamma_\nu-s)+s}{2\,s-1}}()\right)$，定理 1.4 中的收敛速率至少在情形 $\gamma_\nu-s\geq\gamma_0$下看起来是紧的。

设置 (C) 的结果如图 4 所示。这里我们再次观察到全局收敛，尽管我们的保证仅是局部的；定理 $O\left(t^{-\frac{\min(\gamma_0,\gamma_\nu-1)+s}{s-1}}()\right)$ 在 $s=(d+3)/2$ 条件下给出的能量收敛速率 1.7 看起来略微保守（不过，当 d 较大时，该速率在渐近意义下是尖锐的；此时尖锐速率由定理 4.8 对所有 d给出）。我们还观察到，Wasserstein (W) 和 Wasserstein–Fisher–Rao (WFR) 梯度流的表现相近；这是因为这里将 ν 选为概率测度。如果 ν 不是概率测度，而是质量不同于 1 的非负测度，那么 (W) 不可能收敛到极小化子，而 (WFR) 则可以（并且局部上确实会收敛）。

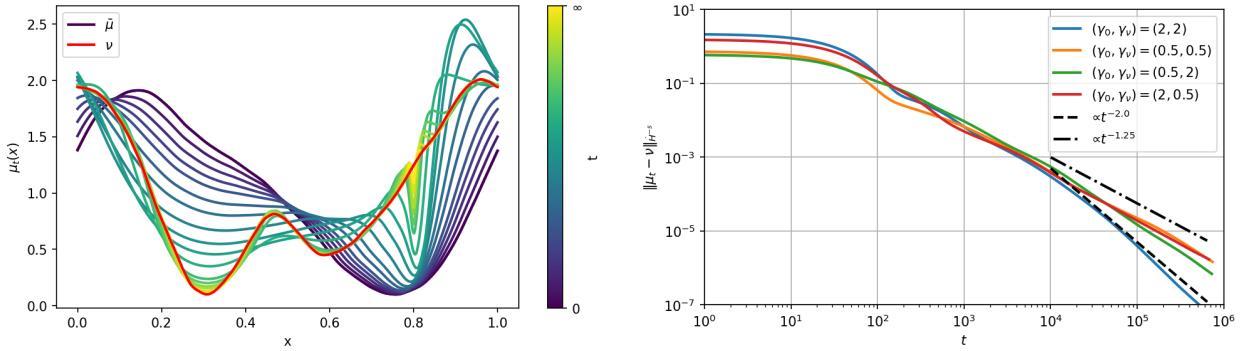
1.5 证明思路

Let 下面介绍主要结果的证明思路，即定理 1.2 和 1.　　4.

为了获得定量的能量衰减率，我们沿着演化 μ_t寻找一个局部 Łojasiewicz 梯度不等式，即

$$\int_{\mathbb{T}^d}|\nabla K_s*(\mu_t-\nu)|^2\mu_t\geq c\|\mu_t-\nu\|_{\dot{H}^{-s}}^{2\beta}\qquad\forall t\geq 0,\quad c>0,\,\beta\geq 1. \tag{1.21}$$

事实上，(1.6) 与 (1.21) 结合起来可得：当 $\|\mu_t-\nu\|_{\dot{H}^{-s}}^2$ 时，呈指数衰减 $\sim e^{-ct}$；若 $\beta=1$，则呈多项式衰减 $\sim(1+Ct)^{-1/(\beta-1)}$。若 $\beta>1$。我们指出，对于我们的能量泛函，这类 Łojasiewicz 梯度不等式并非在整个测度空间上都成立；正如我们将要看到的，它们只在该空间的某个适当子区域内成立。证明的主要难点在于说明，在对



（a）密度的演化　　（b）变化的源与目标下界

图 3：使用带有迎风格式的有限体积离散对 $s = 2$ 和 $d = 1$ 进行积分的情形。我们将 γ_0 （分别为 γ_ν）记为满足 $\bar{\mu} \in H^{\gamma_0}$ （分别为 $\nu \in H^{\gamma_\nu}$）且密度具有下界 0.2 的最大标量。（a）沿着 μ_t 的演化过程中的密度快照（此处 (γ_0, γ_ν) = （2,2））。我们可以观察到最大值原理的缺失（参见 $x = 0.9$）以及高频分量的出现（参见 $x = 0.8$）。在 $s = 1$的情形中，这两种现象均不存在。（b） $\dot{H}^{-2}$ 范数中的收敛速率。虚线表示渐近速率的近似拟合。

初始测度与目标测度作出适当假设后，梯度流始终被限制在这样一个良好区域内。这需要能量耗散恒等式 (1.6) 与针对高阶导数的一些精细能量估计之间进行非平凡的相互配合（见定理 2.13）。

当 $s = 1$ 时，(1.21) 在 $\beta = 1$ 且 $c = \alpha > 0$ 的条件下成立，这得益于最大值原理 $\inf \mu_t \geq \alpha_0$。然而，在 $s > 1$ 的情形下，同样简单的估计因两个原因而行不通。首先，最大值原理不适用于 $s > 1$；其次，即便我们有 $\inf \mu_t \geq c > 0$，(1.21) 右侧得到的量反而会涉及低阶范数 $\|\mu_t - \nu\|^2_{\dot{H}^{1-2s}}$。这些障碍使得对 $s > 1$ 情形的分析与 $s = 1$ 情形大不相同。

因此，若要沿着流寻找类似于 (1.21) 的不等式，就不得不假设 H^γ 具有更高阶正则性， $\mu_t - \nu$ 取充分大的 γ，并按如下方式插值齐次范数：

$$\|\mu_t - \nu\|_{\dot{H}^{-s}} \leq \|\mu_t - \nu\|^{1-\theta}_{\dot{H}^{1-2s}} \|\mu_t - \nu\|^{\theta}_{\dot{H}^{\gamma}}, \qquad \theta = \frac{s-1}{\gamma + 2s - 1}. \tag{1.22}$$

由 (1.22) 可得

$$\int_{\mathbb{T}^d} |\nabla K_s * (\mu_t - \nu)|^2 \mu_t \geq \inf \mu_t \left(\|\mu_t - \nu\|^2_{\dot{H}^\gamma}\right)^{1-\beta} \left(\|\mu_t - \nu\|^2_{\dot{H}^{-s}}\right)^{\beta}, \qquad \beta = 1 + \frac{s-1}{\gamma + s}.$$

因此，为了得到（1.21），进而得到 $\|\mu_t - \nu\|^2_{\dot{H}^{-s}}$ 以指数 $-(\gamma + s)/(s-1)$ 的多项式衰减，我们需要确保对于所有 $t \geq 0$，$\inf \mu_t$ 和 $\|\mu_t - \nu\|_{\dot{H}^\gamma}$ 分别始终有下界和上界。这时，对高阶范数 $\|\mu_t - \nu\|^2_{\dot{H}^\gamma}$ 的能量估计便发挥了关键作用，同时还需满足初始偏差足够小这一假设。

为便于阐述，我们通过考察扰动 $\sigma_t := \mu_t - \nu$ 的线性化方程 $\partial_t \sigma_t = \operatorname{div}(\nu \nabla (-\Delta)^{-s} \sigma_t)$，来说明 Łojasiewicz 梯度不等式随时间传播的主要机制。 $\|\sigma_t\|^2_{\dot{H}^\gamma}$ 的演化由以下公式决定

$$\frac{d}{dt}\|\sigma_t\|^2_{\dot{H}^\gamma} = -2 \int_{\mathbb{T}^d} \nu \nabla (-\Delta)^{-s} \sigma_t \cdot \nabla (-\Delta)^{\gamma} \sigma_t.$$

在最简单的情形下 $\nu = 1$，我们只需对拉普拉斯算子的某个幂$(\gamma + s)/2$ 分部积分，即可得出右侧等于 $-2\|\sigma_t\|^2_{\dot{H}^{\gamma - s + 1}}$。当 ν 不是常数时，我们也采用类似的论证，但必须处理导数作用于 ν所产生的误差项。得益于 Kato–Ponce 交换子估计（我们在第A.2节将其推广到周期情形），这些误差项可以

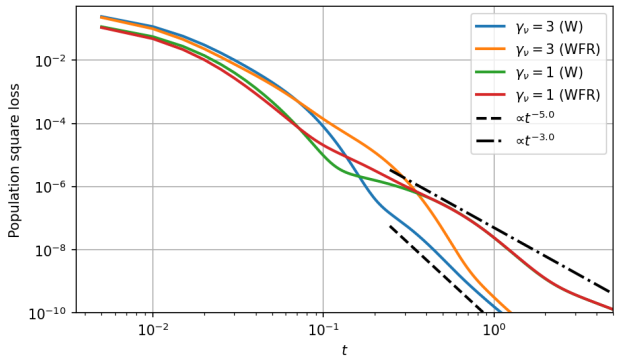


图 4：浅层 ReLU 神经网络的情形，其中 $d = 1$，通过对总体损失进行梯度下降实现，使用宽度为 800 的“学生”和“教师”神经网络分别对测度 μ_t和 ν 进行离散化（换言之，这是每个测度包含800个粒子的偏微分方程相互作用粒子系统逼近）。我们将 $\bar{\mu}$ 初始化为均匀密度（$\gamma_0 = +\infty$），并改变目标正则性 $\dot{}$（我们标出满足 γ_ν 的最大 $\nu \in H^{\gamma_\nu}$）。我们比较 Wasserstein（W）和 Wasserstein–Fisher–Rao（WFR）动力学。尽管（WFR）的能量衰减在初始化时稍快（得益于能量耗散公式（1.18）中的额外项），但总体而言其收敛速度并不更快。虚线表示渐近速率的近似拟合。

对照 由 σ_t 的低阶范数控制，前提是对 ν 假设一定的高阶正则性 ：

$$\frac{d}{dt}\|\sigma_t\|_{\dot{H}^\gamma}^2 \le -2\big(\min_{\mathbb{T}^d}\nu\big)\|\sigma_t\|_{\dot{H}^{\gamma-s+1}}^2 + C\|\nu\|_{H^{\gamma+s}}\|\sigma_t\|_{\dot{H}^{\gamma-s+1}}\|\sigma_t\|_{\dot{H}^{\gamma-s}}.$$

假设 $\nu > 0$，在齐次范数之间插值，并使用 Young 不等式，我们可以吸收右侧误差项中具有危险性的部分。再考虑$\|\sigma_t\|_{\dot{H}^{-s}}^2$的耗散恒等式，最终得到以下微分不等式系统:

$$\frac{d}{dt}\|\sigma_t\|_{\dot{H}^\gamma}^2 \le C\|\sigma_t\|_{\dot{H}^{-s}}^2, \qquad \frac{d}{dt}\|\sigma_t\|_{\dot{H}^{-s}}^2 \le -2\min_{\mathbb{T}^d}\nu\left(\|\sigma_t\|_{\dot{H}^\gamma}^2\right)^{1-\beta}\left(\|\sigma_t\|_{\dot{H}^{-s}}^2\right)^{\beta}, \qquad \beta = 1 + \frac{s-1}{\gamma+s}. \quad (1.23)$$

当 γ 足够大时，(1.23给出了 $\|\sigma_t\|_{\dot{H}^\gamma}^2$ 关于时间的一致有界性，以及指数为 $-(\gamma+s)/(s-1)$ 的多项式衰减，对于 $\|\sigma_t\|_{\dot{H}^{-s}}^2$，从而得出线性化情形下所需的收敛结果。最后，为了控制非线性效应，我们需要假设初始扰动 $\sigma_0 = \bar{\mu} - \nu$足够小，由此得到定理1.4中的局部性。

2 适定性理论

本节旨在证明定理1.1。在固定记号并于第1.4节引入方程 (2.1) 的解的概念后，我们将在第2.2 节证明方程 (1.4) 在空间 $\mathscr{X}_s(\mathbb{T}^d)$中的极大解的存在性与唯一性，并给出延拓判据。最后，在第2.3 节中，我们证明 Hölder 正则性和索博列夫正则性会从数据传播至整个极大存在区间。

2.1 预备知识

本节介绍全文中与方程 (1.4) 的解相关的记号和基本定义。在第 2.1.1 节中，我们回顾环面上的齐次索оболев空间和 Riesz 核。在第 2.1.2 节中，我们定义 Riesz 核均值差异泛函 $\mathscr{E}_s^\nu$ 以及概率测度空间上的 Wasserstein 梯度流相关概念。最后，在第 2.1.3 节中，我们介绍主动标量方程 (1.4) 的弱解，并证明只要它们属于局部适定性类 $\mathscr{X}_s(\mathbb{T}^d)$，就与 $\mathscr{E}_s^\nu$ 的 Wasserstein 梯度流一致。

2.1.1 $\mathbb{T}^d$上的索np列夫空间与Riesz核

设 $d \geq 1$ 为整数，并令 $\mathbb{T}^d \cong \mathbb{R}^d/\mathbb{Z}^d$ 表示 d 维环面。我们考虑 $L^2(\mathbb{T}^d)$ 的标准傅里叶标准正交基，记为 $\{e^{2\pi ik\cdot x}\}_{k\in\mathbb{Z}^d}$。对于任意周期分布 $f \in \mathscr{D}'(\mathbb{T}^d)$ 和任意 $k \in \mathbb{Z}^d$，我们用 $\hat{} f_k = \langle f, e^{2\pi ik\cdot x}\rangle \in \mathbb{C}$ 表示 f 的第 k 个傅里叶系数，其中 $\langle\cdot,\cdot\rangle$ 是对偶配对；我们用记号 $\sum_{k\in\mathbb{Z}^d} \hat{} f_k e^{2\pi ik\cdot x}$ 表示 f 的傅里叶展开。

对于 $\gamma \in \mathbb{R}$， f 的齐次 γ-索об列夫半范数定义为

$$\|f\|_{\dot{H}^\gamma(\mathbb{T}^d)} := \left(\sum_{k\in\mathbb{Z}^d\setminus\{0\}} (2\pi|k|)^{2\gamma}|\hat{f}_k|^2\right)^{1/2} \qquad \forall f \in \mathscr{D}'(\mathbb{T}^d), \quad \forall \gamma \in \mathbb{R}.$$

齐次索об列夫空间 $\dot{H}^\gamma(\mathbb{T}^d)$由满足$\|\cdot\|_{\dot{H}^\gamma}$ 有限的零均值分布构成。当$\gamma \geq 0$时，我们还考虑由下式定义的（非齐次）索об列夫空间 $H^\gamma(\mathbb{T}^d) \subseteq L^2(\mathbb{T}^d)$

$$H^\gamma(\mathbb{T}^d) := \left\{f \in L^2(\mathbb{T}^d) : \|f\|_{H^\gamma(\mathbb{T}^d)} := \|f\|_{L^2(\mathbb{T}^d)} + \|f\|_{\dot{H}^\gamma(\mathbb{T}^d)} < \infty\right\} \qquad \forall \gamma \geq 0.$$

对于 $s > 0$，分数阶拉普拉斯算子$(-\Delta)^s$ 是定义在分布上的傅里叶乘子，定义为

$$(-\Delta)^s f := \sum_{k\in\mathbb{Z}^d} (2\pi|k|)^{2s}\hat{f}_k e^{2\pi ik\cdot x} \qquad \forall f \in \mathscr{D}'(\mathbb{T}^d), \quad \forall s > 0.$$

$(-\Delta)^s$ 的零平均值基本解（即在$\mathbb{T}^d$中满足$(-\Delta)^s K_s = \delta_0 - 1$ 的分布解，其中 δ_0 表示位于0处的狄拉克δ函数）是Riesz核 $K_s \in \mathscr{D}'(\mathbb{T}^d)$，由以下傅里叶展开定义

$$K_s(x) := \sum_{k\in\mathbb{Z}^d\setminus\{0\}} (2\pi|k|)^{-2s}e^{2\pi ik\cdot x} \qquad \forall s > 0. \tag{2.1}$$

因此，对于每个 $f \in \dot{H}^{-s}(\mathbb{T}^d)$，$(-\Delta)^s u = f$ 在 $\mathbb{T}^d$ 中的零均值解为 $u = (-\Delta)^{-s} f := K_s * f \in \dot{H}^s(\mathbb{T}^d)$，其中卷积 $*$ 按分布的意义理解。特别地，齐次 Sobolev 范数 $\|\cdot\|_{\dot{H}^{-s}(\mathbb{T}^d)}$ 可写为

$$\|f\|^2_{\dot{H}^{-s}(\mathbb{T}^d)} = \langle f, K_s * f\rangle \qquad \forall f \in \dot{H}^{-s}(\mathbb{T}^d). \tag{2.2}$$

关于 A.1节中 K_s的精确正则性性质，请参阅该节。

2.1.2 概率测度瓦瑟斯坦空间中的黎曼能量梯度流

我们回顾概率测度空间中的分析理论的一些概念，并给出与负阶索оболев范数相关的核均值差异（KMD）的梯度流的精确定义。更多详情请参阅 [AGS08] 。

我们将 球面上概率测度的空间 $\mathscr{P}(\mathbb{T}^d)$ 配备 2-瓦瑟斯坦距离 e

$$W_2(\mu_1,\mu_2) := \left(\min_{\gamma\in\Gamma(\mu_1,\mu_2)} \int_{\mathbb{T}^d\times\mathbb{T}^d} |x-y|^2 d\gamma(x,y)\right)^{1/2}, \qquad \mu_1,\mu_2 \in \mathscr{P}(\mathbb{T}^d),$$

其中 $\Gamma(\mu_1, \mu_2) \subset \mathscr{P}(\mathbb{T}^d \times \mathbb{T}^d)$ 是 μ_1与 μ_2之间耦合的集合。众所周知，$(\mathscr{P}(\mathbb{T}^d)\ W_2)$ 是一个紧致度量空间，其收敛概念与测度的弱-$*$ 收敛一致。

给定一个区间 $I \subseteq \mathbb{R}$，我们称概率测度曲线 $\mu : I \to \mathscr{P}(\mathbb{T}^d)$ 是绝对连续的，并记作 $\mu \in AC(I; \mathscr{P}(\mathbb{T}^d))$，如果存在某个 $g \in L^1(I)$ 使得

$$W_2(\mu_b,\mu_a) \leq \int_a^b g(t)dt \qquad \forall a,b \in I,\ a < b. \tag{2.3}$$

在这种情况下，度量导数是良定义的

$$|\mu_t'| := \lim_{r\to t} \frac{W_2(\mu_r,\mu_t)}{|r-t|} < \infty \qquad \text{for a.e. } t \in I,$$

并且 $g(t) = |\mu_t'|$ 是使上述条件（2.3）成立的最小可能选择。

绝对连续曲线满足以下刻画： $\mu : I \to \mathscr{P}(\mathbb{T}^d)$ 绝对连续，当且仅当它在弱-$*$ 拓扑下连续，且存在一个博雷尔向量场$v_t(x)$: $I \times \mathbb{T}^d \to \mathbb{R}^d$，使得对于几乎处处的 $t \in I$，有 $v_t \in L^2(\mathbb{T}^d, \mu_t; \mathbb{R}^d)$，并且连续性方程

$$\partial_t \mu + \operatorname{div}(\mu v) = 0 \qquad \text{in } I \times \mathbb{T}^d$$

在分布意义下成立，等价地，

$$\int_{\mathbb{T}^d} \varphi d\mu_b = \int_{\mathbb{T}^d} \varphi d\mu_a + \int_a^b \int_{\mathbb{T}^d} \nabla\varphi \cdot v_r d\mu_r dr \qquad \forall a, b \in I,\ a < b, \qquad \forall \varphi \in C^\infty(\mathbb{T}^d).$$

在这种情况下，存在唯一的 v，称为 μ 的切向量场，使得

$$\|v_t\|_{L^2(\mathbb{T}^d, \mu_t; \mathbb{R}^d)} = |\mu_t'| \qquad \text{for a.e. } t \in I.$$

现在引入本文所研究的定义在 $\mathscr{P}(\mathbb{T}^d)$ 上的能量泛函类。对于每个 $s \geq 1$ 以及任意给定的概率测度 $\nu \in \mathscr{P}(\mathbb{T}^d)$，我们定义 $\mathscr{E}_s^\nu$: $\mathscr{P}(\mathbb{T}^d) \to [0, \infty]$ 为

$$\mathscr{E}_s^\nu(\mu) := \frac{1}{2}\|\mu - \nu\|_{\dot{H}^{-s}}^2 = \frac{1}{2}\left(\sup_{\substack{f \in C^\infty(\mathbb{T}^d) \\ \|f\|_{\dot{H}^s} \leq 1}} \int_{\mathbb{T}^d} f(d\mu - d\nu) \right)^2,$$

. . 其中，最后一个等号由对偶性 $H^s - H^{-s}$ 以及光滑函数在 H^s 中的稠密性得到。泛函 $\mathscr{E}_s^\nu$ 是适当的（$\mathscr{E}_s^\nu(\nu) = 0$），并且关于 Wasserstein 度量下半连续。事实上，如上所见，它可以写成一族连续函数的上确界。

对于每个 $\mu \in \{\mathscr{E}_s^\nu < \infty\}$，$\mathscr{E}_s^\nu$在 μ处的次微分是集合 $\partial\mathscr{E}_s^\nu(\mu)$，由所有满足下述条件的向量场 $\xi \in L^2(\mathbb{T}^d, \mu; \mathbb{R}^d)$ 构成:

$$\mathscr{E}_s^\nu(\eta) - \mathscr{E}_s^\nu(\mu) \geq \inf_{\gamma \in \Gamma_{\text{opt}}(\mu,\eta)} \int_{\mathbb{T}^d \times \mathbb{T}^d} \xi(x) \cdot (y - x) d\gamma(x, y) + o\left(W_2(\mu, \eta)\right) \qquad \forall \eta \in \mathscr{P}(\mathbb{T}^d), \tag{2.4}$$

其中 $\Gamma_{\text{opt}}(\mu, \eta) \subset \mathscr{P}(\mathbb{T}^d \times \mathbb{T}^d)$ 是 μ 与 η之间的最优耦合集合。这对应于 [AGS08，公式 10.3 中的约化次微分概念。12]。接下来，我们给出能量 $\mathscr{E}_s^\nu$的梯度流定义:

定义 2.1。若 $\mu \in AC([0,T); \mathscr{P}(\mathbb{T}^d))$ 是从 $\bar{\mu} \in \mathscr{P}(\mathbb{T}^d)$ 出发的 $\mathscr{E}_s^\nu$的梯度流，且$\mu_0 = \bar{\mu}$，并且 μ 的切向量场 v 满足

$$-v_t \in \partial\mathscr{E}_s^\nu(\mu_t) \qquad \textit{for a.e. } t \in (0, T).$$

This 这一概念恰好对应于 [AGS08，定义 11.1。1] 应用于泛函 $\mathscr{E}_s^\nu$.

2.1.3 主动标量方程

我们引入本文所考虑的 (1.4) 的弱解概念。对于任意区间 $I \subseteq \mathbb{R}$以及任意概率测度曲线 $I \ni t \mapsto \mu_t \in \mathscr{P}(\mathbb{T}^d)$，如果 $\mu \in C_{w^*}(I; \mathscr{P}(\mathbb{T}^d))$ 满足 μ 关于弱 $*$ 拓扑随时间连续，则记作该情形。

定义 2.2. 设 $\bar{\mu}, \nu \in \mathscr{P}(\mathbb{T}^d)$ 且 $s \in [1, \infty)$。若概率测度曲线 $\mu \in C_{w^*}([0,T); \mathscr{P}(\mathbb{T}^d))$ 满足主动标量方程 *(1.4)*，即 $\mu_0 = \bar{\mu}$, $v_t := -\nabla K_s * (\mu_t - \nu)$ 满足 $v \in L^1_{\text{loc}}((0,T) \times \mathbb{T}^d, \mu; \mathbb{R}^d)$，且连续性方程 $\partial_t \mu + \operatorname{div}(\mu v) = 0$ 在分布意义下成立，等价地

$$\int_{\mathbb{T}^d} \varphi d\mu_t = \int_{\mathbb{T}^d} \varphi d\bar{\mu} + \int_0^t \int_{\mathbb{T}^d} \nabla\varphi \cdot v_r d\mu_r dr \qquad \forall t \in (0, T), \quad \forall \varphi \in C^\infty(\mathbb{T}^d).$$

对于任意 $s \geq 1$，(1.4) 的局部适定性所处的函数空间见 (1.5)。回顾一下，对于任意 $p \in [1,\infty)$ 和 $q \in [1,\infty]$，洛伦兹空间 $L^{p,q}(\mathbb{T}^d)$ 定义为满足以下条件的函数 $f : \mathbb{T}^d \to \mathbb{R}$ 所组成的集合

$$\|f\|_{L^{p,q}(\mathbb{T}^d)} := p^{1/q} \|\lambda \mathscr{L}^d(\{|f| > \lambda\})^{1/p}\|_{L^q((0,\infty), d\lambda/\lambda)} < \infty. \tag{2.5}$$

以下是洛伦兹空间的一些已知性质： $L^{p,p} = L^p$ 和 $\|f\|_{L^{p,p}} = \|f\|_{L^p}$； $L^{p,q} \subset L^{p,r}$ 若 $q \leq r$ 且 $\|f\|_{L^{p,r}} \lesssim_{p,q,r} \|f\|_{L^{p,q}}$； $L^r \subset L^{p,q}$ 对每个 $r > p$ 均成立。在后文中，我们将根据定理 2.2 考虑方程 (1.4) 的解，这些解在时间上局部有界，且属于空间 $\mathscr{X}_s(\mathbb{T}^d)$，即 $\mu \in L^\infty_{\mathrm{loc}}([0,T); \mathscr{X}_s(\mathbb{T}^d))$。当 $s \geq d/2+1$ 时，这不会带来任何额外条件，因为根据定义， μ_t 在每个时刻都是概率测度。我们通常考虑延拓至极大存在时间的解：

定义 2.3。设 $s \in [1,\infty)$, $\bar{\mu}, \nu \in \mathscr{P} \cap \mathscr{X}_s(\mathbb{T}^d)$，且 $\mu \in L^\infty_{\mathrm{loc}}([0,T); \mathscr{X}_s(\mathbb{T}^d))$ 是根据定理 2.2 得到的方程 (1.4) 的一个解。若满足以下条件，则称 μ 是空间 $\mathscr{X}_s(\mathbb{T}^d)$ 中方程 (1.4) 的极大解：设 $\tilde{T} \geq T$ 和 $\tilde{\mu} \in L^\infty_{\mathrm{loc}}([0,\tilde{T}); \mathscr{X}_s(\mathbb{T}^d))$ 是方程 (1.4) 的一个解，且对所有 $t \in [0,T)$ 都有 $\tilde{\mu}_t = \mu_t$。˜ 那么 $T = T$ 且 $\tilde{\mu} = \mu$。

现在我们证明，属于 $L^\infty_{\mathrm{loc}}([0,T); \mathscr{X}_s(\mathbb{T}^d))$ 这一类的方程 (1.4的解是 $\mathscr{E}^\nu_s$ 的 Wasserstein 梯度流，这由定理 2.1得出。我们先给出以下引理，它保证由主动标量方程 (1.4)规定的向量场 v_t 属于 $\mathscr{E}^\nu_s$ 在 μ_t 处的次微分，只要 $\mu_t \in \mathscr{X}_s(\mathbb{T}^d)$。

引理 2.4。 设 $s \geq 1$，且 $\mu, \nu \in \mathscr{P} \cap \mathscr{X}_s(\mathbb{T}^d)$。则， $\nabla K_s * (\mu - \nu) \in \partial \mathscr{E}^\nu_s(\mu)$。

证明。 记 $\xi := \nabla K_s * (\mu - \nu)$。由定理 A.2 和定理 A.4， $\xi \in C(\mathbb{T}^d; \mathbb{R}^d) \subset L^2(\mathbb{T}^d, \mu_t; \mathbb{R}^d)$ 具有连续模

$$|\xi(x) - \xi(y)| \lesssim_{d,s,M} |x-y| \log(2 + |x-y|^{-1}), \qquad M := \max\{\|\mu\|_{\mathscr{X}_s}, \|\nu\|_{\mathscr{X}_s}\}.$$

特别地，由中值定理可得

$$|K_s * (\mu-\nu)(y) - K_s * (\mu-\nu)(x) - \xi(x) \cdot (y-x)| \lesssim_{d,\ s,\ M} |x-y|^2 \log(2 + |x-y|^{-1}) \qquad \forall x, y \in \mathbb{T}^d. \tag{2.6}$$

我们来验证条件（2.4）成立。取 $\eta \in \mathscr{P}(\mathbb{T}^d)$ 使得 $\mathscr{E}^\nu_s(\eta) < \infty$ 且 $\gamma \in \Gamma_{\mathrm{opt}}(\mu, \eta)$。然后，利用（2.2）可得

$$\begin{aligned}
\mathscr{E}^\nu_s(\eta) - \mathscr{E}^\nu_s(\mu) &= \frac{1}{2}\|\eta - \nu\|^2_{\dot{H}^{-s}} - \frac{1}{2}\|\mu - \nu\|^2_{\dot{H}^{-s}} = \langle K_s * (\mu - \nu), \eta - \mu \rangle + \frac{1}{2}\|\eta - \mu\|^2_{\dot{H}^{-s}} \\
&\geq \langle K_s * (\mu - \nu), \eta - \mu \rangle = \int_{\mathbb{T}^d \times \mathbb{T}^d} (K_s * (\mu-\nu)(y) - K_s * (\mu-\nu)(x))\, d\gamma(x,y) \\
&\geq \int_{\mathbb{T}^d \times \mathbb{T}^d} \xi(x) \cdot (y-x) d\gamma(x,y) - C(d,s,M) \int_{\mathbb{T}^d \times \mathbb{T}^d} |x-y|^2 \log(2 + |x-y|^{-1}) d\gamma(x,y) \\
&\geq \int_{\mathbb{T}^d \times \mathbb{T}^d} \xi(x) \cdot (y-x) d\gamma(x,y) - C(d,s,M) W_2^2(\mu,\eta) \log\left(2 + W_2^{-1}(\mu,\eta)\right),
\end{aligned}$$

其中在最后两步中，我们使用了（2.6）以及凹函数 $t \mapsto t\log(2+t^{-1/2})$ 的詹森不等式。由于 $\gamma \in \Gamma_{\mathrm{opt}}(\mu, \eta)$ 的任意性，我们得到（2.4）。 □

命题 2.5。设 $s \geq 1$, $\bar{\mu}, \nu \in \mathscr{P} \cap \mathscr{X}_s(\mathbb{T}^d)$，以及 $\mu \in L^\infty_{\mathrm{loc}}([0,T); \mathscr{X}_s(\mathbb{T}^d))$ 是根据定理 2.2 得到的（1.4）的一个解。那么， μ 是根据定理 2.1、以初始数据 $\bar{\mu}$ 为初值的 $\mathscr{E}^\nu_s$ 梯度流。此外， $\mu \in Lip_{\mathrm{loc}}([0,T); (\mathscr{P}(\mathbb{T}^d) W_2))$，且能量耗散恒等式（1.6）成立。

证明。令 $v_t = -\nabla K_s * (\mu_t - \nu)$ 为解 μ 生成的向量场。根据定理 2.4， $v_t \in -\partial \mathscr{E}^\nu_s(\mu_t) \subset L^2(\mathbb{T}^d, \mu_t; \mathbb{R}^d)$ 对所有 $t \in [0,T)$ 均成立，因此 $\mu \in AC([0,T); \mathscr{P}(\mathbb{T}^d))$，并且根据定理 2.1，它是从 $\bar{\mu}$ 出发的 $\mathscr{E}^\nu_s$ 梯度流。此外，由于 $\mu \in L^\infty_{\mathrm{loc}}([0,T); \mathscr{X}_s(\mathbb{T}^d))$，根据定理 A.4 v，它是局部

有界于 $[0,T)\times\mathbb{T}^d$，这意味着 $\mu\in \mathrm{Lip}_{\mathrm{loc}}\left([0,T);\left(\mathscr{P}(\mathbb{T}^d)\,W_2\right)\right)$。为证明（1.6），取 $t\in(0,T)$，并利用（2.2）和方程（1.4）计算：

$$\begin{aligned}\frac{\mathscr{E}_s^\nu(\mu_{t+h})-\mathscr{E}_s^\nu(\mu_t)}{h}&=\left\langle K_s*(\mu_t-\nu)+\frac{1}{2}K_s*(\mu_{t+h}-\mu_t),\frac{\mu_{t+h}-\mu_t}{h}\right\rangle\\&=\fint_t^{t+h}\left\langle K_s*(\mu_t-\nu)+\frac{1}{2}K_s*(\mu_{t+h}-\mu_t),\partial_r\mu_r\right\rangle dr\\&=\fint_t^{t+h}\left\langle K_s*(\mu_t-\nu)+\frac{1}{2}K_s*(\mu_{t+h}-\mu_t),\mathrm{div}(\mu_r\nabla K_s*(\mu_r-\nu))\right\rangle dr\\&=-\fint_t^{t+h}\left\langle \nabla K_s*(\mu_t-\nu)+\frac{1}{2}\nabla K_s*(\mu_{t+h}-\mu_t),\mu_r\nabla K_s*(\mu_r-\nu)\right\rangle dr.\end{aligned}$$

现在，根据定理 A.4，由于 μ 关于时间弱 $*$ 连续，$\nabla K_s*(\mu_r-\nu)\to\nabla K_s*(\mu_t-\nu)$ 在 $r\to t$ 趋于 $\nabla K_s*(\mu_{t+h}-\mu_t)\to 0$ 在 $h\to 0$时一致收敛。因此，我们得到

$$\frac{\mathscr{E}_s^\nu(\mu_{t+h})-\mathscr{E}_s^\nu(\mu_t)}{h}=-\int_{\mathbb{T}^d}|\nabla K_s*(\mu_t-\nu)|^2d\mu_t+o(h)\qquad\text{as } h\to 0,$$

由此可得（1.6）。 □

2.2 弱类中的局部适定性

本节证明方程（1.4）在类 $\mathscr{X}_s(\mathbb{T}^d)$中的（极大）解的存在性与唯一性。具体而言，我们在第 2.2.1 节为 $\mathscr{X}_s(\mathbb{T}^d)$中的解建立定量稳定性估计，由此特别可得唯一性。随后在第 2.2.2节证明存在性。最后，第 2.2.3 节致力于构造极大解并证明延拓判据。

2.2.1 唯一性与稳定性

在下述命题中，我们证明，对于数据的变化，在有界时间区间上，解关于 Wasserstein 距离一致稳定。稳定性不等式的精确偏离速率（根据 s 的选择，可以是指数或双指数）见定理 2.7。

命题 2.6。（唯一性与稳定性）设 $s\geq 1$，并设 μ^1，$\mu^2\in L^\infty([0,T)$；$\mathscr{X}_s(\mathbb{T}^d))$ 是方程（1.4）的两个解，其初始测度与目标测度分别为 $\bar\mu^1$，$\bar\mu^2$，ν^1，$\nu^2\in\mathscr{P}\cap\mathscr{X}_s(\mathbb{T}^d)$。令

$$M:=\max\left\{\sup_{t\in[0,T)}\|\mu_t^1\|_{\mathscr{X}_s},\sup_{t\in[0,T)}\|\mu_t^2\|_{\mathscr{X}_s},\|\nu^1\|_{\mathscr{X}_s},\|\nu^2\|_{\mathscr{X}_s}\right\}.$$

则下列结论成立：

i) (Uniqueness). If $\bar\mu^1=\bar\mu^2$ and $\nu^1=\nu^2$, then $\mu_t^1=\mu_t^2$ for all $t\in[0,T)$.

ii)（稳定性）。否则，下述界成立：

$$W_2(\mu_t^1,\mu_t^2)\leq\Omega(t)\qquad\forall t\in[0,T),$$

其中 $\Omega=\Omega\left(d,s,M,W_2(\bar\mu^1,\bar\mu^2)\ W_2(\nu^1,\nu^2)\right):[0,\infty)\to[0,\infty)$ 是一个连续递增函数，满足 $\Omega(0)=W_2(\bar\mu^1,\bar\mu^2)$。此外，随着$W_2(\bar\mu^1,\bar\mu^2)$ 和 $W_2(\nu^1,\nu^2)$ 趋于零，Ω 在有界区间上一致趋于零。

证明。令 $v_t^1=-\nabla K_s*(\mu_t^1-\nu^1)$ 和 $v_t^2=-\nabla K_s*(\mu_t^2-\nu^2)$ 分别为由两个解 μ^1 和 μ^2生成的向量场。根据定理 A.2，

$$|v_t^i(x)-v_t^i(y)|\lesssim_{d,s}M\omega_s(|x-y|)\qquad\forall x,y\in\mathbb{T}^d,\quad\forall t\in[0,T),\tag{2.7}$$

其中，连续模 ω_s 由（A.4）给出，根据 s的选择，它是利普希茨或对数利普希茨的。根据 ω_s 奥斯古德型的性质[5], 以及 [AB08], 中的结果，这两个解必然是拉格朗日的，也就是说，与向量场 $X^1, X^2\colon [0,T)\times\mathbb{T}^d \to \mathbb{T}^d$相关联的流映射 v^1, v^2是良定义的，并给出如下表示

$$\mu_t^i = (X_t^i)_{\#}\bar{\mu}^i \qquad \forall t\in[0,T), \quad i=1,2.$$

设 $\bar{\gamma}\in\mathscr{P}(\mathbb{T}^d\times\mathbb{T}^d)$ 是初始测度 $\bar{\mu}^1$ 和 $\bar{\mu}^2$之间二次代价最优传输问题的一个最优耦合。我们考虑耦合 $\gamma_t := (X_t^1,\ X_t^2)_{\#}\bar{\gamma}$，其边缘为 μ_t^1 和 μ_{2}^t。我们有

$$W_2(\mu_t^1,\mu_t^2) \le \left(\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}|x-y|^2 d\gamma_t(x,y)\right)^{1/2} = \left(\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}|X_t^1(x)-X_t^2(y)|^2 d\bar{\gamma}(x,y)\right)^{1/2} =: \rho(t).$$

我们将为时间区间 $[0,T)$ 内的量 $\rho(t)$ 推导一个合适的微分不等式，积分后即可得到对 $W_2(\mu_t^1,\ \mu_t^2)$ 的所需控制。由柯西–施瓦茨不等式和 $L^2(\mathbb{T}^d\times\mathbb{T}^d,\ \bar{\gamma};\mathbb{R}^d)$ 中的三角不等式可得

$$\begin{aligned}\rho'(t) &\le \left(\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}|v_t^1(X_t^1(x))-v_t^2(X_t^2(y))|^2 d\bar{\gamma}(x,y)\right)^{1/2}\\ &\le \underbrace{\left(\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}|v_t^1(X_t^1(x))-v_t^2(X_t^1(x))|^2 d\bar{\gamma}(x,y)\right)^{1/2}}_{(I)} + \underbrace{\left(\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}|v_t^2(X_t^1(x))-v_t^2(X_t^2(y))|^2 d\bar{\gamma}(x,y)\right)^{1/2}}_{(II)}.\end{aligned}$$

为了给 (I) 取一个上界，我们首先利用 $\bar{\gamma}$ 上的边缘条件、变量代换以及 $L^2(\mathbb{T}^d,\ \mu_t;\ \mathbb{R}^d)$ 中的三角不等式，得到

$$(I) = \left(\int_{\mathbb{T}^d}|v_t^1-v_t^2|^2 d\mu_t^1\right)^{1/2} \le \left(\int_{\mathbb{T}^d}|\nabla K_s*(\mu_t^1-\mu_t^2)|^2 d\mu_t^1\right)^{1/2} + \left(\int_{\mathbb{T}^d}|\nabla K_s*(\nu^1-\nu^2)|^2 d\mu_t^1\right)^{1/2}.$$

现在，在 $s\in[1,d/2+1)$ 的情形下，我们使用指数为 $p = d/(2s-2)\ \in(1,\ \infty]$ 和$q=p'=d/(d-2s+2)\in[1,\infty)$ 的赫尔德不等式，并结合定理 A.3 中的界，得到

$$\begin{aligned}(I) &\le \|\mu_t^1\|_{L^p}^{1/2}\|\nabla K_s*(\mu_t^1-\mu_t^2)\|_{L^{2q}} + \|\mu_t^1\|_{L^p}^{1/2}\|\nabla K_s*(\nu^1-\nu^2)\|_{L^{2q}}\\ &\lesssim_{d,s} \|\mu_t^1\|_{L^p}^{1/2}\max\{\|\mu_t^1\|_{L^p},\|\mu_t^2\|_{L^p}\}^{1/2}W_2(\mu_t^1,\mu_t^2) + \|\mu_t^1\|_{L^p}^{1/2}\max\{\|\nu^1\|_{L^p},\|\nu^2\|_{L^p}\}^{1/2}W_2(\nu^1,\nu^2)\\ &\lesssim MW_2(\mu_t^1,\mu_t^2) + MW_2(\nu^1,\nu^2).\end{aligned} \tag{2.8}$$

在情形 $s\ge d/2+1$ 下，我们使用定理 A.4 中的一致有界，并结合 μ_t^1 上的单位质量条件，从而得到

$$(I) \lesssim_{d,s} \begin{cases} W_2(\mu_t^1,\mu_t^2)+W_2(\nu^1,\nu^2) & \text{if } s>\frac{d}{2}+1,\\ W_2(\mu_t^1,\mu_t^2)\log\left(2+W_2(\mu_t^1,\mu_t^2)^{-1}\right)+W_2(\nu^1,\nu^2)\log\left(2+W_2(\nu^1,\nu^2)^{-1}\right) & \text{if } s=\frac{d}{2}+1.\end{cases} \tag{2.9}$$

为了估计(II)，我们转而使用向量场 v_t^2 在定理 A.2 中的连续模以及詹森不等式:

$$(II) \lesssim_{d,s} M\left(\int_{\mathbb{T}^d}\int_{\mathbb{T}^d}\omega_s^2(|X_t^1(x)-X_t^2(y)|)d\bar{\gamma}(x,y)\right)^{1/2} \le \begin{cases} M\rho(t) & \text{if } s\notin\{1,d/2+1\},\\ M\rho(t)\log\left(2+\rho(t)^{-1}\right) & \text{if } s\in\{1,d/2+1\}.\end{cases} \tag{2.10}$$

综合(2.8)、(2.9)与(2.10)，并注意到 $W_2(\mu_t^1,\mu_t^2)\le\rho(t)$，我们最终得到

$$\rho'(t) \lesssim_{d,s} \begin{cases} M\rho(t)+MW_2(\nu^1,\nu^2) & \text{if } s\notin\{1,d/2+1\},\\ M\rho(t)\log\left(2+\rho(t)^{-1}\right)+MW_2(\nu^1,\nu^2)\log\left(2+W_2(\nu^1,\nu^2)^{-1}\right) & \text{if } s\in\{1,d/2+1\}.\end{cases} \tag{2.11}$$

通过对该微分不等式积分，即得所需结论（见定理 2.7 中连续模Ω的精确表达式）。

□

[5]一个连续模 $\omega\colon\ [0,\infty)\to[0,\infty)$ 称为奥斯古德型，如果 $\int_0^1 \omega^{-1}(r)dr=\infty$.

注 2.7. 通过对式（2.11）积分，可得到定理 2.6 中函数 Ω的显式表达式。当 $s \notin \{1, d/2+1\}$时，根据格朗沃尔不等式可得

$$W_2(\mu_t^1, \mu_t^2) \le \rho(t) \le W_2(\bar{\mu}^1, \bar{\mu}^2) e^{c(d,s)Mt} + W_2(\nu^1, \nu^2) \left(e^{c(d,s)Mt} - 1 \right).$$

在这种情况下，解之间的接近程度至多以指数速度丧失，这与向量场是一致利普希茨的事实相符。在临界情形 $s \in \{1, d/2+1\}$，精确积分（2.11）会更加复杂。然而，我们可以分析有界区间内 Ω 的行为，其中 $W_2(\bar{\mu}^1, \bar{\mu}^2)$ 和$W_2(\nu^1, \nu^2)$ 较小，并注意到，对于 $\rho(t) + W_2(\nu^1, \nu^2) \le 1/2$,

$$\begin{aligned} \left(\rho(t) + W_2(\nu^1, \nu^2)\right)' &\lesssim_{d,s} M\rho(t) \log\left(2 + \rho(t)^{-1}\right) + MW_2(\nu^1, \nu^2) \log\left(2 + W_2(\nu^1, \nu^2)^{-1}\right) \\ &\lesssim M(\rho(t) + W_2(\nu^1, \nu^2)) \log\left((\rho(t) + W_2(\nu^1, \nu^2))^{-1}\right). \end{aligned}$$

因此,

$$W_2(\mu_t^1, \mu_t^2) \le \rho(t) \le \left(W_2(\bar{\mu}^1, \bar{\mu}^2) + W_2(\nu^1, \nu^2)\right)^{\exp(-c(d,s)Mt)} - W_2(\nu^1, \nu^2)$$

对于所有满足右端有界于 $1/2$ 的时刻 $t \ge 0$ 。在这种情况下，偏离至多呈双指数增长，这与向量场的对数Lipschitz正则性相符。■

2.2.2 解的存在性

接下来，我们在空间 $\mathscr{X}_s(\mathbb{T}^d)$ 中构造（1.4）的解，时间区间取较短长度。证明通过在拉格朗日流层面进行皮卡迭代完成。

命题 2.8（存在性）。设 $s \ge 1$ 和 $\bar{\mu}$, $\nu \in \mathscr{P} \cap \mathscr{X}_s$（$\mathbb{T}^d$）。则存在 $T > 0$ 以及一个解$\mu \in L^\infty$（$[0,T)$； $\mathscr{X}_s$（$\mathbb{T}^d$）满足方程 （1.4）。

证明。令 $T > 0$为一个待稍后选取的充分小的正数。我们递归地构造一列近似解 $\mu^n \in L^\infty$（$[0,T)$； $\mathscr{X}_s$（$\mathbb{T}^d$）如下。令

$$\mu_t^0 = \bar{\mu} \qquad \forall t \in [0,T).$$

于是，对于每个 $n \ge 0$，分别记 $v_t^n = -\nabla K_s * (\mu_t^n - \nu)$ 和 X^n为与第 n个近似解 μ^n相关联的向量场和流映射，我们定义

$$\mu_t^{n+1} := (X_t^n)_{\#}\bar{\mu} \qquad \forall t \in [0,T).$$

我们将证明，只要取 T足够小， $\{\mu^n\}_{n\ge 0}$在 L^∞（$[0,T)$ $\mathscr{X}_s$（$\mathbb{T}^d$）中一致有界； C_{w^*}（$[0,T)$； $\mathscr{P}$（$\mathbb{T}^d$）在中相对紧，并且该序列的任意弱极限都是方程（1.4）的解。

步骤 1： 我们首先用归纳法证明

$$\|\mu^n\|_{L^\infty([0,T);\mathscr{X}_s)} \le 2\|\bar{\mu}\|_{\mathscr{X}_s} \qquad \forall n \ge 0, \tag{2.12}$$

若 T足够小。注意，当 $s \ge d/2+1$时，这一条件自动满足，因为连续性方程保证总质量守恒。下面讨论 $s \in [1, d/2+1)$的情形。不等式（2.12）在 $n = 0$时显然成立。假设它对某个 $n \ge 0$成立，下面证明它对 $n+1$也成立。

若 $s = 1$，则有 div $v_t^n = \mu_t^n - \nu$，因此对所有 $t \in [0,T)$，以下界成立：

$$\|\mu_t^{n+1}\|_{L^\infty} \le \|\bar{\mu}\|_{L^\infty} \exp\left(\int_0^t \|\operatorname{div} v_r^n\|_{L^\infty} dr \right) \le \|\bar{\mu}\|_{L^\infty} \exp\left(T(2\|\bar{\mu}\|_{L^\infty} + \|\nu\|_{L^\infty})\right) \le 2\|\bar{\mu}\|_{L^\infty},$$

如果 T足够小，这里我们使用了归纳假设。在 $s \in (1, d/2+1)$的情形下，则由定理 A.2， v_t^n具有利普希茨连续模，可改写为

$$\|\nabla v_t^n\|_{L^\infty} \lesssim_{d,s} \|\mu_t^n - \nu\|_{L^{p,1}} \qquad \forall t \in [0,T).$$

特别地，对于所有 $t \in [0,T)$，流映射 X_t^n及其逆映射（X_t^n）$^{-1}$的 Lipschitz常数可由下式界定

$$\max\{\|\nabla (X_t^n)^{-1}\|_{L^\infty}, \|\nabla X_t^n\|_{L^\infty}\} \le \exp\left(\int_0^t \|\nabla v_r^n\|_{L^\infty} dr\right) \le \exp\left(C(d,s)T(\|\bar{\mu}\|_{L^{p,1}} + \|\nu\|_{L^{p,1}})\right) =: L,$$

因此，根 据表示公式 $\mu_t^{n+1} = (\bar{\mu} \det \nabla X_t^n) \circ (X_t^n)^{-1}$，对于所有 $t \in [0,T)$，我们推得

$$\|\mu_t^{n+1}\|_{L^{p,1}} = \left\|(\bar{\mu} \det \nabla X_t^n) \circ (X_t^n)^{-1}\right\|_{L^{p,1}} \le L^d \|\bar{\mu} \circ (X_t^n)^{-1}\|_{L^{p,1}} \le L^{d+1+1/p} \|\bar{\mu}\|_{L^{p,1}} \le 2\|\bar{\mu}\|_{L^{p,1}},$$

只要 T 选取得足够小，这里我们使用了 $L^{p,1}$-拟范数（2.5），以及以下事实:

$$\left|\{|\bar{\mu} \circ (X_t^n)^{-1}| > \lambda\}\right| = \left|(X_t^n)^{-1}(\{|\bar{\mu}| > \lambda\})\right| \le L|\{|\bar{\mu}| > \lambda\}| \qquad \forall \lambda > 0,$$

这由 $(X_t^n)^{-1}$ 上的 Lipschitz界 推出。

步骤 2： 接下来，如有必要，将 T 取得更小，我们证明

$$\lim_{n\to\infty} \sup_{k\ge n} \sup_{t\in[0,T)} W_2(\mu_t^k, \mu_t^{k+1}) = 0. \tag{2.13}$$

令 $M := \max\{\|\bar{\mu}\|_{\mathscr{X}_s}, \|\nu\|_{\mathscr{X}_s}\}$。接着定义量

$$\rho^n(t) := \left(\int_{\mathbb{T}^d} |X_t^{n+1} - X_t^n|^2 d\bar{\mu}\right)^{1/2} \ge W_2(\mu_t^n, \mu_t^{n+1}),$$

论证与.....完全相同证明定理2.6时完全相同，并利用（2.12）中的一致有界性，可得

$$\frac{d}{dt}\rho^n(t) \lesssim_{d,s} \begin{cases} M\rho^n(t) + M\rho^{n-1}(t) & \text{if } s \notin \{1, d/2+1\}, \\ M\rho^n(t)\log\left(2 + \frac{1}{\rho^n(t)}\right) + M\rho^{n-1}(t)\log\left(2 + \frac{1}{\rho^{n-1}(t)}\right) & \text{if } s \in \{1, d/2+1\}. \end{cases}$$

因此，令

$$\tilde{\rho}^n(t) := \sup_{k\ge n} \rho^k(t), \qquad \omega_s(r) := \begin{cases} r & \text{if } s \notin \{1, d/2+1\}, \\ r\log(2 + r^{-1}) & \text{if } s \in \{1, d/2+1\}, \end{cases}$$

我们得到

$$\tilde{\rho}^{n+1}(t) \le C\int_0^t \omega_s(\tilde{\rho}^n(r))dr, \quad C = C(d,s,M) \qquad \forall t \in [0,T), \quad \forall n \ge 0. \tag{2.14}$$

由（2.14）可知，只要 $\sup_{t\in[0,T)} \tilde{\rho}^n(t) \to 0$ 当 $n \to \infty$时， T 取值足够小。首先注意到，存在一个常数 $K > 0$ 使得

$$\omega_s(r) \lesssim r\log(2 + r^{-1}) \le \varepsilon + L_\varepsilon r, \qquad L_\varepsilon := K\log(\varepsilon^{-1}), \qquad \forall \varepsilon \in (0, 1/2), \quad \forall r \in (0,\infty). \tag{2.15}$$

给定 $n \ge 1$ 和 $\varepsilon \in (0,1/2)$，我们可以反复使用 (2.14) 并结合 (2.15)，得到对每个 $t \in [0,T)$，

$$\begin{aligned}
\tilde{\rho}^n(t) &\le C\int_0^t \omega_s(\tilde{\rho}^n(r))dr \le C\int_0^T (\varepsilon + L_\varepsilon \tilde{\rho}^n(r))dr \\
&\le \varepsilon CT + C^2 L_\varepsilon \int_0^T \tilde{\rho}^{n-1}(r)dr \\
&\le \cdots \le \varepsilon CT + \varepsilon\frac{C^2 L_\varepsilon T^2}{2} + \cdots + \varepsilon\frac{C^n L_\varepsilon^{n-1} T^n}{n!} + C^n L_\varepsilon^n \int_0^T \int_0^{t_1} \cdots \int_0^{t_{n-1}} \tilde{\rho}^0(t_n) dt_n dt_{n-1} \ldots d_{t_1} \\
&\le \varepsilon CT \sum_{k=0}^{n-1} \frac{(CL_\varepsilon T)^k}{k!} + \bar{C}\frac{(CL_\varepsilon T)^n}{n!} \le \varepsilon CT\exp(CKT\log(\varepsilon^{-1})) + \bar{C}\frac{(CKT\log(\varepsilon^{-1}))^n}{n!},
\end{aligned}$$

其中在倒数第二步中，我们使用了由$\mathbb{T}^d$的有界性得到的一致有界 $\tilde{\rho}^0 \le \bar{C} = \bar{C}(d)$。取 $\varepsilon = e^{-n}$，我们得到

$$\sup_{t\in[0,T)} \tilde{\rho}(t) \le CT\exp\left(-n(1-CKT)\right) + \bar{C}\frac{(CKTn)^n}{n!}.$$

然后，利用斯特林不等式 $\log(n^n/n!) \lesssim n$，我们可以找到一个足够小且与 n 无关的 T，使得

$$\sup_{t\in[0,T)} \tilde{\rho}(t) \le (C+\bar{C})e^{-n/2} \qquad \forall n \ge 0,$$

由此完成了这一步的证明。

步骤 3： 在最后这一步中，我们证明 $\{\mu^n\}_{n\ge0}$ 收敛到方程 μ（1.4）在 $C_{w^*}([0,T)$ 的一个解；$\mathscr{P}(\mathbb{T}^d))$ 在取子列的意义下成立。首先，由步骤 1、定理 A.2和定理 A.4可知，向量场 v_t^n 一致有界，并且关于同一个对数-利普希茨连续模一致连续。因此，流映射 $\{X^n\}_{n\ge0}$ 在 $[0,T) \times \mathbb{T}^d$上一致连续。由阿尔泽拉-阿斯科利定理，存在一个子列$n_k \to \infty$和一个连续映射 $X: [0,T) \times \mathbb{T}^d \to \mathbb{T}^d$，使得

$$\lim_{k\to\infty} \|X^{n_k} - X\|_{C([0,T)\times\mathbb{T}^d)} = 0.$$

定义 $\mu_t := (X_t)_{\#}\bar{\mu}$，我们可推出

$$\mu_t^{n_k} \overset{*}{\rightharpoonup} \mu_t \qquad \forall t\in[0,T), \qquad \mu \in C_{w^*}([0,T);\mathscr{P}(\mathbb{T}^d)) \cap L^\infty([0,T);\mathscr{X}_s(\mathbb{T}^d)). \tag{2.16}$$

然后，根据步骤 2，我们还可以推得

$$\mu_t^{n_k-1} \overset{*}{\rightharpoonup} \mu_t \qquad \forall t\in[0,T).$$

最后，令 $v_t := -\nabla K_s * (\mu_t - \nu)$，定理 A.4 表明

$$\|v_t^{n_k-1} - v_t\|_{C(\mathbb{T}^d)} \to 0 \qquad \forall t\in[0,T). \tag{2.17}$$

考虑到（2.16）、（2.17）以及 v^n，v的一致有界性，我们可以使用控制收敛定理，在分布形式两侧取极限

$$\int_{\mathbb{T}^d} \varphi\mu_t^{n_k} = \int_{\mathbb{T}^d} \varphi\bar{\mu} + \int_0^t \int_{\mathbb{T}^d} \nabla\varphi \cdot v_r^{n_k-1}\mu_r^{n_k} dr \qquad \forall t\in[0,T), \quad \forall \varphi \in C^\infty(\mathbb{T}^d).$$

这表明 μ 是（1.4）的一个解，从而完成证明。 □

2.2.3 极大解与延拓判据

在以下命题中，我们将证明上述构造的局部解可以延拓至某个极大存在时间，并给出一个延拓判据。

命题 2.9。设 $s \ge 1$ and $\bar{\mu},\nu \in \mathscr{P}\cap\mathscr{X}_s(\mathbb{T}^d)$。根据定理 2.3，方程 (1.4) 存在唯一的极大解 $u \in L^\infty_{\mathrm{loc}}([0,T);\mathscr{X}_s(\mathbb{T}^d))$。此外，若 $s \ge d/2+1$，则有 $T=\infty$。反之，若 $s\in[1,d/2+1)$，则 T 有限当且仅当 $\limsup_{t\to T^-}\|\mu_t\|_{L^p} = +\infty$，其中 $p = d/(2s-2) \in (1,\infty]$。

证明。考虑根据定理 2.2 得到的方程 (1.4) 的所有解构成的族 $\mathscr{S}_s(\bar{\mu},\nu)$，并将 T 定义为 $\mathscr{S}_s(\bar{\mu},\nu)$ 中所有元素的存在时间的上确界。根据定理 2.6，若 $\mu^1 \in L^\infty_{\mathrm{loc}}([0,T^1))$ 和 $\mu^2 \in L^\infty_{\mathrm{loc}}([0,T^2))$ 是方程 (1.4) 的两个解，则对所有 $t\in[0,T^1\wedge T^2)$，都有 $\mu_t^1 = \mu_t^2$。因此，可以定义曲线 $\mu: [0,T) \to \mathscr{P}(\mathbb{T}^d)$，使得对每个 $t\in[0,T)$，$\mu_t = \hat{\mu}_t$，其中 $\hat{\mu} \in \mathscr{S}_s(\bar{\mu},\nu)$ 是任意一个在某个时间 $\hat{T} > t$ 之前的解（根据 T 的定义，这样的解存在）。注意，$\mu \in L^\infty_{\mathrm{loc}}([0,T);\mathscr{X}_s(\mathbb{T}^d))$ 是方程 (1.4) 的解，因为它在局部时间范围内与 $\mathscr{S}_s(\bar{\mu},\nu)$ 中的元素一致。此外，根据构造，μ 是定理 2.3 意义下的极大解。极大解的唯一性直接由定理 2.6 得出。

为证明该陈述的最后一部分，我们首先观察到：如果 $T<\infty$和 $[0,T) \ni t \mapsto \mu_t \in (\mathscr{P}(\mathbb{T}^d)\ W_2)$ 是利普希茨连续的，那么存在一个极限测度 $\mu_* \in \mathscr{P}\cap\mathscr{X}_s(\mathbb{T}^d)$ 使得 $\mu_t \overset{*}{\rightharpoonup} \mu_*$当 $t\to T^-$时。将定理 2.8 应用于从时间 μ_*开始的初始条件 T，我们可以找到一个解

$\tilde{\mu}$ 在 $[T, T+\varepsilon)$中，对某个 $\varepsilon > 0$。将 μ 和 $\tilde{\mu}$ 进行拼接，可得到 μ的非平凡延拓，这与其极大性矛盾。另一方面，根据定理2.5 ，我们知道 μ 满足能量耗散恒等式（1.6）。特别地，由于定理 A.4，对于每个 $t \in (0, T)$，度量导数 $|\mu_t'|$ 可以被估计为

$$
\begin{aligned}
|\mu_t'| &= \left(\int_{\mathbb{T}^d} |\nabla K_s * (\mu_t - \nu)|^2 d\mu_t \right)^{1/2} \\
&\leq \|\nabla K_s * (\mu_t - \nu)\|_{L^\infty} \lesssim_{d,s} \begin{cases} \|\mu\|_{L^p} + \|\nu\|_{L^p} \quad \text{for} \quad p = \frac{d}{2s-2} & \text{if } s \in [1, \frac{d}{2} + 1), \\ 1 & \text{if } s \geq \frac{d}{2} + 1. \end{cases}
\end{aligned}
$$

因此，当 $s \geq d/2 + 1$ 时，我们总有 $\mu \in \mathrm{Lip}([0,T); (\mathscr{P}(\mathbb{T}^d) W_2))$；而对于 $s \in [1, d/2+1)$，只要 $\mu \in L^\infty([0,T); L^p)$，同样成立。证明至此结束。 □

2.3 正则性的传播

本节将证明，赫尔德正则性和索博列夫正则性会从数据传播到方程（1.4）的解，且在其极大存在时间内保持成立。第 2.3.1 节讨论赫尔德正则性的情形。特别地，我们将推导出，光滑数据会产生方程（1.4）的光滑解。结合定理 2.13 中的先验估计，我们将在第 2.3.2 节证明索伯列夫正则性的传播。

2.3.1 赫尔德正则性的传播

在下面的命题中，我们证明赫尔德正则性的传播。采用类似的策略，可以证明任何"优于 $\mathscr{X}_s(\mathbb{T}^d)$"的正则性都会传播至$\mathscr{X}_s(\mathbb{T}^d)$-解的最大存在时间（见定理 2.11）。

命题 2.10. 设 $s \geq 1$, $\bar{\mu}, \nu \in \mathscr{P} \cap \mathscr{X}_s(\mathbb{T}^d)$ 和 $\mu \in L^\infty_{\mathrm{loc}}([0,T); \mathscr{X}_s(\mathbb{T}^d))$ 是定理 *2.9* 给出的方程 *(1.4)* 的唯一极大解。设 $k \in \mathbb{N} \cup \{0\}$ 和 $\alpha \in (0,1)$。若 $\bar{\mu}, \nu \in C^{k,\alpha}(\mathbb{T}^d)$，则 $\mu \in L^\infty_{\mathrm{loc}}([0,T); C^{k,\alpha}(\mathbb{T}^d))$ 。

证明。 显然，只需证明对于任意给定的 $\tau \in (0,T)$，都有 $\mu \in L^\infty([0,\tau); C^{k,\alpha}(\mathbb{T}^d))$。令 $v_t := -\nabla K_s * (\mu_t - \nu)$ 和 X 分别为与解 μ 相关联的向量场和流映射。我们将证明分为若干步骤。

步骤 1： 我们首先证明，对于某个较小的 $0 < \tau' \leq \tau$， $\mu \in L^\infty([0,\tau'); C^{0,\alpha})$。我们按照定理 2.8 的证明，通过构造一列近似解 $\mu^n \in L^\infty([0,\tau'); C^{0,\alpha})$ 来论证，使得

$$
\mu_t^0 = \bar{\mu}, \qquad \mu_t^{n+1} = (X_t^n)_\# \bar{\mu} \qquad \forall t \in [0,\tau'), \quad \forall n \geq 0.
$$

我们声称存在 $\tau' \in (0, \tau]$ 使得

$$
\|\mu^n\|_{L^\infty([0,\tau');C^{0,\alpha})} \leq 2\|\bar{\mu}\|_{C^{0,\alpha}} \qquad \forall n \geq 0. \tag{2.18}
$$

一旦证明了（2.18），令 $n \to \infty$ 趋于极限，即可推出 $\mu \in L^\infty([0,\tau'); C^{0,\alpha})$，如所求。注意，当 $n = 0$ 时，（2.18）显然成立。假设它对某个 $n \geq 0$ 成立，下面证明它对 $n+1$ 也成立。由算子 $\nabla(-\Delta)^{-s} : C^{0,\alpha} \to C^{1,\alpha}$ 的连续性，我们得到

$$
\|v_t^n\|_{C^{1,\alpha}} \leq C(d,s,\alpha)\|\mu_t - \nu\|_{C^{0,\alpha}} \leq C(d,s,\alpha)\left(\|\nu\|_{C^{0,\alpha}} + 2\|\bar{\mu}\|_{C^{0,\alpha}}\right) =: L \qquad \forall t \in [0,\tau').
$$

因此，对恒等式 $\frac{d}{dt}\nabla X_t^n = \nabla v_t^n \nabla X_t^n$ 两边取范数并对时间积分，得到

$$
\max\{\|\nabla X_t^n\|_{C^{0,\alpha}}, \|\nabla (X_t^n)^{-1}\|_{C^{0,\alpha}}\} \leq \exp\left(\int_0^t \|\nabla v_r\|_{C^{0,\alpha}} dr \right) \leq \exp(L\tau') \qquad \forall t \in [0,\tau').
$$

特别是 特别是，利用表示公式 $\mu_t^{n+1} = (\bar{\mu} \det \nabla X_t^n) \circ (X_t^n)^{-1}$，我们得出结论

$$
\|\mu_t^{n+1}\|_{C^{0,\alpha}} \leq \left(1 + \|\nabla (X_t^n)^{-1}\|_{C^{0,\alpha}}^\alpha\right) \|\nabla X_t^n\|_{C^{0,\alpha}}^d \|\bar{\mu}\|_{C^{0,\alpha}} \leq \exp\left((\alpha + d)L\tau'\right) \|\bar{\mu}\|_{C^{0,\alpha}} \leq 2\|\bar{\mu}\|_{C^{0,\alpha}},
$$

前提是我们选择的 $\tau' > 0$ 足够小。

步骤 2： 令 $\tau'' \in [\tau', \tau]$ 为解在 $C^{0,\alpha}$ 中存在的最大时间，

$$\tau'' := \sup\{t \in [\tau', \tau] : \mu \in L^\infty([0,t); C^{0,\alpha})\}.$$

在这一步中，我们证明 $\tau'' = \tau$。按照定理 2.9 中的延拓论证，我们只需证明$\mu \in L^\infty([0,\tau''); C^{0,\alpha})$。令

$$M := \|\nu\|_{\mathscr{X}_s} + \sup_{t\in[0,\tau)} \|\mu_t\|_{\mathscr{X}_s}.$$

由于 $\mu \in L^\infty_{\mathrm{loc}}([0,\tau''); C^{0,\alpha})$，我们有 $v \in L^\infty_{\mathrm{loc}}([0,\tau''); C^{1,\alpha})$。令 $\tilde{\mu}_t := \mu_t \circ X_t$。由于 μ满足方程（1.4），且向量场是正则的，我们得到

$$\frac{d}{dt}\tilde{\mu}_t = -\tilde{\mu}_t \operatorname{div} v_t \circ X_t \qquad \forall t \in (0,\tau''). \tag{2.19}$$

我们首先证明存在某个有限常数 $K > 0$，使得

$$\|\mu_t\|_{L^\infty} = \|\tilde{\mu}_t\|_{L^\infty} \le K \qquad \forall t \in [0,\tau''). \tag{2.20}$$

当 $s = 1$时，根据假设，我们已知 $\|\mu_t\|_{L^\infty}$（因此 $\|\tilde{\mu}_t\|_{L^\infty}$）在 $[0,\tau) \supseteq [0,\tau'')$ 中一致有界。当 $s \in (1, d/2+1]$ 时，由（2.19）和定理 A.5，令 $p = d/(2s-2) \in [1,\infty)$ 可得

$$\begin{aligned}\frac{d}{dt}\|\tilde{\mu}_t\|_{L^\infty} &\le \|\tilde{\mu}_t\|_{L^\infty}\|\nabla^2 K_s * (\mu_t - \nu)\|_{L^\infty} \\ &\lesssim_{d,s} \|\tilde{\mu}_t\|_{L^\infty}\|\mu_t - \nu\|_{L^p} \log\left(1 + \frac{\|\mu_t - \nu\|_{L^\infty}}{\|\mu_t - \nu\|_{L^p}}\right) \lesssim_{d,s,M} \|\tilde{\mu}_t\|_{L^\infty} \log\left(2 + \|\nu\|_{L^\infty} + \|\tilde{\mu}\|_{L^\infty}\right).\end{aligned}$$

因此， $\|\tilde{\mu}\|_{L^\infty}$ 随时间至多呈双指数增长，并且在 $[0,\tau'')$上一致有界。最后，当 $s > d/2+1$ 时同样成立，这是因为 $\nabla^2 K_s * (\mu_t - \nu)$ 由定理 A.2 给出的一致有界，以及对（2.19）应用格朗沃尔不等式。

为证明 $\mu \in L^\infty([0,\tau''); C^{0,\alpha})$，我们区分两种情况： $s > 1$和 $s = 1$。若 $s > 1$，由于 $\nabla v_t = \nabla^2 K_s * (\mu_t - \nu)$，Schauder 估计给出

$$\begin{aligned}\|\nabla v_t\|_{L^\infty} &\lesssim_{d,s} \|\mu_t\|_{L^\infty} + \|\nu\|_{L^\infty}, \\ \|\nabla v_t\|_{C^{0,\alpha}} &\lesssim_{d,s,\alpha} \|\mu_t\|_{C^{0,\alpha}} + \|\nu\|_{C^{0,\alpha}}.\end{aligned} \qquad \forall t \in [0,\tau''). \tag{2.21}$$

因此，根据（2.20），我们得到

$$\|\nabla v_t\|_{L^\infty}, \operatorname{Lip}(X_t), \operatorname{Lip}(X_t^{-1}) \le \tilde{K} \qquad \|\mu_t\|_{C^{0,\alpha}} \le \|\tilde{\mu}_t\|_{C^{0,\alpha}}(1 + \tilde{K}^\alpha) \qquad \forall t \in [0,\tau''), \tag{2.22}$$

其中某个有限常数为 $K > 0$。由（2.19），利用（2.20）、（2.21）和（2.22），对每个 $t \in (0, \tau'')$，我们得到

$$\begin{aligned}\frac{d}{dt}\|\tilde{\mu}_t\|_{C^{0,\alpha}} &\le \|\tilde{\mu}_t \nabla v_t \circ X_t\|_{C^{0,\alpha}} \\ &\le \|\tilde{\mu}_t\|_{C^{0,\alpha}}\|\nabla v_t\|_{L^\infty} + \|\tilde{\mu}_t\|_{L^\infty}\left(\|\nabla v_t\|_{L^\infty} + \|\nabla v_t\|_{C^{0,\alpha}} \operatorname{Lip}(X_t)^\alpha\right) \\ &\lesssim_{d,s,\alpha,K,\tilde{K}} \|\tilde{\mu}_t\|_{C^{0,\alpha}} + 1.\end{aligned}$$

因此，利用格朗沃尔不等式，我们可推出 $\tilde{\mu} \in L^\infty([0,\tau''); C^{0,\alpha})$，并由 (2.22)，有 $\mu \in L^\infty([0,\tau''); C^{0,\alpha})$。当 $s = 1$时，有 $-$散度 $v_t = \mu_t - \nu$，因此由 (2.19)，对所有 $t \in (0,\tau'')$，

$$\frac{d}{dt}\|\tilde{\mu}_t\|_{C^{0,\alpha}} \le \|\tilde{\mu}_t(\tilde{\mu}_t - \nu \circ X_t)\|_{C^{0,\alpha}} \lesssim_M \|\tilde{\mu}_t\|_{C^{0,\alpha}} + \|\nu\|_{C^{0,\alpha}} \operatorname{Lip}(X_t)^\alpha.$$

由格朗沃尔不等式可得，对于所有 $t \in [0,\tau'')$，

$$\|\mu_t\|_{C^{0,\alpha}} \le \operatorname{Lip}(X_t^{-1})^\alpha \|\tilde{\mu}_t\|_{C^{0,\alpha}} \lesssim_{M,\tau} \operatorname{Lip}(X_t^{-1})^\alpha \left(1 + \|\nu\|_{C^{0,\alpha}} \int_0^t \operatorname{Lip}(X_r)^\alpha dr\right). \tag{2.23}$$

另一方面，我们知道

$$\mathrm{Lip}(X_t)^\alpha, \mathrm{Lip}(X_t^{-1})^\alpha \le \exp\left(\alpha \int_0^t \|\nabla v_r\|_{L^\infty} dr\right). \tag{2.24}$$

特别地，结合（A.5）中的定理以及（2.23）和（2.24），对于所有 $t \in [0, \tau'')$，我们得到

$$\begin{aligned}\|\nabla v_t\|_{L^\infty} &\lesssim_{d,\alpha} \|\mu_t - \nu\|_{L^\infty} \left(1 + \log\left(\frac{\|\mu_t - \nu\|_{C^{0,\alpha}}}{\|\mu_t - \nu\|_{L^\infty}}\right)\right) \\ &\lesssim_M 1 + \log\left(2 + \|\nu\|_{C^{0,\alpha}} + \|\mu_t\|_{C^{0,\alpha}}\right) \lesssim_{M,\alpha,\tau,\|\nu\|_{C^{0,\alpha}}} 1 + \int_0^t \|\nabla v_r\|_{L^\infty} dr.\end{aligned}$$

再应用一次格朗沃尔不等式可得 $v \in L^\infty([0,\tau''); C^{0,1})$，由此最终推得 $\mu \in L^\infty([0,\tau''); C^{0,\alpha})$，这归功于（2.23）和（2.24）。

步骤 3 $\mu \in L^\infty([0,\tau); C^{k,\alpha})$：最后，我们通过自举论证证明 $\mu \in L^\infty([0,\tau); C^{k',\alpha})$（$0 \le k' < k$）。假设对于某个 k'，我们已知 $\nabla(-\Delta)^{-s} : C^{k',\alpha} \to C^{k'+1,\alpha}$。那么，由 v 的连续性，我们可推得 $v \in L^\infty([0,\tau); C^{k'+1,\alpha})$，进而得到 $X, X^{-1} \in L^\infty([0,\tau); C^{k'+1,\alpha})$。函数 $\tilde{\mu}_t = \mu_t \circ X_t$ 满足 $\partial_t \tilde{\mu}_t = -\tilde{\mu}_t[(-\Delta)^{1-s}(\mu_t - \nu)] \circ X_t$。因此，利用霍尔德空间中的复合与乘积法则，以及 $\tilde{\mu}$ 在 $C^{k',\alpha}$ 中的一致有界性和 X_t 在 $C^{k'+1,\alpha}$ 中的一致有界性，可得

$$\frac{d}{dt}\|\tilde{\mu}_t\|_{C^{k'+1,\alpha}} \le \|\tilde{\mu}_t[(-\Delta)^{1-s}(\mu_t - \nu)] \circ X_t\|_{C^{k'+1,\alpha}} \lesssim \|\tilde{\mu}_t\|_{C^{k'+1,\alpha}} + \|\nu\|_{C^{k'+1,\alpha}}.$$

由格朗沃尔不等式可推出 $\tilde{\mu} \in L^\infty([0,\tau); C^{k'+1,\alpha})$，再与逆流复合可得 $\mu \in L^\infty([0,\tau); C^{k'+1,\alpha})$。至此，提升论证完成，证明也随之结束。 □

注 2.11（一般正则性的传播）。定理 2.10 证明的关键在于，只要 μ_t 比 $\mathscr{X}_s(\mathbb{T}^d)$ 稍具更高的正则性，它就会生成利普希茨向量场 v_t。这是借助格朗沃尔不等式传播（Hölder）正则性所需的关键条件。同样的论证也适用于传播其他类型的正则性。例如，可以采用类似策略证明：若 $s \in (1, d/2+1]$，$p = d/(2s-2)$，$r > p$，且 $\bar{\mu}, \nu \in L^r(\mathbb{T}^d)$，则 $\mu \in L^\infty_{\mathrm{loc}}([0,T); L^r)$，其中 T 是 $\mathscr{X}_s$-解的最大存在时间。若 $s > d/2+1$，对所有 $r \ge 1$ 也同样成立。此外，利用同一技巧，我们可以传播 C^k,1-正则性，其中 $k \in \mathbb{N} \cup \{0\}$ 任意，只要 $s > 1$。■

注 2.12（时间正则性）。对于极大解 $\mu \in L^\infty([0,T); \mathscr{X}_s(\mathbb{T}^d))$ 的（1.4），定理2.10 所提供的（时间上一致的）空间正则性可用于证明时间正则性。例如，记由 μ分别生成的速度场和流映射为 v 和 X，则有：

$$\bar{\mu}, \nu \in C^\infty \implies \mu \in C^\infty([0,T) \times \mathbb{T}^d),\ v \in C^\infty([0,T) \times \mathbb{T}^d; \mathbb{R}^d),\ X \in C^\infty([0,T) \times \mathbb{T}^d; \mathbb{T}^d).$$

事实上，在此情形下，定理 2.10 给出 $\mu \in L^\infty_{\mathrm{loc}}([0,T); C^k)$ 对所有 $k \in \mathbb{N}$均成立，其中 $T > 0$ 是解 μ的最大存在时间。由此进一步可推出，对所有 $k \in \mathbb{N}$，有 $v \in L^\infty_{\mathrm{loc}}([0,T); C^k)$。接下来，对 $\frac{d}{dt}X = v \circ X$关于时间和空间求导，并按照导数阶数使用归纳法，可以推出$X, X^{-1} \in C^\infty([0,T) \times \mathbb{T}^d; \mathbb{T}^d)$，最后，由表示公式 $\mu_t = (\bar{\mu}\,\mathrm{det}\,\nabla X_t) \circ X_t^{-1}$，可得$\mu \in C^\infty([0,T) \times \mathbb{T}^d)$以及 $v \in C^\infty([0,T) \times \mathbb{T}^d; \mathbb{R}^d)$。■

2.3.2 索伯列夫正则性的传播

在下一个引理中，我们将推导方程 (1.4) 的解在齐次索оболев空间中的能量估计。作为应用，我们将证明索伯列夫正则性的传播。这些估计将在定理1.2 和定理 1.4中的光滑收敛结果的证明中发挥基础性作用。

引理 2.13。设 $s \ge 1$，$\gamma > d/2$，以及 $\bar{\mu}, \nu \in \mathscr{P} \cap C^\infty(\mathbb{T}^d)$。令 $\mu \in C^\infty([0,T) \times \mathbb{T}^d)$ 为定理 (1.4) 以及定理 *2.9* 和定理 *2.12*给出的极大解。则以下能量估计对

某个 $C=C(d,\ s,\ \gamma)>0$ 以及所有 $t\in(0,\ T)$：

$$\frac{d}{dt}\|\mu_t\|_{\dot{H}^\gamma}^2 \le C\left(\|\nabla^2(-\Delta)^{-s}(\mu_t-\nu)\|_{L^\infty}+\|\nu\|_{H^\gamma}\right)\left(\|\mu_t\|_{\dot{H}^\gamma}^2+\|\nu\|_{\dot{H}^\gamma}^2\right), \tag{2.25}$$

$$\begin{aligned}\frac{d}{dt}\|\mu_t-\nu\|_{\dot{H}^\gamma}^2 \le{}& C\|\nabla^2(-\Delta)^{-s}(\mu_t-\nu)\|_{L^\infty}\|\mu_t-\nu\|_{\dot{H}^\gamma}^2 \\ &-2\big(\min_{\mathbb{T}^d}\nu\big)\|\mu_t-\nu\|_{\dot{H}^{\gamma-s+1}}^2+C\|\nu\|_{\dot{H}^{\gamma+s}}\|\mu_t-\nu\|_{\dot{H}^{\gamma-s}}\|\mu_t-\nu\|_{\dot{H}^{\gamma-s+1}}.\end{aligned} \tag{2.26}$$

证明。 下文中，对于所有 $\beta\in\mathbb{R}$，我们使用记号 $D^\beta:=(-\Delta)^{\beta/2}$，并考虑满足方程的零均值扰动 $\sigma_t:=\mu_t-\nu$

$$\partial_t\sigma_t=\operatorname{div}(\nu\nabla(-\Delta)^{-s}\sigma_t)+\operatorname{div}(\sigma_t\nabla(-\Delta)^{-s}\sigma_t)\qquad\forall t\in(0,T).$$

步骤 1： 我们首先证明（2.25）。利用方程（1.4）、分部积分并展开分数阶导数，我们得到三项：

$$\begin{aligned}\frac{d}{dt}\frac{\|\mu_t\|_{\dot{H}^\gamma}^2}{2}&=\int_{\mathbb{T}^d}D^\gamma\operatorname{div}\left(\mu_t\nabla D^{-2s}\sigma_t\right)D^\gamma\mu_t\\&=-\int_{T^d}D^\gamma\mu_t\nabla D^{-2s}\sigma_t\cdot\nabla D^\gamma\mu_t\\&\quad+\int_{\mathbb{T}^d}\left[D(D^\gamma\mu_t\nabla D^{-2s}\sigma_t)-D^{\gamma+1}\mu_t\nabla D^{-2s}\sigma_t\right]\cdot\nabla D^{\gamma-1}\mu_t\\&\quad-\int_{\mathbb{T}^d}\left[D^{\gamma+1}(\mu_t\nabla D^{-2s}\sigma_t)-D^{\gamma+1}\mu_t\nabla D^{-2s}\sigma_t\right]\cdot\nabla D^{\gamma-1}\mu_t=:I+II+III.\end{aligned}$$

对于 I，我们使用恒等式 $2D^\gamma\mu_t\nabla D^\gamma\mu_t=\nabla(D^\gamma\mu_t)^2$, 以及分部积分：

$$I=\frac{1}{2}\int_{\mathbb{T}^d}\Delta D^{-2s}\sigma_t(D^\gamma\mu_t)^2\le C\|\nabla^2D^{-2s}\sigma_t\|_{L^\infty}\|\mu_t\|_{\dot{H}^\gamma}^2. \tag{2.27}$$

对于 II，我们使用柯西–施瓦茨不等式和（A.9）：

$$\begin{aligned}II&\le\|D(D^\gamma\mu_t\nabla D^{-2s}\sigma_t)-D^{\gamma+1}\mu_t\nabla D^{-2s}\sigma_t\|_{L^2}\|\nabla D^{\gamma-1}\mu_t\|_{L^2}\\&\lesssim_{d,\gamma}\|\nabla^2D^{-2s}\sigma_t\|_{L^\infty}\|D^\gamma\mu_t\|_{L^2}\|\nabla D^{\gamma-1}\mu_t\|_{L^2}\lesssim_{d,\gamma}\|\nabla^2D^{-2s}\sigma_t\|_{L^\infty}\|\mu_t\|_{\dot{H}^\gamma}^2.\end{aligned} \tag{2.28}$$

然后我们写出 $\mu_t=\sigma_t+\nu$，并进一步将 III 分成两项：

$$\begin{aligned}III=&-\int_{\mathbb{T}^d}\left[D^{\gamma+1}(\sigma_t\nabla D^{-2s}\sigma_t)-D^{\gamma+1}\sigma_t\nabla D^{-2s}\sigma_t\right]\cdot\nabla D^{\gamma-1}\mu_t\\&-\int_{\mathbb{T}^d}\left[D^{\gamma+1}(\nu\nabla D^{-2s}\sigma_t)-D^{\gamma+1}\nu\nabla D^{-2s}\sigma_t\right]\cdot\nabla D^{\gamma-1}\mu_t=:III_1+III_2.\end{aligned}$$

为了界定 III_1，我们使用柯西–施瓦茨不等式，然后将（A.11）应用于函数 $\phi_t=D^{-2s}\sigma_t$：

$$\begin{aligned}III_1&\le\|D^{\gamma+1}(D^{2s}\phi_t\nabla\phi_t)-D^{\gamma+1+2s}\phi_t\nabla\phi_t\|_{L^2}\|\nabla D^{\gamma-1}\mu_t\|_{L^2}\\&\lesssim_{d,s,\gamma}\|\nabla^2\phi_t\|_{L^\infty}\|D^{\gamma+2s}\phi_t\|_{L^2}\|\mu_t\|_{\dot{H}^\gamma}=\|\nabla^2D^{-2s}\sigma_t\|_{L^\infty}\|\sigma_t\|_{\dot{H}^\gamma}\|\mu_t\|_{\dot{H}^\gamma}.\end{aligned} \tag{2.29}$$

为了界定 III_2，我们使 用柯西–施瓦茨不等式，以及（A.8）连同索伯列夫嵌入 $H^\gamma\hookrightarrow L^\infty$ ：

$$\begin{aligned}III_2&\le\|D^{\gamma+1}(\nu\nabla D^{-2s}\sigma_t)-D^{\gamma+1}\nu\nabla D^{-2s}\sigma_t\|_{L^2}\|\nabla D^{\gamma-1}\mu_t\|_{L^2}\\&\lesssim_{d,s,\gamma}\left(\|D^\gamma\nu\|_{L^2}\|\nabla^2D^{-2s}\sigma_t\|_{L^\infty}+\|\nu\|_{L^\infty}\|\nabla D^{\gamma+1-2s}\sigma_t\|_{L^2}\right)\|\mu_t\|_{\dot{H}^\gamma}\\&\lesssim_d\|\nu\|_{\dot{H}^\gamma}\|\nabla^2D^{-2s}\sigma_t\|_{L^\infty}\|\mu_t\|_{\dot{H}^\gamma}+\|\nu\|_{H^\gamma}\|\sigma_t\|_{\dot{H}^\gamma}\|\mu_t\|_{\dot{H}^\gamma}.\end{aligned} \tag{2.30}$$

综合（2.27）、（2.28）、（2.29）和（2.30），可得（2.25）。

步骤 2： 接下来，我们证明（2.26）。对 $\|\sigma_t\|_{\dot{H}^\gamma}^2$ 关于时间求导，可得两项之和：

$$\frac{d}{dt}\frac{\|\sigma_t\|_{\dot{H}^\gamma}^2}{2} = \underbrace{\int_{\mathbb{T}^d} D^\gamma \operatorname{div}(\nu\nabla D^{-2s}\sigma_t) D^\gamma \sigma_t}_{\text{(Lin)}} + \underbrace{\int_{\mathbb{T}^d} D^\gamma \operatorname{div}(\sigma_t \nabla D^{-2s}\sigma_t) D^\gamma \sigma_t}_{\text{(NonLin)}}.$$

我们首先处理（线性项），它源于线性化方程（见第1.5节）。通过分部积分，然后在被积函数的两个因子之间分配分数阶导数，可得

$$\begin{aligned}
\text{(Lin)} &= -\int_{\mathbb{T}^d} D^\gamma\left(\nu\nabla D^{-2s}\sigma_t\right)\cdot\nabla D^\gamma\sigma_t \\
&= -\int_{\mathbb{T}^d} D^{\gamma+s}\left(\nu\nabla D^{-2s}\sigma_t\right)\cdot\nabla D^{\gamma-s}\sigma_t \\
&= \underbrace{-\int_{\mathbb{T}^d}\nu|\nabla D^{\gamma-s}\sigma_t|^2}_{(\text{Lin})_{\text{main}}} + \underbrace{(-1)\int_{\mathbb{T}^d}\left[D^{\gamma+s}\left(\nu\nabla D^{-2s}\sigma_t\right) - \nu D^{\gamma+s}\nabla D^{-2s}\sigma_t\right]\cdot\nabla D^{\gamma-s}\sigma_t}_{(\text{Lin})_{\text{err}}}.
\end{aligned}$$

现在，对于主项（线性项）$_{\text{main}}$，我们只需用其最小值对 ν 进行下界估计，即得

$$(\text{Lin})_{\text{main}} \le -\big(\min_{\mathbb{T}^d}\nu\big)\int_{\mathbb{T}^d}|\nabla D^{\gamma-s}\sigma_t|^2 = -\big(\min_{\mathbb{T}^d}\nu\big)\|\sigma_t\|_{\dot{H}^{\gamma-s+1}}^2. \tag{2.31}$$

对于误差项（线性项）$_{\text{err}}$，利用柯西–施瓦茨不等式和 Kato–Ponce 交换子估计（A.8），可得

$$\begin{aligned}
(\text{Lin})_{\text{err}} &\le \left\|D^{\gamma+s}\left(\nu\nabla D^{-2s}\sigma_t\right) - \nu D^{\gamma+s}\nabla D^{-2s}\sigma_t\right\|_{L^2}\|\nabla D^{\gamma-s}\sigma_t\|_{L^2} \\
&\lesssim_{d,s} \left(\|\nabla\nu\|_{L^\infty}\|D^{\gamma+s-1}\nabla D^{-2s}\sigma_t\|_{L^2} + \|D^{\gamma+s}\nu\|_{L^2}\|\nabla D^{-2s}\sigma_t\|_{L^\infty}\right)\|\sigma_t\|_{\dot{H}^{\gamma-s+1}} \\
&\lesssim_{d,s,\gamma} \|\nu\|_{\dot{H}^{\gamma+s}}\|\sigma_t\|_{\dot{H}^{\gamma-s}}\|\sigma_t\|_{\dot{H}^{\gamma-s+1}},
\end{aligned} \tag{2.32}$$

其中，在最后一步中我们还使用了嵌入 $\dot{H}^{\gamma+s-1}\hookrightarrow C^0$。

非线性项（NonLin）可类似于步骤 1 进行处理。拆分后

$$\begin{aligned}
\text{(NonLin)} &= -\int_{T^d} D^\gamma\sigma_t\nabla D^{-2s}\sigma_t\cdot\nabla D^\gamma\sigma_t \\
&\quad + \int_{\mathbb{T}^d}\left[D(D^\gamma\sigma_t\nabla D^{-2s}\sigma_t) - D^{\gamma+1}\sigma_t\nabla D^{-2s}\sigma_t\right]\cdot\nabla D^{\gamma-1}\sigma_t \\
&\quad - \int_{\mathbb{T}^d}\left[D^{\gamma+1}(\sigma_t\nabla D^{-2s}\sigma_t) - D^{\gamma+1}\sigma_t\nabla D^{-2s}\sigma_t\right]\cdot\nabla D^{\gamma-1}\sigma_t \\
&=: (\text{NonLin})_1 + (\text{NonLin})_2 + (\text{NonLin})_3,
\end{aligned}$$

我们分别可以从步骤 1 得到 $(\text{NonLin})_1$, $(\text{NonLin})_2$, 和 $(\text{NonLin})_3$ 的估计为 I，II，以及 III_1，从而得到

$$(\text{NonLin}) \lesssim_{d,s,\gamma} \|\nabla^2 D^{-2s}\sigma_t\|_{L^\infty}\|\sigma_t\|_{\dot{H}^\gamma}^2. \tag{2.33}$$

综合 (2.31)、(2.32) 和 (2.33)，得到 (2.26)。 □

命题 2.14。 设 $s\ge 1$，$\gamma > d/2$，$\bar{\mu}$，$\nu\in\mathscr{P}\cap H^\gamma(\mathbb{T}^d)$，并设 $\mu\in L^\infty_{\text{loc}}([0,T);\ \mathscr{X}_s(\mathbb{T}^d))$ 是由定理 (1.4) 给出的唯一极大解，2.9。那么，$\mu\in L^\infty_{\text{loc}}([0,T);\ H^\gamma(\mathbb{T}^d))$。

证明。 我们利用定理 2.25) 中的先验能量估计，并通过光滑解进行逼近。2.13。设 $\bar{\mu}^n$，$\nu^n\in\mathscr{P}\cap C^\infty(\mathbb{T}^d)$ 满足

$$\|\bar{\mu}^n - \bar{\mu}\|_{H^\gamma}\to 0, \quad \|\nu^n-\nu\|_{H^\gamma}\to 0 \qquad \text{as } n\to\infty, \tag{2.34}$$

并且令 μ^n 为方程（1.4）在初始测度与目标测度 $\bar{\mu}^n$ 和 ν^n下的极大解，其极大存在时间为 T^n。我们将证明分为三个步骤。

步骤 1： 在这一步中，我们推导出 T^n的一个一致下界，以及 μ_t^n的索伯列夫界。由 $\nabla^2(-\Delta)^{-s}$ 的连续性：$\dot{H}^\gamma \to C^0$，由（2.25）可得

$$\frac{d}{dt}\left(\|\mu_t^n\|_{\dot{H}^\gamma}^2 + \|\nu^n\|_{H^\gamma}^2\right) \le C\left(\|\mu_t^n\|_{\dot{H}^\gamma}^2 + \|\nu^n\|_{H^\gamma}^2\right)^{\frac{3}{2}} \qquad \forall t \in (0, T^n). \tag{2.35}$$

对该微分不等式积分，并考虑到定理 2.9中的延拓判据，我们得到以下界：

$$T^n \ge T_{\min}^n := C^{-1}\left(\|\bar{\mu}^n\|_{\dot{H}^\gamma}^2 + \|\nu^n\|_{H^\gamma}^2\right)^{-1/2}, \tag{2.36}$$

$$\|\mu_t^n\|_{\dot{H}^\gamma} \le \frac{\left(\|\bar{\mu}^n\|_{\dot{H}^\gamma}^2 + \|\nu^n\|_{H^\gamma}^2\right)^{1/2}}{1 - C\left(\|\bar{\mu}^n\|_{\dot{H}^\gamma}^2 + \|\nu^n\|_{H^\gamma}^2\right)^{1/2} t} \qquad \forall t \in [0, T_{\min}^n). \tag{2.37}$$

步骤 2： 接下来，我们证明对于某个较小的 $\mu \in L^\infty([0,\tau);\ H^\gamma(\mathbb{T}^d))$ 成立。由（2.34）和（2.36），我们得到

$$\liminf_{n\to\infty} T^n \ge \liminf_{n\to\infty} T_{\min}^n = C^{-1}\left(\|\bar{\mu}\|_{\dot{H}^\gamma}^2 + \|\nu\|_{H^\gamma}^2\right)^{-1/2} =: T_{\min} > 0.$$

定义 $\tau := (T_{\min}/2) \wedge T$，定理 2.6 中的稳定性结果保证测度$\mu_t^n \overset{*}{\rightharpoonup} \mu_t$的弱-$*$ 收敛，其中 $n \to \infty$对所有 $t \in [0,\tau)$ 均成立。因此，由于根据（2.37），近似解 μ^n 在 $L^\infty\left([0,\tau;\ H^\gamma(\mathbb{T}^d)\right)$中一致有界，我们得到 $\mu \in L^\infty([0,\tau);\ H^\gamma(\mathbb{T}^d))$，并且

$$\mu_t^n \rightharpoonup \mu_t \quad \text{in } H^\gamma(\mathbb{T}^d), \qquad \mu_t^n - \nu^n \to \mu_t - \nu \quad \text{in } \dot{H}^{\gamma'}(\mathbb{T}^d) \quad \forall \gamma' < \gamma \qquad \forall t \in [0,\tau). \tag{2.38}$$

由（2.38）以及 $\nabla^2(-\Delta)^{-s}$ 的连续性：$\dot{H}^{\gamma'} \to C^0$，对于 $\gamma' \in \left(\frac{d}{2}, \gamma\right)$，我们可以对（2.25）的积分形式取极限，从而得到对所有 $t \in [0,\tau)$，

$$\left(\|\mu_t\|_{\dot{H}^\gamma}^2 + \|\nu\|_{H^\gamma}^2\right)^{1/2} \le \left(\|\bar{\mu}\|_{\dot{H}^\gamma}^2 + \|\nu\|_{H^\gamma}^2\right)^{1/2} \exp\left(C\|\nu\|_{H^\gamma} t + \int_0^t C\|\nabla^2(-\Delta)^{-s}(\mu_r - \nu)\|_{L^\infty} dr\right). \tag{2.39}$$

步骤 3： 设 $\tau' \in [\tau, T]$为满足以下条件的最大时刻：$\mu \in L^\infty_{\mathrm{loc}}([0,\tau');\ H^\gamma(\mathbb{T}^d))$ 以及 (2.39) 在 $[0,\tau')$中成立。在最后这一步中，我们证明 $\tau' = T$，从而完成证明。We m假设 $\tau' < \infty$，否则便无须证明。我们必然有

$$\limsup_{t\to(\tau')^-} \|\mu_t\|_{H^\gamma} = \infty. \tag{2.40}$$

事实上，若有 $\mu \in L^\infty([0,\tau');\ H^\gamma)$，仿照定理 2.9 的证明进行推理，可得 μ_t在 H^γ中当 $t \to (\tau')^-$时存在弱极限；从时刻 τ'开始重复步骤 2，我们便能在 H^γ 中延拓解，并使 (2.39) 在最大时刻 τ'之后仍成立。

反设 $\tau' < T$，则由矛盾可知。根据 T 的定义，存在某个常数 $K > 0$，使得

$$\|\mu_t\|_{\mathscr{X}_s} \le K \qquad \forall t \in [0,\tau'). \tag{2.41}$$

由定理A.5 和（2.41）可得

$$\|\nabla^2(-\Delta)^{-s}(\mu_t - \nu)\|_{L^\infty} \lesssim_{d,s,\gamma,K} 1 + \log\left(2 + \|\mu_t - \nu\|_{\dot{H}^\gamma}\right).$$

结合（2.39）可得

$$\|\nabla^2(-\Delta)^{-s}(\mu_t - \nu)\|_{L^\infty} \lesssim_{d,s,\gamma,K,\|\bar{\mu}\|_{H^\gamma},\|\nu\|_{H^\gamma},\tau'} 1 + \int_0^t \|\nabla^2(-\Delta)^{-s}(\mu_r - \nu)\|_{L^\infty} dr \qquad \forall t \in [0,\tau').$$

因此，根据格朗沃尔不等式，$\|\nabla^2(-\Delta)^{-s}(\mu_t - \nu)\|_{L^\infty}$ 在 $[0,\tau')$上一致有界，特别地$\mu \in L^\infty([0,\tau');\ H^\gamma)$ 根据（2.39）。这与（2.40）矛盾，从而完成证明。 □

定理*1.1*的证明。根据定理 $\mathscr{X}_s(\mathbb{T}^d)$，2.3中证明了2.9中的极大解的存在性与唯一性，以及爆破判据。定理2.10和定理2.14分别证明了 Hölder 正则性和索博列夫正则性从数据传播到极大存在区间内的所有时刻。

□

3 定量收敛结果

本节证明 Riesz 核均值差异的定量收敛结果。定理1.2 （库仑情形 $s = 1$）在第3.1节证明，而定理1.4 （ $s > 1$情形）在第3.2节证明。

3.1 情形 $s = 1$

我们从情形 $s = 1$开始，该情形对应于库仑相互作用能。在第 3.1.1 节中，我们首先建立最大值原理；这是库仑动力学的一个重要结构特征，而对于 $s > 1$则不成立。最大值原理、耗散恒等式（1.6）以及高阶能量估计（2.26）是定理 1.2证明的主要组成部分，该证明在第3.1.2节中给出。最后，在第3.1.3节中，我们说明，初始测度的下界并非获得某种形式的目标指数收敛所必需（见定理 1.3）。

3.1.1 最大值原理

命题3.1。 设 $s = 1$、 $\bar{\mu}$、 $\nu \in \mathscr{P} \cap L^\infty(\mathbb{T}^d)$以及 $\mu \in L^\infty_{\mathrm{loc}}([0,T);\ L^\infty(\mathbb{T}^d))$ 是由定理(1.4) 给出的2.9的极大解。那么， $T = \infty$ 以及

$$\min\{\inf \bar{\mu}, \inf \nu\} \le \inf \mu_t \le \sup \mu_t \le \max\{\sup \bar{\mu}, \sup \nu\} \qquad \forall t \in [0,\infty). \tag{3.1}$$

证明。根据定理 1.1， T有限当且仅当 $\limsup_{t\to T^-} \|\mu_t\|_{L^\infty} = \infty$。因此，只需证明存在区间内每个时刻都成立的最大值原理 (3.1)。通过对初始测度与目标测度进行光滑化论证，并利用定理 2.6 的稳定性结果，我们可以将证明限制为验证 $\bar{\mu}$、 $\nu \in C^\infty$ 的不等式。事实上，本质上界和下界在弱收敛下极限中仍得以保留。在这种情况下，根据定理 2.12， μ 及其对应的向量场 v 和流映射 X 在时空中光滑。我们使用特征线法：只需证明 $\tilde{\mu}_t$ 的上界和下界：$= \mu_t \circ X_t$。由 $\partial_t \mu_t + \mathrm{div}(\mu_t v_t) = 0$，以及 $v_t = -\nabla(-\Delta)^{-1}(\mu_t - \nu)$，可推得 $\tilde{\mu}$ 满足

$$\frac{d}{dt}\tilde{\mu}_t(x) = -(\mu_t \operatorname{div} v_t) \circ X_t(x) = -\tilde{\mu}_t(x)(\tilde{\mu}_t(x) - \nu \circ X_t(x)) \qquad \forall x \in \mathbb{T}^d,\ \forall t \in (0,T). \tag{3.2}$$

我们只证明 (3.1) 中的上界，下界的证明类似。反设存在 $x_0 \in \mathbb{T}^d$和 $t_0 \in (0,T)$，使得 $\tilde{\mu}_{t_0}(x_0) > \max\{\max \bar{\mu}, \max \nu\}$. 令 t_1 为满足$\tilde{\mu}_t(x_0) = \max\{\max \bar{\mu}, \max \nu\}$的最大 $t \in [0, t_0)$。于是，对每个 $t \in (t_1, t_0)$，都有 $\tilde{\mu}_t(x_0) > \max \nu$，并且根据 (3.2)，它是 $[t_1, t_0)$ 上的单调递减函数，矛盾。 □

3.1.2 定理1.2 的证明

定理 1.2的证明。 在定理 3.1 中，我们已经证明解 μ 关于时间是全局的，并且最大值原理 (1.7) 成立。特别地，解一致有界：

$$\|\mu_t\|_{L^\infty} \le \max\{\sup \bar{\mu}, \sup \nu\} =: M \qquad \forall t \in [0,\infty).$$

我们将剩余结论的证明分为四个步骤。

步骤 1： 在这一步中，我们证明 $\|\mu_t - \nu\|_{\dot{H}^{-1}} \to 0$ 当 $t \to \infty$。结合一致的 L^∞-界，这意味着弱-$*$ 收敛到 L^∞ 中的目标，并完成第 i) 点的证明。我们称

$$\mathscr{P}_M(\mathbb{T}^d) := \{\eta \in \mathscr{P} \cap L^\infty(\mathbb{T}^d) : \|\eta\|_{L^\infty} \le M\}.$$

我们首先断言

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \qquad \|\eta - \nu\|_{\dot{H}^{-1}} \ge \varepsilon \implies \int_{\mathbb{T}^d} |\nabla K_1 * (\eta - \nu)|^2 \eta \ge \delta \qquad \forall \eta \in \mathscr{P}_M(\mathbb{T}^d). \tag{3.3}$$

反设存在 $\varepsilon > 0$ 和 $\{\eta_j\}_{j\ge 0} \subset \mathscr{P}_M(\mathbb{T}^d)$，使得

$$\|\eta_j - \nu\|_{\dot{H}^{-1}} \ge \varepsilon \quad \forall j \ge 0 \qquad \text{and} \qquad \int_{\mathbb{T}^d} |\nabla K_1 * (\eta_j - \nu)|^2 \eta_j \to 0 \quad \text{as } j \to \infty.$$

抽取一个子列后，我们有 $W_2(\eta_j,\ \eta_\infty) \to 0$ 当 $j \to \infty$ 对于某个 $\eta_\infty \in \mathscr{P}_M(\mathbb{T}^d)$。特别地，由于定理 A.3 以及一致有界 $\|\eta_j\|_{L^\infty} \le M$，我们得到 $\|\eta_j - \eta_\infty\|_{\dot{H}^{-1}} \to 0$。现在我们证明

$$\int_{\mathbb{T}^d} |\nabla K_1 * (\eta_\infty - \nu)|^2 \eta_\infty = 0. \tag{3.4}$$

利用 $(a+b)^2 \le 2a^2 + 2b^2$，我们可以给出上界

$$\int_{\mathbb{T}^d} |\nabla K_1 * (\eta_\infty - \nu)|^2 \eta_\infty \le \underbrace{\int_{\mathbb{T}^d} |\nabla K_1 * (\eta_\infty - \nu)|^2 (\eta_\infty - \eta_j)}_{I_j^1}$$
$$+ \underbrace{2\int_{\mathbb{T}^d} |\nabla K_1 * (\eta_\infty - \eta_j)|^2 \eta_j}_{I_j^2} + \underbrace{2\int_{\mathbb{T}^d} |\nabla K_1 * (\eta_j - \nu)|^2 \eta_j}_{I_j^3}.$$

我们有 $I_j^1 \to 0$ 因为 $|\nabla K_1 * (\eta_\infty - \nu)|^2 \in L^\infty$ 由定理 A.4 以及 $\eta_j \rightharpoonup \eta_\infty$ 在 L^2中。另一方面，$I_j^2 \le 2M\|\eta_j - \eta_\infty\|_{\dot{H}^{-1}}^2 \to 0$ 并且 $I_j^3 \to 0$ 根据假设。因此，我们有（3.4），特别地，$\nabla K_1 * (\eta_\infty - \nu)(x) = 0$ 对 η_∞- 几乎处处成立。$x \in \mathbb{T}^d$ 因此，由于索博列夫正则性，

$$\nu(x) - \eta_\infty(x) = \operatorname{div} \nabla K_1 * (\eta_\infty - \nu)(x) = 0 \qquad \text{for } \eta_\infty\text{-a.e. } x \in \mathbb{T}^d,$$

这进而意味着 η_∞ 与 ν 实际上相同，因为它们具有相同的总质量，且 $\eta_\infty \ll \mathscr{L}^d$。这是一个矛盾，因为 $\|\eta_\infty - \nu\|_{\dot{H}^{-1}} = \lim_{j\to\infty}\|\eta_j - \nu\|_{\dot{H}^{-1}} \ge \varepsilon > 0$，命题（3.3）得证。

现在注意到 $\mu_t \in \mathscr{P}_M(\mathbb{T}^d)$ 并且 $\|\mu_t - \nu\|_{\dot{H}^{-1}}$ 根据定理 2.5 随时间递减。若有

$$\varepsilon := \lim_{t\to\infty} \|\mu_t - \nu\|_{\dot{H}^{-1}}^2 > 0,$$

由能量耗散恒等式（1.6）可得

$$\frac{d}{dt}\|\mu_t - \nu\|_{\dot{H}^{-1}}^2 = -2\int_{\mathbb{T}^d} |\nabla K_1(\mu_t - \nu)|^2 \mu_t \le -2\delta \qquad \text{for a.e. } t \in (0,\infty),$$

其中 $\delta > 0$ 由断言（3.3）给出。这便得到了所需的矛盾。

步骤 2： 从现在起，我们假设 $\bar{\mu},\ \nu \ge \alpha > 0$。在这一步中，我们证明 i) 点。由（1.7）中的最大值原理可知

$$\inf \mu_t \ge \alpha \qquad \forall t \in [0,\infty).$$

特别地，（1.9）中的第一个不等式可由以下一般事实推出，该事实可由 Benamou-Brenier 公式导出（例如参见 [FG21，习题 A.16，第 137] 页）：

$$W_2(\eta_1,\eta_2) \le \alpha^{-1/2}\|\eta_1 - \eta_2\|_{\dot{H}^{-1}} \qquad \forall \eta_1,\eta_2 \in \mathscr{P}(\mathbb{T}^d),\quad 0 < \alpha \le \eta_1,\eta_2 \ll \mathscr{L}^d.$$

此外，利用能量耗散恒等式（1.6）可得

$$\frac{d}{dt}\|\mu_t - \nu\|_{\dot{H}^{-1}}^2 = -2\int_{\mathbb{T}^d} |\nabla K_1 * (\mu_t - \nu)|^2 \mu_t \le -2\alpha\|\mu_t - \nu\|_{\dot{H}^{-1}}^2 \qquad \forall t \in (0,\infty).$$

对该微分不等式进行积分，便可得到方程（1.9）中的指数衰减。

步骤 3： 接下来证明 ii）。根据定理 2.6，采用逼近论证，与例如定理 3.1的证明中所用的方法类似，只需考虑 $\bar{\mu},\ \nu \in C^\infty$ 的情形，其中$\mu \in C^\infty([0,\infty) \times \mathbb{T}^d)$（见定理 2.12）。令 $v_t = -\nabla K_1 * (\mu_t - \nu)$ 为解 μ 所生成的向量场。利用定理 A.4 和（1.9），我们得到关于 v_t 的以下一致有界：

$$\|v_t\|_{L^\infty} \le C_d M^{\frac{d}{d+1}} W_2(\mu_t,\nu)^{\frac{1}{d+1}} \le C\left(d,\alpha,M,\|\bar{\mu} - \nu\|_{\dot{H}^{-1}}\right) e^{-\frac{\alpha}{d+1}t} \qquad \forall t \in [0,\infty). \tag{3.5}$$

令 $X\colon [0,\infty)\times\mathbb{T}^d\to\mathbb{T}^d$ 为与向量场 v 相关联的流映射。由于 $\|X_t-X_r\|_{L^\infty}\le\int_r^t\|v_u\|_{L^\infty}du$，由（3.5）可推出 $(X_t)_{t\ge0}$ 在一致范数下是柯西列，当 $t\to\infty$时，因此 X_t 在 $t\to\infty$ 时一致收敛到某个连续映射 $X_\infty:\mathbb{T}^d\to\mathbb{T}^d$。更确切地说，

$$\|X_t-X_\infty\|_{L^\infty}\le\int_t^\infty\|v_r\|_{L^\infty}dr\le C\left(d,\alpha,M,\|\bar\mu-\nu\|_{\dot H^{-1}}\right)e^{-\frac{\alpha}{d+1}t}\qquad\forall t\in[0,\infty).\tag{3.6}$$

我们首先假设 $\nu\in C^{0,\beta}$ 其中某个 $\beta\in(0,1)$ 并证明 μ_t 随时间指数快速收敛到 ν，且收敛是在一致范数下进行的。事实上，控制了 $C^{0,\beta}$-范数后，由（3.6）可得

$$\|\nu\circ X_t-\nu\circ X_\infty\|_{L^\infty}\le\|\nu\|_{C^{0,\beta}}\|X_t-X_\infty\|_{L^\infty}^{\beta}\le C\left(d,\alpha,\beta,M,\|\bar\mu-\nu\|_{\dot H^{-1}},\|\nu\|_{C^{0,\beta}}\right)e^{-\kappa t},\tag{3.7}$$

其中 $\kappa=\alpha\beta/(d+1)$。现在我们证明 $\tilde\mu_t:=\mu_t\circ X_t$ 随时间指数快速一致收敛到 $\nu\circ X_\infty$。事实上，我们有

$$\begin{aligned}\frac{d}{dt}(\tilde\mu_t-\nu\circ X_\infty)&=-(\mu_t\operatorname{div}v_t)\circ X_t\\&=-\tilde\mu_t(\tilde\mu_t-\nu\circ X_t)=-\tilde\mu_t(\tilde\mu_t-\nu\circ X_\infty)+\tilde\mu_t(\nu\circ X_t-\nu\circ X_\infty).\end{aligned}$$

积分得

$$\tilde\mu_t-\nu\circ X_\infty=(\bar\mu-\nu\circ X_\infty)\exp\left(-\int_0^t\tilde\mu_r dr\right)+\int_0^t\tilde\mu_r(\nu\circ X_r-\nu\circ X_\infty)\exp\left(-\int_r^t\tilde\mu_u du\right)dr.$$

因此，由于 $\alpha\le\tilde\mu_t\le M$ 对所有时间均成立，并结合（3.7），可得

$$\|\tilde\mu_t-\nu\circ X_\infty\|_{L^\infty}\le C\left(d,\alpha,\beta,M,\|\bar\mu-\nu\|_{\dot H^{-1}},\|\nu\|_{C^{0,\beta}}\right)e^{-\kappa t}.\tag{3.8}$$

最后，结合（3.7）和（3.8），我们得到

$$\begin{aligned}\|\mu_t-\nu\|_{L^\infty}&=\|\tilde\mu_t-\nu\circ X_t\|_{L^\infty}\\&\le\|\tilde\mu_t-\nu\circ X_\infty\|_{L^\infty}+\|\nu\circ X_t-\nu\circ X_\infty\|_{L^\infty}\le C\left(d,\alpha,\beta,M,\|\bar\mu-\nu\|_{\dot H^{-1}},\|\nu\|_{C^{0,\beta}}\right)e^{-\kappa t},\end{aligned}$$

如所需。

现在假设 $\nu\in C(\mathbb{T}^d)$ 仅具有 Dini连续模，即存在一个连续凹非减函数 $\omega:[0,\infty)\to[0,\infty)$，使得

$$\omega(0)=0,\quad\int_0^1\frac{\omega(r)}{r}dr<\infty,\qquad\text{and}\qquad|\nu(x)-\nu(y)|\le\omega(|x-y|)\qquad\forall x,y\in\mathbb{T}^d.$$

那么，由 (3.6) 和变量代换可得

$$\begin{aligned}\int_0^\infty\|\nu\circ X_t-\nu\circ X_\infty\|_{L^\infty}dt&\le\int_0^\infty\omega\left(\|X_t-X_\infty\|_{L^\infty}\right)dt\\&\le\int_0^\infty\omega\left(C\left(d,\alpha,M\right)e^{-\frac{\alpha}{d+1}t}\right)dt\\&\lesssim_{d,\alpha,M}\int_0^{C(d,\alpha,M)}\frac{\omega(r)}{r}dr\lesssim_{d,s,M,\omega}1.\end{aligned}\tag{3.9}$$

由此，重复情形 $\nu\in C^{0,\beta}$ 上完全相同的步骤，我们得到 $\|\mu_t-\nu\|_{L^\infty}\to0$ 当 $t\to\infty$。**步骤 4：** 在这最后一步中，我们证明第 iii 点。设 $\gamma>d/2$。通过通常的逼近论证（例如，参见定理 2.14 的证明），我们可以将问题归结为在假设 $\bar\mu,\nu\in C^\infty(\mathbb{T}^d)$ 的情况下证明该结论，因此 $\mu\in C^\infty([0,\infty)\times\mathbb{T}^d)$。记 $\sigma_t=\mu_t-\nu$，并考虑能量估计（2.26）。在 $\|\sigma_t\|_{\dot H^{\gamma-1}}$ 与 $\|\sigma_t\|_{\dot H^{-1}}$ 之间进行插值，$\|\sigma_t\|_{\dot H^\gamma}$利用 Young 不等式和下界 $\nu\ge\alpha>0$，可将（2.26）右端的第二行估计为

$$-2\big(\min_{\mathbb{T}^d}\nu\big)\|\sigma_t\|_{\dot H^\gamma}^2+C\|\nu\|_{\dot H^{\gamma+1}}\|\sigma_t\|_{\dot H^{\gamma-1}}\|\sigma_t\|_{\dot H^\gamma}\le C(d,\gamma,\alpha,\|\nu\|_{\dot H^{\gamma+1}})\|\sigma_t\|_{\dot H^{-1}}^2.$$

因此，回想起 $\|\sigma_t\|_{\dot{H}^{-1}}$ 由（1.9）给出的指数衰减，能量估计（2.26）可简化为

$$\frac{d}{dt}\|\sigma_t\|_{\dot{H}^\gamma}^2 \le C\|\nabla^2(-\Delta)^{-1}\sigma_t\|_{L^\infty}\|\sigma_t\|_{\dot{H}^\gamma}^2 + Ce^{-2\alpha t},$$

进而对其积分可得

$$\begin{aligned}\|\sigma_t\|_{\dot{H}^\gamma}^2 &\le \|\sigma_0\|_{\dot{H}^\gamma}^2 \exp\left(C\int_0^t \|\nabla^2(-\Delta)^{-1}\sigma_r\|_{L^\infty}dr\right) + C\int_0^t e^{-2\alpha r}\exp\left(C\int_r^t \|\nabla^2(-\Delta)^{-1}\sigma_u\|_{L^\infty}du\right)\\ &\le \left(\|\sigma_0\|_{\dot{H}^\gamma}^2 + C(1-e^{-2\alpha t})\right)\exp\left(C\int_0^t \|\nabla^2(-\Delta)^{-1}\sigma_r\|_{L^\infty}dr\right).\end{aligned} \tag{3.10}$$

此外，由索伯列夫嵌入可知 $\nu \in C^1(\mathbb{T}^d)$。因此，我们可以利用第 ii) 点得到 σ_t 的 L^∞ 范数呈指数衰减：

$$\|\sigma_t\|_{L^\infty} \le C(d,\gamma,\alpha,\|\bar{\mu}\|_{H^\gamma},\|\nu\|_{H^{\gamma+1}})e^{-\frac{\alpha}{d+1}t} \qquad \forall t\in[0,\infty). \tag{3.11}$$

由定理A.5、（3.10）和（3.11），可得 L^∞范数下向量场梯度的一致有界估计如下：

$$\begin{aligned}\|\nabla^2(-\Delta)^{-1}\sigma_t\|_{L^\infty} &\le C(d,\gamma)\|\sigma_t\|_{L^\infty}\left(1+\log\left(1+\frac{\|\sigma_t\|_{\dot{H}^\gamma}}{\|\sigma_t\|_{L^\infty}}\right)\right)\\ &\le C(d,\gamma)\|\sigma_t\|_{L^\infty}\left(1+\log(1+\|\sigma_t\|_{L^\infty}^{-1})+\log(1+\|\sigma_t\|_{\dot{H}^\gamma})\right)\\ &\le C(d,\gamma,\alpha,\|\bar{\mu}\|_{H^\gamma},\|\nu\|_{\dot{H}^{\gamma+1}})e^{-\frac{\alpha}{d+1}t}\left(1+t+\int_0^t\|\nabla^2(-\Delta)^{-1}\sigma_r\|_{L^\infty}dr\right).\end{aligned} \tag{3.12}$$

由格朗沃尔不等式可得

$$\int_0^\infty \|\nabla^2(-\Delta)^{-1}\sigma_r\|_{L^\infty}dr \le C(d,\gamma,\alpha,\|\bar{\mu}\|_{H^\gamma},\|\nu\|_{\dot{H}^{\gamma+1}}),$$

特别地，由（3.10），

$$\|\sigma_t\|_{\dot{H}^\gamma} \le C(d,\gamma,\alpha,\|\bar{\mu}\|_{H^\gamma},\|\nu\|_{H^{\gamma+1}}). \tag{3.13}$$

此时，结合（3.13）中对 $\|\sigma_t\|_{\dot{H}^\gamma}$ 的一致有界性与（1.9）中 $\|\sigma_t\|_{\dot{H}^{-s}}$ 的指数衰减，利用索博列夫插值可推出

$$\|\nabla^2(-\Delta)^{-1}\sigma_t\|_{L^\infty},\ \|\sigma_t\|_{\dot{H}^{\gamma-1}} \le Ce^{-t/C}. \tag{3.14}$$

将 (3.14) 和 (3.13代入能量估计 (2.26)，我们得到

$$\frac{d}{dt}\|\sigma_t\|_{\dot{H}^\gamma}^2 \le (Ce^{-t/C}-2\alpha)\|\sigma_t\|_{\dot{H}^\gamma}^2 + Ce^{-t/C},$$

最后，对该微分不等式积分，可得

$$\begin{aligned}\|\sigma_t\|_{\dot{H}^\gamma}^2 &\le \|\sigma_0\|_{\dot{H}^\gamma}^2\exp\left(\int_0^t\left(Ce^{-r/C}-2\alpha\right)dr\right) + \int_0^t Ce^{-r/C}\exp\left(\int_r^t\left(Ce^{-u/C}-2\alpha\right)du\right)\\ &\le \left(\|\sigma_0\|_{\dot{H}^\gamma}^2 + C(1-e^{-t/C})\right)\exp\left(-2\alpha t + C(1-e^{-t/C})\right).\end{aligned}$$

对上述结果取平方根，我们便得到 $\|\sigma_t\|_{\dot{H}^\gamma}$ 所需的衰减率，从而完成证明。 □

3.1.3 初始测度无下界时的指数收敛 e

在本节中，我们证明假设 $\bar{\mu} \ge \alpha > 0$ 可以放宽，同时仍能得到向目标的指数弱收敛。关键在于下面的引理：只要目标有一致的下界，光滑解支撑中的“孔洞”就会以指数速率“填满”。

引理 3.2 (孔洞的指数填充)。设 $\bar{\mu}$, $\nu \in \mathscr{P} \cap C^\infty(\mathbb{T}^d)$，并假设 $0 < \lambda \le \nu \le \Lambda < \infty$属于 $\mathbb{T}^d$。令 $\mu \in C^\infty([0,\infty) \times \mathbb{T}^d)$ 为方程 (1.4) 的相应解。则有：

$$|\{\mu_t \le a\}| \le |\{\bar{\mu} \le a\}| e^{-(\lambda - a)t} \quad \forall a \in [0, \lambda]; \qquad |\{\mu_t \ge b\}| \le |\{\bar{\mu} \ge b\}| e^{-(b-\Lambda)t} \quad \forall b \in [\Lambda, +\infty).$$

证明。我们使用特征线法。回想一下， $\tilde{\mu}_t := \mu_t \circ X_t$ 满足

$$\frac{d}{dt}\tilde{\mu}_t(x) = \tilde{\mu}_t(x)(\nu \circ X_t(x) - \tilde{\mu}_t(x)).$$

由此方程，考虑到 $\lambda \le \nu \le \Lambda$，我们推得

$$\{\tilde{\mu}_t \le a\} \subseteq \{\tilde{\mu}_r \le a\} \quad \forall a \le \lambda, \forall r \le t; \qquad \{\tilde{\mu}_t \ge b\} \subseteq \{\tilde{\mu}_r \ge b\} \quad \forall b \ge \Lambda, \forall r \le t. \tag{3.15}$$

对于给定的 $a \in [0, \lambda]$，我们有 $\{\mu_t \le a\} = X_t(\{\tilde{\mu}_t \le a\})$。特别地，通过变量替换公式，我们得到

$$|\{\mu_t \le a\}| = |X_t(\{\tilde{\mu}_t \le a\})| = \int_{\{\tilde{\mu}_t \le a\}} \det \nabla X_t(x).$$

现在，令 $x \in \{\tilde{\mu}_t \le a\}$。由（3.15），我们有 $\tilde{\mu}_r(x) \le a$ 对所有 $r \in [0,t]$ 均成立。因此，

$$\det \nabla X_t(x) = \exp\left(\int_0^t \operatorname{div} v_r \circ X_r(x) dr\right) = \exp\left(\int_0^t (\tilde{\mu}_r(x) - \nu \circ X_r(x)) dr\right) \le e^{-(\lambda - a)t}.$$

再次利用（3.15），最终得到

$$|\{\mu_t \le a\}| \le |\{\tilde{\mu}_t \le a\}| e^{-(\lambda-a)t} \le |\{\tilde{\mu}_0 \le a\}| e^{-(\lambda-a)t} = |\{\bar{\mu} \le a\}| e^{-(\lambda-a)t},$$

如所需。当 $b \ge \Lambda$ 时，对上水平集 $\{\mu_t \ge b\}$ 也可作类似论证。 □

定理 3.2 表明，只需假设目标测度有下界，即可指数弱收敛至平衡态。

命题 3.3。设 $\bar{\mu}$, $\nu \in \mathscr{P} \cap L^\infty(\mathbb{T}^d)$，并假设 $\nu \ge \alpha > 0$。令 $\mu \in L^\infty([0,\infty); L^\infty(\mathbb{T}^d))$ 为方程 (1.4)的相应解。则

$$\|\mu_t - \nu\|_{\dot{H}^{-1}} \le e^{-\frac{\alpha t}{2}} \left(\|\bar{\mu} - \nu\|_{\dot{H}^{-s}} + C(d)\left(\|\bar{\mu}\|_{L^\infty} + \|\nu\|_{L^\infty}\right)\left(e^{\frac{\alpha t}{4}} - 1\right)\right) \qquad \forall t \in [0,\infty).$$

证明。 根据初始测度与目标测度上的通常逼近论证，我们可以假设 $\bar{\mu}$, $\nu \in \mathscr{P} \cap C^\infty(\mathbb{T}^d)$。利用能量耗散恒等式（1.6），我们可以得到以下界

$$\begin{aligned}
\frac{d}{dt}\|\mu_t - \nu\|_{\dot{H}^{-1}}^2 &= -2\int_{\mathbb{T}^d} |\nabla K_1 * (\mu_t - \nu)|^2 \mu_t \le -\alpha \int_{\{\mu_t > \alpha/2\}} |\nabla K_1 * (\mu_t - \nu)|^2 \\
&= -\alpha \int_{\mathbb{T}^d} |\nabla K_1 * (\mu_t - \nu)|^2 + \alpha \int_{\{\mu_t \le \alpha/2\}} |\nabla K_1 * (\mu_t - \nu)|^2 \\
&\le -\alpha \|\mu_t - \nu\|_{H^{-1}}^2 + \alpha \|\nabla K_1 * (\mu_t - \nu)\|_{L^\infty}^2 |\{\mu_t \le \alpha/2\}|.
\end{aligned}$$

根据定理 A.4 和最大值原理（1.7）， $\|\nabla K_1 * (\mu_t - \nu)\|_{L^\infty} \lesssim_d \|\bar{\mu}\|_{L^\infty} + \|\nu\|_{L^\infty}$。此外，定理 3.2 保证 $|\{\mu_t \le \alpha/2\}| \le e^{-\alpha t/2}$。因此，我们得到

$$\frac{d}{dt}\|\mu_t - \nu\|_{\dot{H}^{-1}}^2 \le -\alpha \|\mu_t - \nu\|_{H^{-1}}^2 + \alpha C(d)\left(\|\bar{\mu}\|_{L^\infty} + \|\nu\|_{L^\infty}\right)^2 e^{-\frac{\alpha t}{2}},$$

积分后即可得到所需的衰减。 □

3.2 情形 $s>1$。

本节旨在证明定理1.4。如第1.5节所述，该论证依赖于三个主要要素： H^{-s} 能量耗散恒等式（1.6）、 H^{γ} 定理2.13中的能量估计，以及经典索博列夫插值。结合这些要素，在初始数据满足较小偏差假设的条件下，我们能够在动力学过程中建立适当的 Łojasiewicz 梯度不等式。

定理的证明*1.4*。首先注意，只需证明在极大存在区间 $[0,T)$内的所有时刻，估计式（1.11）成立 $T=\infty$；根据定理 1.1 中的延拓判据，这便自动推出 $T=\infty$。其次，根据常见的逼近论证（例如参见定理 2.14 的证明），我们可以将证明（1.11）归结为光滑数据 $\bar{\mu}$， $\nu\in C^{\infty}(\mathbb{T}^d)$ 的情形；此时 $\mu\in C^{\infty}([0,T)\times\mathbb{T}^d)$ 由定理 2.12 得证。

记 $\sigma_t:=\mu_t-\nu$，并考虑式（2.26）中的能量估计。与定理 1.2 证明中的步骤 4 类似，我们可以在 $\|\sigma_t\|_{\dot{H}^{\gamma-s}}$ 与 $\|\sigma_t\|_{\dot{H}^{-s}}$ 之间对 $\|\sigma_t\|_{\dot{H}^{\gamma-s+1}}$进行插值，并结合下界 $\nu\geq\alpha>0$ 使用 Young 不等式，得到

$$-2\big(\min_{\mathbb{T}^d}\nu\big)\|\sigma_t\|^2_{\dot{H}^{\gamma-s+1}}+C\|\nu\|_{\dot{H}^{\gamma+s}}\|\sigma_t\|_{\dot{H}^{\gamma-s}}\|\sigma_t\|_{\dot{H}^{\gamma-s+1}}\leq C(d,s,\gamma,\alpha,\|\nu\|_{\dot{H}^{\gamma+s}})\|\sigma_t\|^2_{\dot{H}^{-s}}.$$

于是，(2.26)可简化为

$$\frac{d}{dt}\|\sigma_t\|^2_{\dot{H}^{\gamma}}\leq C\|\nabla^2(-\Delta)^{-s}\sigma_t\|_{L^{\infty}}\|\sigma_t\|^2_{\dot{H}^{\gamma}}+C\|\sigma_t\|^2_{\dot{H}^{-s}},$$

对其积分可得

$$\begin{aligned}\|\sigma_t\|^2_{\dot{H}^{\gamma}}\leq{}&\|\sigma_0\|^2_{\dot{H}^{\gamma}}\exp\left(C\int_0^t\|\nabla^2(-\Delta)^{-s}\sigma_r\|_{L^{\infty}}dr\right)\\&+C\int_0^t\|\sigma_r\|^2_{\dot{H}^{-s}}\exp\left(C\int_r^t\|\nabla^2(-\Delta)^{-s}\sigma_u\|_{L^{\infty}}du\right)dr.\end{aligned}\tag{3.16}$$

另一方面，能量耗散恒等式(1.6)结合对 $\|\sigma_t\|_{\dot{H}^{-s}}$在 $\|\sigma_t\|_{\dot{H}^{1-2s}}$ 与 $\|\sigma_t\|_{\dot{H}^{\gamma}}$ 之间的索博列夫插值，得到

$$\begin{aligned}\frac{d}{dt}\|\sigma_t\|^2_{\dot{H}^{-s}}&=-2\int_{\mathbb{T}^d}|\nabla(-\Delta)^{-s}\sigma_t|^2\mu_t\\&\leq-2\big(\inf_{\mathbb{T}^d}\mu_t\big)\|\sigma_t\|^2_{\dot{H}^{1-2s}}\leq-2\big(\inf_{\mathbb{T}^d}\mu_t\big)\|\sigma_t\|^{-\frac{2s-2}{\gamma+s}}_{\dot{H}^{\gamma}}\|\sigma_t\|^{2+\frac{2s-2}{\gamma+s}}_{\dot{H}^{-s}}.\end{aligned}\tag{3.17}$$

此时，我们注意到以下几点。一方面，对 $\|\sigma_t\|_{\dot{H}^{\gamma}}$的一致上界，以及对所有 $t\in[0,T)\inf\ \mu_t$的一致下界，通过对 (3.17)积分，即可得到 $\|\sigma_t\|_{\dot{H}^{-s}}$所需的多项式衰减。另一方面， $\|\sigma_t\|_{\dot{H}^{-s}}$ 衰减得足够快（时间可积）意味着，根据 (3.16)， $\|\sigma_t\|_{\dot{H}^{\gamma}}$ 一致有界。在证明的其余部分，我们将证明，只要充分小地选取 $\|\sigma_0\|_{\dot{H}^{-s}}$，这些界在 $[0,T)$上确实一致成立。由此得到 (1.11)。

代入 $M:=\|\sigma_0\|^2_{\dot{H}^{\gamma}}$，根据上述启发式分析，我们需要考虑

$$\tau:=\sup\left\{t>0:\|\sigma_r\|^2_{\dot{H}^{\gamma}}\leq 2M\quad\text{and}\quad\inf_{\mathbb{T}^d}\mu_r\geq\alpha/2,\ \forall r\in[0,t)\right\}\in(0,T].$$

我们的目标是证明，若 $\|\sigma_0\|_{\dot{H}^{-s}}$足够小，则 $\tau=T$。由（3.17）及 τ的定义可得

$$\frac{d}{dt}\|\sigma_t\|^2_{\dot{H}^{-s}}\leq-\kappa\|\sigma_t\|^{2+\frac{2s-2}{\gamma+s}}_{\dot{H}^{-s}}\qquad\forall t\in[0,\tau),$$

其中，为方便起见，我们设 $\beta=(\gamma+s)/(s-1)$ 和 $\kappa:=\alpha(2M)^{-1/\beta}$。积分后得到

$$\|\sigma_t\|^2_{\dot{H}^{-s}}\leq\|\sigma_0\|^2_{\dot{H}^{-s}}(1+Kt)^{-\frac{\gamma+s}{s-1}}\qquad\forall t\in[0,\tau),\tag{3.18}$$

其中 $K := \alpha\beta^{-1}(\|\sigma_0\|^2_{\dot H^{-s}}/2\|\sigma_0\|^2_{\dot H^\gamma})^{1/\beta}$。此外，记 $\eta := \gamma - d/2 > 0$ 和 $\hat\gamma := d/2 + 2 - 2s + \eta/2$，$\nabla^2(-\Delta)^{-s}$ 的有界性：$\dot H^{\hat\gamma} \to L^\infty$，结合索博列夫插值，可得

$$\|\nabla^2(-\Delta)^{-s}\sigma_t\|_{L^\infty} \le C\|\sigma_t\|_{\dot H^{\hat\gamma}} \le C\|\sigma_t\|^{1-\theta}_{\dot H^{-s}}\|\sigma_t\|^{\theta}_{\dot H^\gamma}, \qquad \theta = 1 - \frac{2(s-1)+\eta/2}{\gamma+s}.$$

因此，由（3.18）可得

$$\|\nabla^2(-\Delta)^{-s}\sigma_t\|_{L^\infty} \le C(2M)^\theta\|\sigma_0\|^{1-\theta}_{\dot H^{-s}}(1+Kt)^{-1-\frac{\eta}{4(s-1)}} \qquad \forall t\in[0,\tau),$$

并在时间区间 0 到 τ 上积分，得到

$$\begin{aligned}\int_0^\tau \|\nabla^2(-\Delta)^{-s}\sigma_r\|_{L^\infty}dr &\le C(2M)^\theta\|\sigma_0\|^{1-\theta}_{\dot H^{-s}}\int_0^\infty (1+Kr)^{-1-\frac{\eta}{4(\gamma+s)}}dr \\ &\le C(2M)^\theta\frac{\|\sigma_0\|^{1-\theta}_{\dot H^{-s}}}{K} \le C(d,s,\gamma,M,\alpha)\|\sigma_0\|^{\frac{\eta}{2(\gamma+s)}}_{\dot H^{-s}}.\end{aligned} \tag{3.19}$$

由此，我们推得关于 μ_t 的一致下界。事实上，由于 $\inf \bar\mu \ge \alpha$，且输运 μ_t 的向量场的梯度恰为 $-\nabla^2(-\Delta)^{-s}\sigma_t$，

$$\begin{aligned}\inf_{\mathbb{T}^d}\mu_t &\ge \left(\inf_{\mathbb{T}^d}\bar\mu\right)\exp\left(-C\int_0^t\|\nabla^2(-\Delta)^{-s}\sigma_r\|_{L^\infty}dr\right) \\ &\ge \alpha\exp\left(-C(d,s,\gamma,M,\alpha)\|\sigma_0\|^{\frac{\eta}{2(\gamma+s)}}_{\dot H^{-s}}\right)\end{aligned} \qquad \forall t\in[0,\tau). \tag{3.20}$$

现在，我们通过证明：若 $\|\sigma_0\|^2_{\dot H^{-s}} \le \delta$ 且 $\delta > 0$ 足够小，则 $\tau = T$，从而完成证明。结合（3.18），可证明（1.11）中的估计对所有时刻 $t \in [0,T)$ 均成立，进而推出 $T = \infty$，正如证明开头所指出的。反设 $\tau \in (0,T)$。由（3.20），取 δ 足够小，即可确保对每个$\inf_{\mathbb{T}^d}\mu_t \ge 3\alpha/4$ 均有$t\in[0,\tau)$，因此必有 $\|\sigma_\tau\|^2_{\dot H^\gamma} = 2M$。然而，（3.16）结合（3.18）和（3.19）给出

$$\begin{aligned}\|\sigma_\tau\|^2_{\dot H^\gamma} &\le M\exp\left(C\int_0^\tau\|\nabla^2(-\Delta)^{-s}\sigma_t\|_{L^\infty}dt\right) \\ &\quad + C\int_0^\tau\|\sigma_t\|^2_{\dot H^{-s}}\exp\left(\int_t^\tau C\|\nabla^2(-\Delta)^{-s}\sigma_r\|_{L^\infty}dr\right) \\ &\le M\exp\left(C\|\sigma_0\|^{\frac{\eta}{2(\gamma+s)}}_{\dot H^{-s}}\right)\left(1+\frac{\|\sigma_0\|^2_{\dot H^{-s}}}{M}\int_0^\infty(1+Kt)^{-\frac{\gamma+s}{s-1}}dt\right) \le \frac{3}{2}M,\end{aligned}$$

前提是 $\|\sigma_0\|_{\dot H^{-s}} \le \delta$ 取得足够小。这就产生矛盾，并完成证明。 □

备注 3.4。定理 $\nu \in H^{\gamma+s}(\mathbb{T}^d)$ 中的假设 1.4 仅仅是技术性的，且依赖于我们在第 A.2 节采用的特定版本的 Kato-Ponce 对易子估计。例如，如果我们只使用整数阶导数（即 $s,\gamma\in\mathbb{N}$，这样便可以应用莱布尼茨乘积法则），就可以将这一假设弱化为 $\nu \in H^{\gamma+1}(\mathbb{T}^d)$（参见下文第 1.7 节定理的证明，见第 4.4 节）。此外，如果我们能得到环面上涉及分数阶导数的完整 Kato-Ponce 估计（其形式与 [Li19类似），那么定理 1.2],中的论证还可以进一步推进，只需假设 $\nu \in H^{\gamma+\varepsilon}(\mathbb{T}^d)$ 对某个 $\varepsilon > 0$ 成立。■

4 连续浅层神经网络的收敛

本节将在第1.3节的框架下，将目前为止针对 Riesz KMD 流所发展的理论推广到 ReLU 神经网络的情形。首先，我们说明如何将一般形式下的 Wasserstein 梯度流（1.14），从第1.3节的框架化简为球面上的 Wasserstein–Fisher–Rao 动力学（1.15）。接着，在第4.2 节中，我们研究与反余弦核相关的半正定算子的谱性质。在第4.3节中，我们证明（1.15）在测度类中的全局适定性。最后，在第4.4 节中，我们证明定理 1.7中的局部定量收敛结论。

4.1 化简为 Wasserstein–Fisher–Rao 流

下面说明如何利用问题的对称性，在表达能力仅有少许损失的前提下，将（1.14）化简为在 $\mathbb{S}^d$ 上的等价动力学，其与（1.4）有许多相似之处。我们将分几步说明这一化简（接续第1.3节中的讨论，并使用该节引入的记号）：

从 $\mathbb{R}^{d+2}$ 到 $\sqrt{2}\mathbb{S}^{d+1}$。 利用 2 次齐次性 $\Phi(w,x)$ 关于变量 w，我们可以等价地将式 f_μ （1.13）中的函数表示为关于测度 $\tilde{\mu} \in \mathcal{M}_+$（$\sqrt{2}\mathbb{S}^{d+1}$）的形式；该测度由将 μ 投影到 $\sqrt{2}\mathbb{S}^{d+1}$ 上并在径向变量上赋予二次权重得到：

$$
\begin{aligned}
f_\mu(x) &= \int_{\mathbb{R}^{d+2}} \Phi(w,x) d\mu(w) = \int_{\sqrt{2}\mathbb{S}^{d+1}} \Phi(\theta,x) d\tilde{\mu}(\theta) =: f_{\tilde{\mu}}(x), \\
\int_{\sqrt{2}\mathbb{S}^{d+1}} \psi d\tilde{\mu} &:= \int_{\mathbb{R}^{d+2}} \frac{|w|^2}{2} \psi\left(\sqrt{2}\frac{w}{|w|}\right) d\mu \qquad \forall \psi \in C(\sqrt{2}\mathbb{S}^d).
\end{aligned}
$$

将 μ_t、ν和 $\bar{\mu}$ 分别替换为其在 $\sqrt{2}\mathbb{S}^{d+1}$上的投影后，式（1.14）写为

$$
\left\{
\begin{array}{rcll}
\partial_t \mu_t + \mathrm{div}_{\sqrt{2}\mathbb{S}^{d+1}}(\mu_t v_t) &=& -4\hat{\mathcal{K}}(\mu_t - \nu)\mu_t & \text{in } (0,T)\times\sqrt{2}\mathbb{S}^{d+1}, \\
v_t &=& -\nabla_{\sqrt{2}\mathbb{S}^{d+1}} \hat{\mathcal{K}}(\mu_t - \nu) & \forall t \in (0,T), \\
\mu_0 &=& \bar{\mu}, &
\end{array}
\right. \tag{4.1}
$$

从 $\sqrt{2}\mathbb{S}^{d+1}$ 到 $\{\pm 1\}\times\mathbb{S}^d$。为了进一步化简，我们注意到，任意函数 f_μ，$\mu \in \mathcal{M}_+(\sqrt{2}\mathbb{S}^{d+1})$ 都可以等价地表示为某个 $f_{\mu'}$，其中 $\mu' \in \mathcal{M}_+(\sqrt{2}\mathbb{S}^{d+1})$ 且 $\mu' \subset \{w = (a,b) \in \mathbb{R}\times\mathbb{R}^{d+1} : |a| = |b|\}$。这是由于 $\Phi(w,x)$ 分别关于两个变量 $a \in \mathbb{R}$ 和 $b \in \mathbb{R}^{d+1}$ 具有一次齐次性。此外，动力学 (4.1) 在支撑于锥 $\{|a| = |b|\}$ 中的测度集合内是闭合的；只需证明，当μ_t 和 ν 支撑于该处时，向量场 v_t 与 $\{|a| = |b|\}$ 相切即可。基于这些原因，我们可以将问题化简为 $\bar{\mu},\nu \in \mathcal{M}_+(\sqrt{2}\mathbb{S}^{d+1})$ 是支撑于 $\{|a| = |b|\}$ 中的非负有限测度的情形，因此，(4.1) 的解 $\mu_t \in \mathcal{M}_+(\sqrt{2}\mathbb{S}^{d+1})$ 也将支撑于 $\{|a| = |b|\}$ 中。注意，$\sqrt{2}\mathbb{S}^{d+1}$ 与锥 $\{|a| = |b|\}$ 的交集由 $\mathbb{S}^d$的两个不相交副本组成：

$$
\sqrt{2}\mathbb{S}^{d+1} \cap \{|a| = |b|\} = \{\pm 1\}\times\mathbb{S}^d.
$$

将 μ_t 与一对测度$(\mu_t^+, \mu_t^-) \in \mathcal{M}_+(\mathbb{S}^d)^2$, 类似地对 $\bar{\mu}$ 和 ν进行识别，我们可以将（4.1）写成球面上两个物种的非局部受迫连续性方程 $\mathbb{S}^d$:

$$
\left\{
\begin{array}{rcll}
\partial_t \mu_t^+ + \mathrm{div}_{\mathbb{S}^d}\left(\mu_t^+ v_t\right) &=& -4\mathcal{K}\left((\mu_t^+ - \mu_t^-) - (\nu^+ - \nu^-)\right)\mu_t^+ & \text{in } (0,T)\times\mathbb{S}^d, \\
\partial_t \mu_t^- - \mathrm{div}_{\mathbb{S}^d}\left(\mu_t^- v_t\right) &=& +4\mathcal{K}\left((\mu_t^+ - \mu_t^-) - (\nu^+ - \nu^-)\right)\mu_t^- & \text{in } (0,T)\times\mathbb{S}^d, \\
v_t &=& -\nabla_{\mathbb{S}^d}\mathcal{K}\left((\mu_t^+ - \mu_t^-) - (\nu^+ - \nu^-)\right) & \forall t \in (0,T), \\
\mu_0^+ &=& \bar{\mu}^+, & \\
\mu_0^- &=& \bar{\mu}^-, &
\end{array}
\right. \tag{4.2}
$$

其中 $\mathcal{K}$ 是（1.16）中定义的算子。

从 $\{\pm 1\}\times\mathbb{S}^d$ 到 $\mathbb{S}^d$。 最后，我们证明，在目标函数 $\bar{\mu}^- = \nu^- = 0$ 具有足够正则性的假设下，可以在（$\mu_t^- = 0$）中取（因此也取 f_ν），从而化简为单个物种的演化。在前一步的识别 $\nu = (\nu^+, \nu^-) \in \mathcal{M}_+(\mathbb{S}^d)^2$ 下，我们有$f_\nu = \mathcal{H}(\nu^+ - \nu^-)$，其中

$$
\mathcal{H}(\eta)(x) := \int_{\mathbb{S}^d} (x\cdot y)_+ d\eta(y) \qquad \forall \eta \in \mathcal{M}(\mathbb{S}^d). \tag{4.3}
$$

通过对球谐展开进行一些显式计算，可以证明如下结论（见定理 4.1）：

- 线性函数是球面上唯一能够表示为 $\mathcal{H}(\eta)$ 的奇函数，其中 $\eta \in \mathcal{M}(\mathbb{S}^d)$。特别地，我们能够表示为 f_ν 的所有可能函数，在球面上除加上某个线性函数外，都是偶函数。

- 对于所有 $\gamma \geq 0$，$\mathcal{H}$ 定义了一个从 $H^{\gamma}_{\text{偶函数}}(\mathbb{S}^d)$ 到 $H^{\gamma+s}_{\text{偶函数}}(\mathbb{S}^d)$ 的线性双射连续算子，其中$s=(d+3)/2$。因此，取 $\gamma > d/2$ 并利用索伯列夫嵌入 $H^{\gamma}(\mathbb{S}^d) \hookrightarrow L^{\infty}(\mathbb{S}^d)$，我们得到如下结论：所有函数 $f \in H^{\gamma+s}_{\text{偶函数}}(\mathbb{S}^d)$ 都可以表示为 $f = f_{\nu} + C_f$，其中某个 $\nu = (\nu^+, 0)$、ν^+ 为偶函数，以及某个常数 C_f 满足 $|C_f| \lesssim_{d,\,\gamma} \|f\|_{H^{\gamma+s}}$。

鉴于上述观察，我们可以取 $\bar{\mu}^- = \nu^- = 0$ 和 $\bar{\mu}^+$，ν^+ 为偶函数。这样做的代价是表达能力略有损失：我们只能表示球面上充分正则的偶函数，至多相差一个受控常数。分别将 $\bar{\mu}$、ν和 μ_t 等同于 $\bar{\mu}^+$、ν^+和μ_t^+，则（4.2）变为

$$\begin{cases} \partial_t \mu_t + \operatorname{div}_{\mathbb{S}^d}(\mu_t v_t) &= -4\mathcal{K}(\mu_t - \nu)\mu_t & \text{in } (0,T)\times\mathbb{S}^d, \\ v_t &= -\nabla_{\mathbb{S}^d}\mathcal{K}(\mu_t - \nu) & \forall t \in (0,T), \\ \mu_0 &= \bar{\mu}, \end{cases}$$

正如我们所要证明的。

4.2 球面上的分析与反余弦核算子

在本节中，我们回顾 d维单位球面 $\mathbb{S}^d \subset \mathbb{R}^{d+1}$上的若干分析概念。接下来，我们考虑反余弦核 K：$\mathbb{S}^d \times \mathbb{S}^d \to \mathbb{R}$ 以及相应的卷积算子 $\mathcal{K}$，其定义见（1.16），并得到后者精确的正则化性质。我们建议读者参阅 [Bac17，附录 D] 以及著作 [DX13]，以获得对该主题更广泛的介绍。

4.2.1 球谐函数与乘子算子

我们考虑一个标准正交基 由 $L^2(\mathbb{S}^d)$ 中拉普拉斯-贝尔特拉米算子 $-\Delta_{\mathbb{S}^d}$ 的特征函数构成 ：

$$\{Y_{kj} : k \geq 0,\ 1 \leq j \leq N(d,k)\}, \qquad -\Delta_{\mathbb{S}^d} Y_{kj} = \mu_k Y_{kj}, \quad \mu_k := k(k+d-1),$$

$$N(d,0) = 1, \qquad N(d,k) = \frac{2k+d-1}{k}\binom{k+d-2}{d-1} \quad \forall k \geq 1.$$

对于每个 $k \geq 0$和 $1 \leq j \leq N(d,\ k)$，Y_{kj} 是 $\mathbb{R}^{d+1}$中一个 k 次齐次调和多项式在球面上的限制。因此，Y_{kj} 通常被称为球谐函数。

对于每个分布 $f \in \mathscr{D}'(\mathbb{S}^d)$、每个 $k \geq 0$ 和 $1 \leq j \leq N(d,\ k)$，我们分别将 $\hat{}\, f_{kj} \in \mathbb{R}$ 和 $f_k \in C^{\infty}(\mathbb{S}^d)$ 定义为

$$\hat{f}_{kj} := \langle f, Y_{kj} \rangle \quad 1 \leq j \leq N(d,k), \qquad f_k := \sum_{j=1}^{N(d,k)} \hat{f}_{kj} Y_{kj} \qquad \forall k \geq 0.$$

每个函数 $f \in L^2(\mathbb{S}^d)$ 都可以按照球谐函数标准正交基展开：

$$f = \sum_{k=0}^{\infty}\sum_{j=1}^{N(d,k)} \hat{f}_{kj} Y_{kj} = \sum_{k=0}^{\infty} f_k, \qquad \|f\|^2_{L^2(\mathbb{S}^d)} = \sum_{k=0}^{\infty}\sum_{j=1}^{N(d,k)} |\hat{f}_{kj}|^2 = \sum_{k=0}^{\infty} \|f_k\|^2_{L^2(\mathbb{S}^d)}.$$

对于 $\gamma \in \mathbb{R}$，齐次 γ-索об列夫半范数定义为

$$\|f\|^2_{\dot{H}^{\gamma}(\mathbb{S}^d)} := \sum_{k=1}^{\infty} \mu_k^{\gamma} \|f_k\|^2_{L^2(\mathbb{S}^d)} \qquad \forall f \in \mathscr{D}'(\mathbb{S}^d),$$

其中，$\mu_k = k(k+d-1)$ 是拉普拉斯-贝尔特拉米算子 $-\Delta_{\mathbb{S}^d}$ 的第 k 个特征值。相应的齐次索об列夫空间 $\dot{H}^{\gamma}(\mathbb{S}^d)$ 定义如下：

$$\dot{H}^{\gamma}(\mathbb{S}^d) := \left\{ f \in \mathscr{D}'(\mathbb{S}^d) : f_0 = 0,\ \|f\|^2_{\dot{H}^{\gamma}(\mathbb{S}^d)} < \infty \right\}.$$

当 $\gamma \geq 0$ 时，我们还引入非齐次索оболев空间 $H^{\gamma}(\mathbb{S}^d)$：

$$H^{\gamma}(\mathbb{S}^d) := \left\{ f \in L^2(\mathbb{S}^d) : \|f\|_{H^{\gamma}(\mathbb{S}^d)} := \|f\|_{L^2(\mathbb{S}^d)} + \|f\|_{\dot{H}^{\gamma}(\mathbb{S}^d)} < \infty \right\}.$$

我们称 T 上的线性（可能无界）算子 $L^2(\mathbb{S}^d)$ 为乘子，如果存在一个序列$\{b_k\}_{k\geq 0}\subset\mathbb{R}$ 使得

$$(Tf)_k = b_k f_k \qquad \forall k\geq 0.$$

乘子算子的典型例子是拉普拉斯-贝尔特拉米算子的幂$(-\Delta_{\mathbb{S}^d})^\gamma$，其中 $\gamma\in\mathbb{R}$，对于这些算子 $b_k=\mu_k^\gamma$。注意， $f\in L^2(\mathbb{S}^d)$ 是偶函数（分别是奇函数）[6] 当且仅当 $f_k=0$ 对所有奇函数（分别是偶函数）$k\geq 0$都成立。特别地，乘子算子将偶函数（分别是奇函数）映射为偶函数（分别是奇函数）。

4.2.2 用角导数表示球面上的索оболев空间

对于每个 $1\leq i<j\leq d+1$，坐标 i， j 中的角导数定义如下：

$$\nabla_{(i,j)} := x_i\partial_i - x_j\partial_j = \partial_{\theta_{ij}}, \qquad (x_i,x_j)=r_{ij}(\cos\theta_{ij},\sin\theta_{ij}).$$

角导数的一些有用性质 如下列出 $\nabla_{(i,j)}$ （证明见 [DX13，第 1-3] 章）：

- 令 $\nabla_{\mathbb{S}^d}$ 和 $-\Delta_{\mathbb{S}^d}$ 分别表示 $\mathbb{S}^d$中的切向梯度和拉普拉斯-贝尔特拉米算子。则对于足够光滑的函数 f， g，以下恒等式成立： $\mathbb{S}^d\to\mathbb{R}$:

$$(\nabla_{\mathbb{S}^d}f)_j(x)=\sum_{\substack{1\leq i\leq d+1\\ i\neq j}} x_i\nabla_{(i,j)}f(x), \qquad \nabla_{\mathbb{S}^d}f\cdot\nabla_{\mathbb{S}^d}g=\sum_{1\leq i<j\leq d+1}\nabla_{(i,j)}f\nabla_{(i,j)}g,$$

$$\Delta_{\mathbb{S}^d}f=\sum_{1\leq i<j\leq d+1}\nabla_{(i,j)}^2 f.$$

- 角导数 $\nabla_{(i,j)}$ 使阶数为 k 的球谐函数空间保持不变，对所有 $k\geq 0$ 均成立。特别地， $\nabla_{(i,j)}$ 与任意乘子算子交换。
- 莱布尼茨法则成立： $\nabla_{(i,j)}(fg)=\nabla_{(i,j)}fg+f\nabla_{(i,j)}g$。
- 分部积分成立：$\int_{\mathbb{S}^d} f\nabla_{(i,j)}g=-\int_{\mathbb{S}^d}\nabla_{(i,j)}fg$。

给定多重指标 $\beta=\{\beta_{(i,j)}:\ 1\leq i<j\leq d+1\}$ 且 $\beta_{(i,j)}\in\mathbb{N}\cup\{0\}$，我们记作$|\beta|=\sum_{1\leq i<j\leq d+1}\beta_{(i,j)}$，并将相应的高阶角导数 ∇_β 定义为如下复合

$$\nabla_\beta := \bigcirc_{1\leq i<j\leq d+1}\nabla_{(i,j)}^{\beta_{(i,j)}}.$$

对每一个光滑函数 $f\in C^\infty(\mathbb{S}^d)$、 $\gamma\in\mathbb{N}\cup\{0\}$ 和 $p\in[1,\infty)$，角向索博列夫范数定义如下：

$$\|f\|_{\dot W^{\gamma,p}} := \left(\sum_{|\beta|=\gamma}\|\nabla_\beta f\|_{L^p}^p\right)^{1/p}, \qquad \|f\|_{W^{\gamma,p}} := \sum_{k=0}^{\gamma}\|f\|_{\dot W^{k,p}},$$

在 $p=\infty$ 的情况下，采用上确界范数的经典含义。空间 $W^{\gamma,p}(\mathbb{S}^d)$ 定义为 $C^\infty(\mathbb{S}^d)$ 关于 $\|\cdot\|_{W^{\gamma,p}}$ 的闭包。

如 [DX13第 3.5 节所证明的那样，1], 对每个 $1\leq i<j\leq d+1$，Riesz 变换 $R_{ij}:=\nabla_{(i,j)}(-\Delta_{\mathbb{S}^d})^{-1/2}$在 L^p 中有界，条件是 $1<p<\infty$。更准确地说，我们有

$$\max_{1\leq i<j\leq d+1}\|\nabla_{(i,j)}f\|_{L^p}\approx_{d,p}\|(-\Delta_{\mathbb{S}^d})^{1/2}f\|_{L^p} \qquad \forall f\in W^{1,p}(\mathbb{S}^d),\quad \forall p\in(1,\infty). \tag{4.4}$$

对（4.4）进行迭代应用可得

$$\|f\|_{\dot W^{\gamma,p}}\approx_{d,p,\gamma}\|(-\Delta_{\mathbb{S}^d})^{\gamma/2}f\|_{L^p} \qquad \forall f\in W^{\gamma,p}(\mathbb{S}^d),\quad \forall\gamma\in\mathbb{N}\cup\{0\},\quad \forall p\in(1,\infty).$$

在 $p=2$ 的情况下，我们有 $H^\gamma(\mathbb{S}^d)=W^{\gamma,2}(\mathbb{S}^d)$。此外，齐次 Sobolev 范数与角向索博列夫范数相同：

$$\|f\|_{\dot W^{\gamma,2}}^2=\sum_{|\beta|=\gamma}\|\nabla_\beta f\|_{L^2}^2=\int_{\mathbb{S}^d}f(-\Delta_{\mathbb{S}^d})^\gamma f=\|f\|_{\dot H^\gamma}^2 \qquad \forall f\in W^{\gamma,2}(\mathbb{S}^d),\quad \forall\gamma\in\mathbb{N}\cup\{0\}.$$

[6]这里对称性是 co 相对于原点进行考察。特别地， f 当 $f(x)=f(-x)$ 时为偶函数，当 $f(x)=-f(-x)$ 时为奇函数 .

4.2.3 球面上的 ReLU激活函数与反余弦核

本节研究与 ReLU 激活函数相关的表示算子 $\mathcal{H}$ （见（4.3））的谱性质。我们推导出，式（$\mathcal{K}=\mathcal{H}^2$ （1.16）中的半正定算子与$(-\Delta_{\mathbb{S}^d})^{-\frac{d+3}{2}}$ 相当，至少在作用于正偶频率时如此。稍后在第4.4 节中，我们将利用这一谱分析证明 ReLU 浅层神经网络的局部多项式定量收敛结果（见定理1.7）。

在下面的引理中，我们给出 $\mathcal{H}$ 在球谐函数中的谱分析，该结果由 [Bac17]得出：

引理 4.1 ([Bac17，附录 C.1 和 D。2])。 设 $\mathcal{H}$ 为(4.3)中定义的算子，并设 $s=(d+3)/2$。则 $\mathcal{H}$： $L^2(\mathbb{S}^d)\to L^2(\mathbb{S}^d)$ 是一个乘子算子，满足 $(\mathcal{H}f)_k=\beta_k f_k$，其中

$$(d=1)\qquad \beta_k:=\begin{cases}2, & k=0,\\ \frac{\pi}{2}, & k=1,\\ (-1)^{\frac{k}{2}-1}\frac{2}{k^2-1}, & k\in 2\mathbb{N},\\ 0, & k\in 2\mathbb{N}+1,\end{cases}$$

$$(d\geq 2)\qquad \beta_k:=\begin{cases}\frac{|\mathbb{S}^d|(d-1)}{2\pi d}, & k=0,\\ \frac{|\mathbb{S}^d|(d-1)\Gamma(\frac{d}{2})}{4\sqrt{\pi}\,d\,\Gamma(\frac{d+3}{2})}, & k=1,\\ (-1)^{\frac{k}{2}-1}\frac{|\mathbb{S}^d|(d-1)\Gamma(\frac{d}{2})}{2\pi}\frac{\Gamma(k-1)}{2^k\Gamma(\frac{k}{2})\Gamma(\frac{k+d}{2}+1)}, & k\in 2\mathbb{N},\\ 0, & k\in 2\mathbb{N}+1.\end{cases}$$

此外，对于所有 $k\in 2\mathbb{N}$，都有 $|\beta_k|\approx_d k^{-s}$ 。因此，对于所有 $\gamma\geq 0$， $\mathcal{H}$ 定义了一个从 $H^\gamma_{\mathrm{even}}(\mathbb{S}^d)$ 到 $H^{\gamma+s}_{\mathrm{even}}(\mathbb{S}^d)$的线性连续双射算子。

注4.2。 前述结论也可以由 Goodey–Weil 恒等式得到，该恒等式将球面上的余弦变换 $\mathcal{C}$ 与球面 Radon 变换 $\mathcal{R}$ 联系起来[7] $\mathbb{S}^d$ (参见 [GW92，命题 2。1]:

$$\mathcal{C}^{-1}=\frac{1}{2}(\Delta_{\mathbb{S}^d}+d)\mathcal{R}^{-1},$$

这是因为 $\mathcal{H}=\frac{1}{2}\mathcal{C}$ 作用于偶分布时。 ■

接下来，我们考虑对称核 K： $\mathbb{S}^d\times\mathbb{S}^d\to\mathbb{R}$ 及其对应的卷积算子 $\mathcal{K}$，定义见（1.16）。 K 通常称为“反余弦型核”，因为它可以表示为 $\mathbb{S}^d$ 上 x 与 y之间测地距离的函数，

$$\theta(x,y):=\arccos(x\cdot y)=\mathrm{dist}_{\mathbb{S}^d}(x,y)\in[0,\pi],\qquad x,y\in\mathbb{S}^d.$$

事实上，当s 例如， [CS09、CS11], 表明，存在一个正的常数 $c=c(d)>0$ 使得

$$K(x,y)=J(\theta(x,y))\quad\forall x,y\in\mathbb{S}^d,\qquad J(\theta):=c\,(\sin\theta+(\pi-\theta)\cos\theta)\quad\forall\theta\in[0,\pi].$$

函数 J： $[0,\pi]\to\mathbb{R}$ 光滑、非负、严格递减，并且当 $\theta\to 0^+$， π^- 时具有如下渐近行为：

$$J(\theta)=c\left(\pi-\frac{\pi}{2}\theta^2+\frac{1}{3}\theta^3\right)+o(\theta^3)\quad\text{as }\theta\to 0^+,\qquad J(\theta)=\frac{c}{3}(\pi-\theta)^3+o((\pi-\theta)^3)\quad\text{as }\theta\to\pi^-.$$

因此，在法坐标下可以证明 $K\in C^{2,\,1}(\mathbb{S}^d\times\mathbb{S}^d)$。特别地，对于（ $\mathcal{K}$ 1.16）中的卷积算子，有如下估计：

$$\|\mathcal{K}\eta\|_{C^{2,1}(\mathbb{S}^d)}\lesssim_d\|\eta\|_{\mathcal{M}(\mathbb{S}^d)}\qquad\forall\eta\in\mathcal{M}(\mathbb{S}^d).\tag{4.5}$$

[7]对于任意 $f\in C^\infty(\mathbb{S}^d)$，余弦变换和（球面）Radon 变换分别定义为

$$(\mathcal{C}f)(x)=\int_{\mathbb{S}^d}|x\cdot y|f(y)dy\qquad\text{and}\qquad(\mathcal{R}f)(x)=\int_{\mathbb{S}^d\cap x^\perp}f(v)d\sigma_{u^\perp}(v),\qquad\text{for}\quad x\in\mathbb{S}^d,$$

其中 $\mathbb{S}^{d-1}$ dy d $\mathbb{S}^d$ $d\sigma_{u^\perp}$ v $d-1$ $u^\perp\cap\mathbb{S}^d\cong$ 是上的-维曲面测度，且()表示上的()维曲面测度。

利用核 K 在（1.16）中的特殊结构，可以验证 $\mathcal{K}=\mathcal{H}^2$。这使得$\mathcal{K}$ 成为球谐函数中的正半定乘子算子，其谱行为至少在限制于正偶频率时，与（$-\Delta_{\mathbb{S}^d}$）$^{-\frac{d+3}{2}}$ 相当（参见 [BB21] 中的一些相关结果）：

引理 4.3。 设 $\mathcal{K}$ 为 (1.16)中定义的算子，并设 $s:=\frac{d+3}{2}$。以下结论成立：

i) $\mathcal{K}:\ L^2(\mathbb{S}^d)\to L^2(\mathbb{S}^d)$ 是一个乘子算子，满足 $(\mathcal{K}f)_k=\lambda_k f_k$，其中 $\lambda_k=\beta_k^2$ 和 β_k 由定理 *4.1*给出。特别地，$\lambda_k\approx_d k^{-2s}$ 对所有 $k\in 2\mathbb{N}$均成立，并且 $\mathscr{E}^{\nu}_{\mathbb{S}^d}$ 在(1.19) 中的近似恒等式成立。

ii) 乘子算子 $T:=(-\Delta_{\mathbb{S}^d})^s\mathcal{K}:\ L^2(\mathbb{S}^d)\to L^2(\mathbb{S}^d)$ 限制在零均值偶函数空间 $L^2_{2\mathbb{N}}(\mathbb{S}^d)$上时可逆，并且

$$\|Tf\|_{L^p}\approx_{d,p}\|f\|_{L^p}\qquad\forall f\in L^p\cap L^2_{2\mathbb{N}}(\mathbb{S}^d),\quad\forall p\in(1,\infty).$$

iii) 存在一个常数 $\bar{c}=\bar{c}(d)>0$ 使得

$$\left\|\left(\mathcal{K}-\bar{c}(-\Delta_{\mathbb{S}^d})^{-s}\right)(f)\right\|_{L^2}\lesssim_d\|f\|_{\dot{H}^{-2s-1}}\qquad\forall f\in L^2_{2\mathbb{N}}.$$

证明。 第 i) 点由定理 4.1 以及 $\mathcal{K}=\mathcal{H}^2$这一事实得出。

让我们证明第 ii）点。设 $\mu_k=k(k+d-1)$ 为 $-\Delta_{\mathbb{S}^d}$ 的特征值，则复合函数 $T=(-\Delta_{\mathbb{S}^d})^s\mathcal{K}$ 的符号为

$$(Tf)_k=c_kf_k,\qquad c_k=(k(k+d-1))^s\lambda_k\qquad\forall k\geq 0,$$

并且对所有 $k\in 2\mathbb{N}$，$c_k>0$。我们先考虑 $d\geq 2$ 的情形。令 $g:\ [2,\infty)\to\mathbb{R}$ 为函数

$$g(z)=\frac{|\mathbb{S}^d|^2(d-1)^2\Gamma\left(\frac{d}{2}\right)^2}{4\pi^2}\frac{(z(z+d-1))^{\frac{d+3}{2}}\Gamma(z-1)^2}{2^{2z}\Gamma\left(\frac{z}{2}\right)^2\Gamma\left(\frac{z+d}{2}+1\right)^2}$$

使得 $g(k)=c_k$ 对所有 $k\in 2\mathbb{N}$。得益于 [DX13、定理 3.3。1], 要证明第 ii) 点的内容，我们只需证明 $g(z)$ 在 $z\to\infty$ 时存在正极限，且满足以下 Mikhlin 型条件：

$$\left|\frac{d^j}{dz^j}g(z)\right|\lesssim z^{-j}\qquad\forall z\in[2,\infty),\quad\forall j\in\left\{0,\dots,\left\lceil\frac{d+1}{2}\right\rceil\right\}.\tag{4.6}$$

为证明这一点，可以使用以下形式的斯特林近似公式：

$$\Gamma(x)=\sqrt{2\pi}\,x^{x-\frac{1}{2}}e^{-x}e^{r(x)},\qquad r(x)=\sum_{n=2}^{\infty}\frac{a_n}{(x+1)\cdots(x+n-1)}.$$

这里的系数 $a_n\in\mathbb{R}$ 使得相应的级数对所有 $x>0$都收敛，这使得r（x）成为一个解析余项，并且当 $x\to\infty$ 时趋于零，同时满足 Mikhlin 条件（4.6）。将g（z）展开为 $z\to\infty$，利用上面的斯特林公式可得

$$g(z)=\bar{c}+O(z^{-1}),\qquad\bar{c}:=\frac{2^{d-3}(d-1)^2|\mathbb{S}^d|^2\Gamma\left(\frac{d}{2}\right)^2}{\pi^3}>0.$$

此外，可以将 g 改写为

$$\begin{aligned}g(z)=C(d)&\frac{z^{\frac{d+5}{2}}(z+d-1)^{\frac{d+3}{2}}}{(z-1)^3(z+d+2)^{d+1}}\times\\&\exp\left[-z\left(\log\left(1+\frac{1}{z-1}\right)+\log\left(1+\frac{d+3}{z-1}\right)\right)\right]\times\\&\exp\left[2\left(r(z-1)-r\left(\frac{z}{2}\right)-r\left(\frac{z+d}{2}+1\right)\right)\right].\end{aligned}$$

由于上式右侧的三个因子均满足 Mikhlin 条件 (4.6)，因此它们的乘积也满足该条件。在 $d=1$的情形下，利用 $g(z)=4z^4/(z^2-1)^2$，可用相同的论证。在这种情形下 $g(z)\to\bar{c}:=4$因为 $z\to\infty$。

最后，我们证明第 iii 点。令 $\bar{c}=\bar{c}(d)>0$ 为上述得到的 c_k 的极限。我们有

$$\lambda_k-\bar{c}(k(k+d-1))^{-s}=(g(k)-\bar{c})(k(k+d-1))^{-s}\lesssim_d k^{-2s-1}\qquad\forall k\in 2\mathbb{N},$$

最后，对于每个 $f\in L^2_{2\mathbb{N}}$,

$$\|(\mathcal{K}-\bar{c}(-\Delta_{\mathbb{S}^d})^{-s})(f)\|_{L^2}^2=\sum_{k\in 2\mathbb{N}}(\lambda_k-\bar{c}(k(k+d-1))^{-s})^2\|f_k\|_{L^2}^2\lesssim_d\sum_{k\in 2\mathbb{N}}|k|^{-4s-2}\|f_k\|_{L^2}^2\lesssim_d\|f\|_{H^{-2s-1}}^2,$$

如所需。 □

4.3 Wasserstein–Fisher–Rao 动力学的适定性

本节研究非守恒活性标量方程 (1.15) 的适定性，其中 $\bar{\mu},\ \nu\in\mathcal{M}_+(\mathbb{S}^d)$ 是两个给定的非负测度，而 $\mathcal{K}$ 是式 (1.16) 中定义的卷积算子。下面给出我们所考虑的弱解的精确定义：

定义 *4.4.* 设 $\bar{\mu},\nu\in\mathcal{M}_+(\mathbb{S}^d)$。若非负测度曲线 $\mu\in C_{w^*}([0,T);\mathcal{M}_+(\mathbb{S}^d))$ 满足 *(1.15)*，则称其为解；令 $v_t:=-\nabla_{\mathbb{S}^d}\mathcal{K}(\mu_t-\nu)$，受迫连续性方程 $\partial_t\mu_t+\mathrm{div}_{\mathbb{S}^d}(\mu_t v_t)=-4\mathcal{K}(\mu_t-\nu)\mu_t$ 在分布意义下成立，等价地

$$\int_{\mathbb{S}^d}\varphi d\mu_t=\int_{\mathbb{S}^d}\varphi d\bar{\mu}+\int_0^t\int_{\mathbb{S}^d}\nabla_{\mathbb{S}^d}\varphi\cdot v_r d\mu_r dr-4\int_0^t\int_{\mathbb{S}^d}\varphi\mathcal{K}(\mu_r-\nu)d\mu_r dr\qquad\forall t\in(0,T),\quad\forall\varphi\in C^\infty(\mathbb{S}^d).$$

根据定理4.4，(1.15)的解 μ 在测度空间中局部有界。因此，由(4.5)可知，它生成速度场 $v\in L^\infty_{\mathrm{loc}}([0,T);C^{1,1}(\mathbb{S}^d;\mathbb{R}^{d+1}))$. 特别地，相应的流映射 $X:[0,T)\times\mathbb{S}^d\to\mathbb{S}^d$是良定义的，并且根据标准柯西–利普希茨理论具有 $C^{1,1}$ 正则性。

将 $\varphi\equiv 1$ 作为测试函数代入上文(1.15)的分布形式，我们得到

$$\begin{aligned}\frac{d}{dt}\mu_t(\mathbb{S}^d)&=-4\int_{\mathbb{S}^d}\mathcal{K}(\mu_t-\nu)d\mu_t\\&=-4\int_{\mathbb{S}^d}\mathcal{K}(\mu_t)d\mu_t+4\int_{\mathbb{S}^d}\mathcal{K}(\nu)d\mu_t\le-4\lambda_0\mu_t(\mathbb{S}^d)^2+C(d)\nu(\mathbb{S}^d)\mu_t(\mathbb{S}^d),\end{aligned}$$

其中 $\lambda_0>0$ 是定理 4.3中的 $\mathcal{K}$ 第一特征值，并且我们使用了(4.5)。由这个微分不等式，我们可以推出动力学过程中解的质量具有如下的一致上界：

$$\mu_t(\mathbb{S}^d)\le\max\{\bar{\mu}(\mathbb{S}^d),C(d)\lambda_0^{-1}\nu(\mathbb{S}^d)\}\qquad\forall t\in[0,T).\tag{4.7}$$

在继续讨论 Wasserstein–Fisher–Rao 动力学的适定性结果之前，我们先回顾所谓的有界-利普希茨距离 $d_{BL}:\mathcal{M}(\mathbb{S}^d)\times\mathcal{M}(\mathbb{S}^d)\to[0,\infty)$在有限测度之间的定义：

$$d_{BL}(\mu_1,\mu_2):=\sup\left\{\int_{\mathbb{S}^d}f(d\mu_1-d\mu_2):\|f\|_{L^\infty}\le 1,\mathrm{Lip}(f)\le 1\right\}\qquad\forall\mu_1,\mu_2\in\mathcal{M}(\mathbb{S}^d).$$

在 [PR14] 中证明了 d_{BL} 可度量弱$*$ 收敛于 $\mathcal{M}(\mathbb{S}^d)$。根据 d_{BL} 的定义以及 $C^{2,\ 1}$-正则性，结合式（1.16）中的反余弦核，我们得到

$$\|\mathcal{K}(\mu_1-\mu_2)\|_{C^2(\mathbb{S}^d)}\lesssim_d d_{BL}(\mu_1,\mu_2)\qquad\forall\mu_1,\mu_2\in\mathcal{M}(\mathbb{S}^d).\tag{4.8}$$

本节余下部分将用于证明以下结果：

命题 4.5。 设 $\bar{\mu}$， $\nu \in \mathcal{M}_+(\mathbb{S}^d)$。根据定理 $\mu \in C_{w^*}([0,\infty); \mathcal{M}_+(\mathbb{S}^d))$ 存在唯一的全局解(1.15) *4.4*，且可表示为

$$\mu_t = (X_t)_{\#}(G_t\bar{\mu}), \qquad G_t(x) := \exp\left(-4\int_0^t \mathcal{K}(\mu_r - \nu) \circ X_r(x)dr\right) \qquad \forall t \in [0,\infty), \tag{4.9}$$

其中 $X: [0,\infty) \times \mathbb{S}^d \to \mathbb{S}^d$ 是与速度场 $v_t = -\nabla_{\mathbb{S}^d}\mathcal{K}(\mu_t - \nu)$ 相关联的流映射。此外，能量耗散恒等式 (1.18) 成立。此外:

i) （奇偶性保持）。若 $\bar{\mu}$、ν在 $\mathbb{S}^d$上为偶函数，则对所有 $t \in (0, \infty)$， μ_t都是偶函数 .

ii) （稳定性）。若 $\bar{\mu}^n$、$\nu^n \in \mathcal{M}_+(\mathbb{S}^d)$ 满足 $\bar{\mu}^n \overset{*}{\rightharpoonup} \bar{\mu}$ 且 $\nu^n \overset{*}{\rightharpoonup} \nu$，则相应的 μ_t^n解 (1.15) 在有界时间区间上一致地收敛于 μ_t ，且收敛所采用的距离为有界-利普希茨距离:

$$\sup_{t\in[0,T)} d_{BL}(\mu_t^n, \mu_t) \to 0 \qquad as\ n \to \infty, \qquad \forall T \in (0,\infty).$$

iii) （正则性传播）。若对于某个 $k \in \mathbb{N} \cup \{0\}$和 $\alpha \in (0,1]$,， $\bar{\mu}$、$\nu \in C^{k,\alpha}(\mathbb{S}^d)$ ，则 $\mu \in L^\infty_{\mathrm{loc}}$（$[0,\infty)$；$C^{k,\alpha}(\mathbb{S}^d)$。特别地，若 $\bar{\mu}$、$\nu \in C^\infty(\mathbb{S}^d)$ ，则 $\mu \in C^\infty([0,\infty) \times \mathbb{S}^d)$ 。

证明。 我们将证明分为三个步骤。

步骤 1: 首先证明，若存在一个解 $\mu \in C_{w^*}([0,T); \mathcal{M}_+(\mathbb{S}^d))$，则它必定具有（4.9）的形式。令 v为由 μ 和 $\mathcal{X}$生成的向量场: $[0,T)\times[0,T)\times\mathbb{S}^d \to \mathbb{S}^d$ 为与 v相关联的扩展流映射，其满足

$$\begin{cases} \frac{d}{dt}\mathcal{X}(t,s,x) &= v_t(\mathcal{X}(t,s,x)) & \forall s,t \in (0,T), \\ \mathcal{X}(s,s,x) &= x & \forall s \in [0,T). \end{cases}$$

注意，对于所有 $t \in [0,T)$，都有 $\mathcal{X}_{t,0} \equiv X_t$，且半群性质 $\mathcal{X}_{t,s} \circ \mathcal{X}_{s,r} = \mathcal{X}_{t,r}$ 对所有 $r,s,t \in [0,T)$ 均成立。根据推前的定义，可以证明如下定义的 $\hat{}\mu_t$ 是受迫连续性方程的另一个解；该方程由 μ 求解，且具有相同的速度场和强迫项:

$$\hat{\mu}_t := (\mathcal{X}_{t,0})_{\#}\bar{\mu} + \int_0^t (\mathcal{X}_{t,r})_{\#} g_r\mu_r dr \qquad g_t := -4\mathcal{K}(\mu_t - \nu).$$

于是，由线性性， $\sigma_t := \hat{}\mu_t - \mu_t$ 满足 $\partial_t\sigma_t + \mathrm{div}_{\mathbb{S}^d}(\sigma_t v_t) = 0$，其中 $\sigma_0 = 0$；根据一致 Lipschitz 向量场的标准理论（例如参见 [ABS21，第 16 讲，第 1] 节），可得 $\sigma_t \equiv 0$，即$\hat{}\mu_t = \mu_t$对每个 $t \in [0,T)$ 均成立。特别地，利用扩展流映射的半群性质以及推前算子的线性性:

$$\mu_t = (\mathcal{X}_{t,0})_{\#}\bar{\mu} + \int_0^t (\mathcal{X}_{t,r})_{\#} g_r\mu_r dr = (\mathcal{X}_{t,0})_{\#}\eta_t, \qquad \eta_t := \bar{\mu} + \int_0^t (\mathcal{X}_{0,r})_{\#} g_r\mu_r dr \qquad \forall t \in [0,T).$$

利用上述表达式中的恒等式 $(\mathcal{X}_{0,r})_{\#} g_r\mu_r = g_r \circ \mathcal{X}_{r,0}(\mathcal{X}_{0,r})_{\#}\mu_r = g_r \circ X_r\eta_r$，我们得到

$$\eta_t = \bar{\mu} + \int_0^t g_r \circ X_r\eta_r dr \qquad \forall t \in [0,T),$$

由此我们推出 $\eta_t = G_t\bar{\mu}$，其中 $G_t := \exp\left(\int_0^t g_r \circ X_r dr\right)$，最终对所有 $t \in [0,T)$ 有 $\mu_t = (X_t)_{\#}G_t\bar{\mu}$，即为所求。

由推前表示，能量耗散恒等式（1.18）可通过直接计算得到。

步骤 2: 现在我们证明解关于数据变化的稳定性，且在有界时间区间上是一致的。这同时处理了第 iii 点与唯一性。

对于 $i = 1, 2$，令 $\bar{\mu}^i$、$\nu^i \in \mathcal{M}_+(\mathbb{S}^d)$和 $\mu^i \in C_{w^*}([0,T); \mathcal{M}_+(\mathbb{S}^d))$为（1.15）的解，其中 $T \in (0, \infty)$。我们分别称 v_t^i 和 X_t^i为由解 μ^i生成的向量场和流映射，并且

$$g_t^i := -4\mathcal{K}(\mu_t^i - \nu^i), \qquad G_t^i := \exp\left(\int_0^t g_r^i \circ X_r^i dr\right) \qquad \forall t \in [0,T), \quad i = 1,2.$$

根据步骤 1，我们可以将这些解表示为

$$\mu_t^i = (X_t^i)_\# G_t^i \bar{\mu}^i \qquad \forall t \in [0,T), \quad i = 1,2. \tag{4.10}$$

我们将给出 $d_{BL}(\mu_t^1, \mu_t^2)$ 关于 $t \in [0,T)$ 的一致估计，从而得到第 iii 点中的稳定性，尤其是唯一性，当 $\bar{\mu}^1 = \bar{\mu}^2$ and $\nu^1 = \nu^2$。

我们首先需要一些预备估计。定义

$$M := \bar{\mu}^1(\mathbb{S}^d) + \bar{\mu}^2(\mathbb{S}^d) + \nu^1(\mathbb{S}^d) + \nu^2(\mathbb{S}^d).$$

于是，由（4.7）和（4.5）可得

$$\mu_t^i(\mathbb{S}^d),\ \|v_t^i\|_{C^{1,1}},\ \|\nabla X_t^i\|_{C^{0,1}},\ \|g_t^i\|_{C^{2,1}},\ \|G_t^i\|_{C^{2,1}} \le C(d,M,T), \qquad \forall t \in [0,T), \quad i = 1,2. \tag{4.11}$$

此外，由（4.8）还可得

$$\|g_t^1 - g_t^2\|_{L^\infty},\ \|v_t^1 - v_t^2\|_{L^\infty} \lesssim_d d_{BL}(\mu_t^1, \mu_t^2) + d_{BL}(\nu^1, \nu^2) \qquad \forall t \in [0,T). \tag{4.12}$$

对这些估计积分后，我们得到

$$\|G_t^1 - G_t^2\|_{L^\infty},\ \|X_t^1 - X_t^2\|_{L^\infty} \lesssim_{d,M,T} \int_0^t d_{BL}(\mu_r^1, \mu_r^2)dr + d_{BL}(\nu^1, \nu^2) \qquad \forall t \in [0,T). \tag{4.13}$$

为了估计 $d_{BL}(\mu_t^1, \mu_t^2)$ 关于 $t \in [0,T)$，我们取一个利普希茨函数 $f\colon \mathbb{S}^d \to \mathbb{R}$ 满足 $\|f\|_{L^\infty},\ \mathrm{Lip}(f) \le 1$，并利用表示公式（4.10）以及（4.11）、（4.12）和（4.13）进行计算，

$$\begin{aligned}
\int_{\mathbb{S}^d} f d(\mu_t^1 - \mu_t^2) &= \int_{\mathbb{S}^d} f \circ X_t^1 G_t^1 d\bar{\mu}^1 - \int_{\mathbb{S}^d} f \circ X_t^2 G_t^2 d\bar{\mu}^2 \\
&= \int_{\mathbb{S}^d} (f \circ X_t^1 - f \circ X_t^2) G_t^1 d\bar{\mu}^1 + \int_{\mathbb{S}^d} f \circ X_t^2 (G_t^1 - G_t^2) d\bar{\mu}^1 + \int_{\mathbb{S}^d} f \circ X_t^2 G_t^2 d(\bar{\mu}^1 - \bar{\mu}^2) \\
&\le \left(\|X_t^1 - X_t^2\|_{L^\infty} \|G_t^1\|_{L^\infty} + \|G_t^1 - G_t^2\|_{L^\infty} \right) \bar{\mu}^1(\mathbb{S}^d) + \|f \circ X_t^2 G_t^2\|_{C^{0,1}} d_{BL}(\bar{\mu}^1, \bar{\mu}^2) \\
&\lesssim_{d,M,T} \int_0^t d_{BL}(\mu_r^1, \mu_r^2)dr + d_{BL}(\nu^1, \nu^2) + d_{BL}(\bar{\mu}^1, \bar{\mu}^2).
\end{aligned}$$

由 f 的任意性，我们得到

$$d_{BL}(\mu_t^1, \mu_t^2) \lesssim_{d,M,T} \int_0^t d_{BL}(\mu_r^1, \mu_r^2)dr + d_{BL}(\nu^1, \nu^2) + d_{BL}(\bar{\mu}^1, \bar{\mu}^2) \qquad \forall t \in [0,T).$$

于是，由格朗沃尔不等式可得

$$d_{BL}(\mu_t^1, \mu_t^2) \le d_{BL}(\bar{\mu}^1, \bar{\mu}^2) e^{C(d,M,T)t} + \left(d_{BL}(\bar{\mu}^1, \bar{\mu}^2) + d_{BL}(\nu^1, \nu^2) \right) \left(e^{C(d,M,T)t} - 1 \right) \qquad \forall t \in [0,T),$$

由此完成第 iii 点的证明，并得到唯一性。

步骤 3： 剩下要证明的只有解的存在性，以及正则性和奇偶性从数据中的传播。对于这两项，只需在一个短时间区间 $[0,\tau)$内进行讨论。事实上，按照定理 2.9中的相同论证，只要 $\mu_t(\mathbb{S}^d)$ 保持有界，解就可以延拓，因此可延拓至整个 $[0,\infty)$，这得益于（4.7）中的一致上界。

与通常做法一样，我们通过递归构造一列近似解来进行证明：

$$\mu_t^0 = \bar{\mu}, \qquad \mu_t^{n+1} = (X_t^n)_\# G_t^n \bar{\mu} \qquad \forall t \in [0,\tau), \quad \forall n \ge 0,$$

其中 v_t^n,X_t^n分别表示由 μ_t^n生成的向量场和流映射，并且

$$g_t^n := -4\mathcal{K}(\mu_t^n - \nu), \qquad G_t^n := \exp\left(\int_0^t g_r^n \circ X_r^n dr \right) \qquad \forall t \in [0,\tau), \quad \forall n \ge 0.$$

Fir首先，我们证明只要将 $\tau>0$ 选得足够小，以下结论就成立：

$$\mu_t^n(\mathbb{S}^d)\le 2\bar{\mu}(\mathbb{S}^d)\qquad \forall t\in[0,\tau),\quad \forall n\ge 0. \tag{4.14}$$

由质量的一致有界，应用（4.5），我们将推导出

$$\|v_t^n\|_{C^{1,\ 1}},\quad \|\nabla X_t^n\|_{C^{0,\ 1}},\quad \|\nabla(X_t^n)^{-1}\|_{C^{0,\ 1}},\quad \|g_t^n\|_{C^{2,\ 1}},\quad \|G_t^n\|_{C^{2,\ 1}}\le C(d,\ \tau,\ \bar{\mu}(\mathbb{S}^d)\ \nu(\mathbb{S}^d))\qquad \forall t\in[0,\tau),\quad \forall n\ge 0. \tag{4.15}$$

对于 4.14，当 $n=0$ 时，该界是显然的。假设对某个 $n\ge 0$ 成立，现在证明它对 $n+1$ 也成立。由（4.5），对于所有 $t\in[0,\tau)$，

$$\mu_t^{n+1}(\mathbb{S}^d)=\int_{\mathbb{S}^d}G_t^n d\bar{\mu}\le \|G_t^n\|_{L^\infty}\bar{\mu}(\mathbb{S}^d)\le \exp\left\{C(d)(2\bar{\mu}(\mathbb{S}^d)+\nu(\mathbb{S}^d)\tau\right\}\bar{\mu}(\mathbb{S}^d)\le 2\bar{\mu}(\mathbb{S}^d),$$

只要我们选择 τ 足够小即可。

已知（4.15）后，按照步骤 2 的方法进行推导，可得

$$d_{BL}(\mu_t^n,\mu_t^{n+1})\lesssim_{d,\tau,\bar{\mu}(\mathbb{S}^d),\nu(\mathbb{S}^d)}\int_0^t\left(d_{BL}(\mu_r^n,\mu_r^{n+1})+d_{BL}(\mu_r^{n-1},\mu_r^n)\right)dr\qquad \forall t\in[0,\tau),\quad \forall n\ge 1.$$

由此，按照定理 2.8 证明中步骤 2 的相同论证，并适当减小 τ，我们可得

$$\lim_{n\to\infty}\sup_{k\ge n}\max_{t\in[0,\tau)}d_{BL}(\mu_t^k,\mu_t^{k+1})=0. \tag{4.16}$$

为结束在区间 $[0,\tau)$ 上方程 (1.15 的解的存在性证明，我们仿照定理 2.8 证明中的步骤 3 进行论证。首先，由 (4.15)，X^n 和 G^n 在 $[0,\tau)\times\mathbb{S}^d$上一致有界且等度连续。因此，由阿尔泽拉-阿斯科利定理，我们可以找到连续函数 $X\colon[0,\tau)\times\mathbb{S}^d\to\mathbb{S}^d$ 和 $G\colon[0,\tau)\times\mathbb{S}^d\to[0,\infty)$，以及一个子列 $n_k\to\infty$，使得

$$\lim_{k\to\infty}\left(\|X^{n_k}-X\|_{C([0,\tau)\times\mathbb{S}^d)}+\|G^{n_k}-G\|_{C([0,\tau)\times\mathbb{S}^d)}\right)=0.$$

定义 $\mu_t:=(X_t)_\#G_t\bar{\mu}$，我们可得

$$\mu_t^{n_k}\overset{*}{\rightharpoonup}\mu_t\qquad \forall t\in[0,\tau),\qquad \mu\in C_{w^*}([0,\tau);\mathcal{M}_+(\mathbb{S}^d)). \tag{4.17}$$

然后，由 (4.16) 还可得到

$$\mu_t^{n_k-1}\overset{*}{\rightharpoonup}\mu_t\qquad \forall t\in[0,\tau).$$

最后，令 $g_t:=-4\mathcal{K}(\mu_t-\nu)$ 以及 $v_t:=-\nabla_{\mathbb{S}^d}\mathcal{K}(\mu_t-\nu)$，由 (4.8) 可得

$$\|g_t^{n_k-1}-g_t\|_{C(\mathbb{S}^d)},\ \|v_t^{n_k-1}-v_t\|_{C(\mathbb{S}^d)}\to 0\qquad \forall t\in[0,\tau). \tag{4.18}$$

为证明 μ 是（1.15）的一个解，只需在分布形式的两边应用控制收敛定理取极限

$$\int_{\mathbb{S}^d}\varphi\mu_t^{n_k}=\int_{\mathbb{S}^d}\varphi\bar{\mu}+\int_0^t\int_{\mathbb{S}^d}\nabla_{\mathbb{S}^d}\varphi\cdot v_r^{n_k-1}\mu_r^{n_k}dr+\int_0^t\int_{\mathbb{S}^d}\varphi g_r^{n_k-1}d\mu_r^{n_k}dr\qquad \forall t\in[0,\tau),\quad \forall\varphi\in C^\infty(\mathbb{T}^d),$$

结合（4.17）、（4.18）以及 g^n、g、v^n、v的一致有界性。至此，存在性的证明完成。

关于第 ii) 点，注意到如果 $\bar{\mu}$、ν 是偶函数，那么在上述逼近序列中，对所有 $t\in[0,\tau)$ 和所有 $n\ge 0$，μ_t^n 都是偶函数；这可利用 $\mathcal{K}$ 保持奇偶性的性质，对 n 应用归纳法证明。因此，逼近序列的奇偶性会传递到极限解。

最后，关于第 iv) 点中赫尔德正则性的传播，其证明可仿照定理2.10，利用 $\mathcal{K}$ 在（4.5）中的正则化性质得出。随后，可仿照定理2.12，推导出具有光滑数据的解在时空中的光滑正则性。关于这一点的细节留给读者。

□

4.4 QuWasserstein–Fisher–Rao 动力学的定量收敛

本节证明定理1.7。证明策略是将能量耗散恒等式（1.18）与我们在下一个引理中推导的高阶能量估计相结合（见第1.5节）。为方便起见，后文中我们采用整数参数，并使用经典的角导数代替分数阶导数（见第4.2节中引入的记号）。

引理 *4.6.* 设 $d \in \mathbb{N}$ 为奇函数且 $s := (d+3)/2 \in \mathbb{N}$。设 $\bar{\mu}$、$\nu \in C^\infty(\mathbb{S}^d)$ 为偶非负函数，且 $\mu \in C^\infty([0,\infty) \times \mathbb{S}^d)$ 是由定理 *4.5* 给出的方程 *(1.15)* 的解。对于任意 $\gamma \in \mathbb{N}$、$\gamma \geq s-1$，存在某个常数 $C = C(d,\gamma) > 0$，使得对所有 $t \in (0,\infty)$，以下能量估计成立：

$$\frac{d}{dt}\|\mu_t\|_{\dot{H}^\gamma}^2 \leq C\left(\|\nu\|_{H^\gamma} + \|\mathcal{K}(\mu_t-\nu)\|_{C^2}\right)\left(\|\mu_t\|_{H^\gamma}^2 + \|\nu\|_{H^\gamma}^2\right), \tag{4.19}$$

$$\begin{aligned}\frac{d}{dt}\|\mu_t\|_{\dot{H}^\gamma}^2 \leq{}& C\|\mathcal{K}(\mu_t-\nu)\|_{C^2}\left(\|\mu_t\|_{H^\gamma}^2 + \|\nu\|_{H^\gamma}^2\right) \\ &- \frac{1}{C}\left(\min_{\mathbb{S}^d}\nu\right)\|\mu_t-\nu\|_{\dot{H}^{\gamma-s+1}}^2 + C\|\nu\|_{H^\gamma}\|\mu_t-\nu\|_{\dot{H}^{\max\{\gamma,\frac{d}{2}+1\}-s}}\|\mu_t-\nu\|_{\dot{H}^{\gamma-s+1}} \\ &+ C\|\nu\|_{H^\gamma}\|\nu\|_{\dot{H}^{\gamma+1}}\|\mu_t-\nu\|_{\dot{H}^{\gamma+1-2s}}.\end{aligned} \tag{4.20}$$

证明。 为方便起见，记 $\sigma_t := \mu_t - \nu$ 为相对于目标的扰动。利用方程 (1.15) 并对散度项进行分部积分，可得

$$\begin{aligned}\frac{d}{dt}\frac{\|\mu_t\|_{\dot{H}^\gamma}^2}{2} &= \sum_{|\beta|=\gamma}\int_{\mathbb{S}^d}\nabla_\beta\partial_t\mu_t\nabla_\beta\mu_t \\ &= \underbrace{-\sum_{|\beta|=\gamma}\int_{\mathbb{S}^d}\nabla_\beta\left[\mu_t\nabla_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\right]\cdot\nabla_{\mathbb{S}^d}\nabla_\beta\mu_t}_{I} + \underbrace{(-4)\sum_{|\beta|=\gamma}\int_{\mathbb{S}^d}\nabla_\beta\left[\mu_t\mathcal{K}(\sigma_t)\right]\nabla_\beta\mu_t}_{II}.\end{aligned}$$

我们先详细说明如何界定 I；它是主导阶项，因为 II 的导数阶数少两阶。

步骤 1： 由 I 可知，我们将含有作用于 μ_t 的 $2\gamma+1$ 导数的项单独提出：

$$I = \underbrace{-\sum_{|\beta|=\gamma}\int_{\mathbb{S}^d}\nabla_\beta\mu_t\nabla_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\cdot\nabla_{\mathbb{S}^d}\nabla_\beta\mu_t}_{I_1} + \underbrace{(-1)\sum_{0\leq|\beta_1|\leq\gamma-1}\sum_{|\beta_2|=\gamma+1-|\beta_1|}\int_{\mathbb{S}^d}\nabla_{\beta_1}\mu_t\nabla_{\beta_2}\mathcal{K}(\sigma_t)\nabla_{\beta_1+\beta_2}\mu_t}_{I_2}.$$

U利用恒等式 $\nabla_{\mathbb{S}^d}(\nabla_\beta\mu_t)^2 = 2\nabla_\beta\mu_t\nabla_{\mathbb{S}^d}\nabla_\beta\mu_t$ 和分部积分，我们得到

$$I_1 = \frac{1}{2}\sum_{|\beta|=\gamma}\int_{\mathbb{S}^d}\Delta_{\mathbb{S}^d}\mathcal{K}(\sigma_t)|\nabla_\beta\mu_t|^2 \leq C\|\nabla_{\mathbb{S}^d}^2\mathcal{K}(\sigma_t)\|_{L^\infty}\|\mu_t\|_{\dot{H}^\gamma}^2. \tag{4.21}$$

接下来，写作 $\mu_t = \nu + \sigma_t$，我们进一步将 I_2 分为三项：

$$\begin{aligned}I_2 = &-\sum_{0\leq|\beta_1|\leq\gamma-1}\sum_{|\beta_2|=\gamma+1-|\beta_1|}\int_{\mathbb{S}^d}\nabla_{\beta_1}\sigma_t\nabla_{\beta_2}\mathcal{K}(\sigma_t)\nabla_{\beta_1+\beta_2}\mu_t \\ &-\sum_{0\leq|\beta_1|\leq\gamma-1}\sum_{|\beta_2|=\gamma+1-|\beta_1|}\int_{\mathbb{S}^d}\nabla_{\beta_1}\nu\nabla_{\beta_2}\mathcal{K}(\sigma_t)\nabla_{\beta_1+\beta_2}\sigma_t \\ &-\sum_{0\leq|\beta_1|\leq\gamma-1}\sum_{|\beta_2|=\gamma+1-|\beta_1|}\int_{\mathbb{S}^d}\nabla_{\beta_1}\nu\nabla_{\beta_2}\mathcal{K}(\sigma_t)\nabla_{\beta_1+\beta_2}\nu =: I_{2,1}+I_{2,2}+I_{2,3}.\end{aligned}$$

步骤 2： 在这一步中，我们证明关于 $I_{2,1}$的以下估计：

$$I_{2,1} \leq C\|\nabla_{\mathbb{S}^d}^2\mathcal{K}(\sigma_t)\|_{L^\infty}\|\sigma\|_{H^\gamma}\|\mu_t\|_{\dot{H}^\gamma}. \tag{4.22}$$

对于定义 $I_{2,1}$的每个积分，通过分部积分，我们可以将一个导数从 $\nabla_{\beta_1+\beta_2}\mu_t$ 转移到乘积 $\nabla_{\beta_1}\sigma_t\nabla_{\beta_2}\sigma_t$上。这样，我们可以将 $I_{2,1}$ 写成有限个形如$\nabla_{\alpha_1}\sigma_t\nabla_{\alpha_2}\mathcal{K}(\sigma_t)\nabla_{\alpha_3}\mu_t$的项之和的积分，其中

$$0 \le |\alpha_1| \le \gamma, \quad 2 \le |\alpha_2| \le \gamma+2, \quad |\alpha_1|+|\alpha_2| = \gamma+2, \quad |\alpha_3| = \gamma. \tag{4.23}$$

那么，为了得出结论 （4.22），只需证明对每个 α_1, α_2 满足（4.23）时，以下结论成立：

$$\|\nabla_{\alpha_1}\sigma_t\nabla_{\alpha_2}\mathcal{K}(\sigma_t)\|_{L^2} \le C\|\nabla^2_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|_{L^\infty}\|\sigma_t\|_{H^\gamma}. \tag{4.24}$$

我们假设 $0 \le |\alpha_1| < \gamma$ 和 $2 < |\alpha_2| \le \gamma+2$，因为由赫尔德不等式可轻易推出 $|\alpha_1| = \gamma$ 和 $|\alpha_2| = 2$ 的情形。我们写作 $\sigma_t = (\sigma_t)_0 + T^{-1}(-\Delta_{\mathbb{S}^d})^s\mathcal{K}(\sigma_t)$，其中 $T = (-\Delta_{\mathbb{S}^d})^s\mathcal{K}$是定理 4.3 第 ii) 点中的零阶算子，而 $(\sigma_t)_0$ 是 σ_t的均值。若 $|\alpha_1| > 0$，则 $\nabla_{\alpha_1}\sigma_t = T^{-1}(-\Delta_{\mathbb{S}^d})^s\nabla_{\alpha_1}\mathcal{K}(\sigma_t)$，因此，应用赫尔德不等式、定理 4.3 第 ii) 点以及 Gagliardo–Nirenberg插值不等式，可得

$$\begin{aligned}
\|\nabla_{\alpha_1}\sigma_t\nabla_{\alpha_2}\mathcal{K}(\sigma_t)\|_{L^2} &\le \|T^{-1}(-\Delta_{\mathbb{S}^d})^s\nabla_{\alpha_1}\mathcal{K}(\sigma_t)\|_{L^{2\frac{\gamma+2s-2}{\gamma+2s-|\alpha_2|}}}\|\nabla_{\alpha_2}\mathcal{K}(\sigma_t)\|_{L^{2\frac{\gamma+2s-2}{|\alpha_2|-2}}} \\
&\lesssim \|\nabla^{2s+|\alpha_1|}_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|_{L^{2\frac{\gamma+2s-2}{\gamma+2s-|\alpha_2|}}}\|\nabla^{|\alpha_2|}_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|_{L^{2\frac{\gamma+2s-2}{|\alpha_2|-2}}} \\
&\lesssim \|\nabla^2_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|^{\frac{|\alpha_2|-2}{\gamma+2s-2}}_{L^\infty}\|\nabla^{2s+\gamma}_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|^{\frac{\gamma+2s-|\alpha_2|}{\gamma+2s-2}}_{L^2}\|\nabla^2_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|^{\frac{\gamma+2s-|\alpha_2|}{\gamma+2s-2}}_{L^\infty}\|\nabla^{2s+\gamma}_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|^{\frac{|\alpha_2|-2}{\gamma+2s-2}}_{L^2} \\
&= \|\nabla^2_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|_{L^\infty}\|\nabla^{2s+\gamma}_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|_{L^2} \\
&\lesssim \|\nabla^2_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|_{L^\infty}\|\sigma_t\|_{\dot{H}^\gamma}.
\end{aligned}$$

若 $|\alpha_1| = 0$，则相同的界仍然成立，只需使用全范数 $\|\sigma_t\|_{H^\gamma}$ 代替 $\|\sigma_t\|_{\dot{H}^\gamma}$。至此，式（4.22）的证明完成。

步骤 3： 接下来，我们证明以下关于$I_{2,\ 2}$的界：

$$I_{2,2} \le -\frac{1}{C}\big(\min_{\mathbb{S}^d}\nu\big)\|\sigma_t\|^2_{\dot{H}^{\gamma-s+1}} + C\|\nu\|_{H^\gamma}\left(\|\sigma_t\|_{\dot{H}^{\max\{\gamma,\frac{d}{2}+1\}-s}}\|\sigma_t\|_{\dot{H}^{\gamma-s+1}} + \|\nabla^2_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|_{L^\infty}\|\sigma_t\|_{H^\gamma}\right). \tag{4.25}$$

在$I_{2,\ 2}$定义中的每一项里，我们使用分部积分将 s 个导数从$\nabla_{\beta_1+\beta_2}\sigma_t$ 分配到乘积 $\nabla_{\beta_1}\nu\nabla_{\beta_2}\mathcal{K}(\sigma_t)$上，并根据落在 ν上的导数数量对结果进行拆分：

$$\begin{aligned}
I_{2,2} &= -\sum_{0\le|\alpha_1|\le\gamma+s-1}\ \sum_{|\alpha_2|=\gamma+s+1-|\alpha_1|}\int_{\mathbb{S}^d}\nabla_{\alpha_1}\nu\nabla_{\alpha_2}\mathcal{K}(\sigma_t)(-\Delta_{\mathbb{S}^d})^{-s}\nabla_{\alpha_1+\alpha_2}\sigma_t \\
&= -\sum_{|\alpha|=\gamma+s+1}\int_{\mathbb{S}^d}\nu\nabla_\alpha\mathcal{K}(\sigma_t)(-\Delta_{\mathbb{S}^d})^{-s}\nabla_\alpha\sigma_t \\
&\quad -\sum_{1\le|\alpha_1|\le\gamma-1}\ \sum_{|\alpha_2|=\gamma+s+1-|\alpha_1|}\int_{\mathbb{S}^d}\nabla_{\alpha_1}\nu\nabla_{\alpha_2}\mathcal{K}(\sigma_t)(-\Delta_{\mathbb{S}^d})^{-s}\nabla_{\alpha_1+\alpha_2}\sigma_t \\
&\quad -\sum_{\gamma\le|\alpha_1|\le\gamma+s-1}\ \sum_{|\alpha_2|=\gamma+s+1-|\alpha_1|}\int_{\mathbb{S}^d}\nabla_{\alpha_1}\nu\nabla_{\alpha_2}\mathcal{K}(\sigma_t)(-\Delta_{\mathbb{S}^d})^{-s}\nabla_{\alpha_1+\alpha_2}\sigma_t =: I_{2,2,1}+I_{2,2,2}+I_{2,2,3}.
\end{aligned}$$

我们首先考虑 $I_{2,\ 2,\ 1}$。令 $\bar{c} = \bar{c}(d) > 0$ 为定理 4.3第 iii 点中出现的常数。将 $\bar{c}(-\Delta_{\mathbb{S}^d})^{-s}$ 加上再减去 $\mathcal{K}$，并利用柯西–施瓦茨不等式，得到

$$\begin{aligned}
I_{2,2,1} &= -\bar{c}\sum_{|\alpha|=\gamma+s+1}\int_{\mathbb{S}^d}\nu\left((-\Delta_{\mathbb{S}^d})^{-s}\nabla_\alpha\sigma_t\right)^2 \\
&\quad -\sum_{|\alpha|=\gamma+s+1}\int_{\mathbb{S}^d}\nu\left(\mathcal{K}-\bar{c}(-\Delta_{\mathbb{S}^d})^{-s}\right)(\nabla_\alpha\sigma_t)(-\Delta_{\mathbb{S}^d})^{-s}(\nabla_\alpha\sigma_t) \\
&\le -\frac{1}{C}\big(\min_{\mathbb{S}^d}\nu\big)\|\sigma_t\|^2_{\dot{H}^{\gamma-s+1}} + C\|\nu\|_{L^\infty}\|\sigma_t\|_{\dot{H}^{\gamma-s}}\|\sigma_t\|_{\dot{H}^{\gamma-s+1}}.
\end{aligned} \tag{4.26}$$

为了给 $I_{2,\ 2,\ 2}$ 作界，我们对其定义中出现的每个积分使用赫尔德不等式和索伯列夫嵌入。若 $\gamma-|\alpha_1|>d/2$，由于 $|\alpha_2|\le\gamma+s$，

$$\|\nabla_{\alpha_1}\nu\nabla_{\alpha_2}\mathcal{K}(\sigma_t)\|_{L^2}\le\|\nabla_{\alpha_1}\nu\|_{L^\infty}\|\nabla_{\alpha_2}\mathcal{K}(\sigma_t)\|_{L^2}\le C\|\nu\|_{\dot H^\gamma}\|\sigma_t\|_{\dot H^{\gamma-s}}.$$

假设 $\gamma-|\alpha_1|<d/2$。然后，令 $p=2d/(d-2\gamma+2|\alpha_1|)$ 与 $q=2d/(d-2r+2|\alpha_2|)$，其中 $r=d/2+s+1$，我们得到

$$\|\nabla_{\alpha_1}\nu\nabla_{\alpha_2}\mathcal{K}(\sigma_t)\|_{L^2}\le\|\nabla_{\alpha_1}\nu\|_{L^p}\|\nabla_{\alpha_2}\mathcal{K}(\sigma_t)\|_{L^q}\le C\|\nu\|_{\dot H^\gamma}\|\sigma_t\|_{\dot H^{\frac{d}{2}-s+1}}.$$

因此，

$$I_{2,2,2}\le C\|\nu\|_{\dot H^\gamma}\|\sigma_t\|_{\dot H^{\max\{\gamma,\frac{d}{2}+1\}-s}}\|\sigma_t\|_{\dot H^{\gamma-s+1}}. \tag{4.27}$$

关于 $I_{2,2,3}$，通过分部积分将 $|\alpha_1|-\gamma$ 导数从 $\nabla_{\alpha_1}\nu$ 分配到另外两项，我们将其化为估计与定义 $I_{2,1}$ 的积分类似的积分，其中以 ν 代替 μ_t。上文步骤 2 的同样论证随后导出如下估计：

$$I_{2,2,3}\le C\|\nabla^2_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|_{L^\infty}\|\sigma_t\|_{\dot H^\gamma}\|\nu\|_{\dot H^\gamma}. \tag{4.28}$$

综合（4.26）、（4.27）与（4.28），我们得到（4.25）。

步骤 4：最后，我们估计$I_{2,3}$：

$$I_{2,3}\le C\|\nu\|_{H^\gamma}\|\nu\|_{\dot H^{\gamma+1}}\|\sigma_t\|_{\dot H^{\gamma+1-2s}}. \tag{4.29}$$

该策略类似于对 $I_{2,2,2}$ 所采用的策略：我们对定义中出现的每个积分使用赫尔德不等式和索伯列夫嵌入。若 $\gamma-|\beta_1|>d/2$，由于 $|\beta_2|\le\gamma+1$，我们有

$$\|\nabla_{\beta_1}\nu\nabla_{\beta_2}\mathcal{K}(\sigma_t)\|_{L^2}\le\|\nabla_{\beta_1}\nu\|_{L^\infty}\|\nabla_{\beta_2}\mathcal{K}(\sigma_t)\|_{L^2}\le C\|\nu\|_{H^\gamma}\|\sigma_t\|_{\dot H^{\gamma+1-2s}}.$$

假设 $\gamma-|\beta_2|<d/2$。令 $p=2d/(d-2\gamma+2|\beta_1|)$ 和 $q=2d/(d-2r+2|\beta_2|)$，其中 $r=d/2+1$，可得

$$\|\nabla_{\beta_1}\nu\nabla_{\beta_2}\mathcal{K}(\sigma_t)\|_{L^2}\le\|\nabla_{\beta_1}\nu\|_{L^p}\|\nabla_{\beta_2}\mathcal{K}(\sigma_t)\|_{L^q}\le C\|\nu\|_{H^\gamma}\|\sigma_t\|_{\dot H^{\frac{d}{2}+1-2s}}.$$

由于根据假设 $\gamma>d/2$，第一种情形右侧的界更大，因此我们得到（4.29）。如果将一个导数预先分配给 $\nabla_{\beta_1+\beta_2}\nu$ 中的另外两项，同样的方法可得

$$I_{2,3}\le C\|\nu\|^2_{H^\gamma}\|\sigma_t\|_{\dot H^{\gamma+2-2s}}. \tag{4.30}$$

步骤 5：综合（4.21）、（4.22）、（4.25）、（4.29）和（4.30），可得

$$\begin{aligned}I\le{}&\|\nabla^2_{\mathbb{S}^d}\mathcal{K}(\sigma_t)\|_{L^\infty}\left(\|\mu_t\|^2_{H^\gamma}+\|\nu\|^2_{H^\gamma}\right)\\&-\frac{1}{C}\big(\min_{\mathbb{S}^d}\nu\big)\|\sigma_t\|^2_{\dot H^{\gamma-s+1}}+C\|\nu\|_{H^\gamma}\|\sigma_t\|_{\dot H^{\max\{\gamma,\frac{d}{2}+1\}-s}}\|\sigma_t\|_{\dot H^{\gamma-s+1}}\\&+C\|\nu\|_{H^\gamma}\min\{\|\nu\|_{\dot H^{\gamma+1}}\|\sigma_t\|_{\dot H^{\gamma+1-2s}},\|\nu\|_{\dot H^\gamma}\|\sigma_t\|_{\dot H^{\gamma+2-2s}}\}.\end{aligned} \tag{4.31}$$

要对 II 作界，可以沿用处理 I 的相同步骤；由于 II 比 I 少两个导数，这些步骤会更加简单：

$$II\le C\|\mathcal{K}(\sigma_t)\|_{L^\infty}\left(\|\mu_t\|^2_{H^\gamma}+\|\nu\|^2_{H^\gamma}\right)+C\|\nu\|_{H^\gamma}\|\sigma_t\|^2_{\dot H^{\gamma-s}}+C\|\nu\|^2_{H^\gamma}\|\sigma_t\|_{\dot H^{\gamma-2s}}. \tag{4.32}$$

最后，（4.31）和（4.32）合起来即可推出（4.19）和（4.20）。 □

结合能量估计（4.19）与（4.7）中的质量上界，我们可以控制光滑解的 H^γ-范数的（指数）增长，该增长仅取决于 d、γ、以及数据的 H^γ-范数 $\bar\mu$、ν。特别地，利用通常的逼近论证（例如参见定理 2.14的证明），我们得到如下结论：

推论 4.7。 设 $d\in\mathbb{N}$ 为奇函数，且设 $s=(d+3)/2\in\mathbb{N}$。设 $\gamma\in\mathbb{N}$ 满足 $\gamma\ge s-1$，令$\bar\mu$，$\nu\in H^\gamma(\mathbb{S}^d)$ 为偶非负函数，并设 μ 为定理 (1.15) 给出的解，该定理由 *4.5*给出。那么，$\mu\in L^\infty_{\mathrm{loc}}([0,\infty);H^\gamma(\mathbb{S}^d))$。

定理 1.7 的证明。 通过逼近，只需在假设 $\bar{\mu}, \nu \in C^\infty(\mathbb{S}^d)$ 的情况下证明$\mu \in C^\infty([0,\infty) \times \mathbb{S}^d)$。

In 后文中，为方便起见，我们将 σ_t 称为：$= \mu_t - \nu$。此外，我们引入以下量

$$E_\gamma(t) := \|\mu_t\|_{\dot{H}^\gamma}^2 + \mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t) + \|\nu\|_{H^\gamma}^2 \qquad \forall t \in [0,\infty).$$

注意，由（1.19），$E_\gamma(t) \approx_{d,\ \gamma} \|\mu_t\|_{\dot{H}^\gamma}^2 + \|\nu\|_{H^\gamma}^2$。

考虑能量估计（4.20）右端的第二行。由于 $\max\{\gamma, \frac{d}{2}+1\} - s \in (-s,\ \gamma - s + 1)$，我们可以对 $\|\sigma_t\|_{\dot{H}^{\max\{\gamma,\frac{d}{2}+1\}-s}}$在$\|\sigma_t\|_{\dot{H}^{-s}}$和$\|\sigma_t\|_{\dot{H}^{\gamma-s+1}}$之间使用索博列夫插值、Young 不等式以及下界 $\nu \geq \alpha > 0$，从而将正项吸收到负项中，但会产生一个低阶项：

$$-\frac{1}{C}\big(\min_{\mathbb{T}^d} \nu\big)\|\sigma_t\|_{\dot{H}^{\gamma-s+1}}^2 + C\|\nu\|_{H^\gamma}\|\sigma_t\|_{\dot{H}^{\max\{\gamma,\frac{d}{2}+1\}-s}}\|\sigma_t\|_{\dot{H}^{\gamma-s+1}} \leq C\|\sigma_t\|_{\dot{H}^{-s}}^2 \leq C\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t),$$

其中，最后一步使用了（1.19）。将上述估计代入（4.20），并利用以下事实：由（1.18），$\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t)$ 随时间递减，我们得到

$$\frac{d}{dt}E_\gamma(t) \leq C\|\mathcal{K}(\sigma_t)\|_{C^2}E_\gamma(t) + C\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t) + C\|\sigma_t\|_{\dot{H}^{\gamma+1-2s}}.$$

积分后，由此可得

$$E_\gamma(t) \leq \left(E_\gamma(0) + C\int_0^t \left(\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_r) + \|\sigma_r\|_{\dot{H}^{\gamma+1-2s}}\right) dr\right) \exp\left(C\int_0^t \|\mathcal{K}(\sigma_r)\|_{C^2} dr\right). \tag{4.33}$$

另一方面，能量耗散恒等式（1.18）一旦与（1.19）相结合，并对 $\|\sigma_t\|_{\dot{H}^{-s}}$ 在 $\|\sigma_t\|_{\dot{H}^{1-2s}}$ 与 $\|\sigma_t\|_{\dot{H}^\gamma}$ 之间作索博列夫插值，可得

$$\begin{aligned}
\frac{d}{dt}\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t) &= -4\int_{\mathbb{S}^d} \left(|\nabla_{\mathbb{S}^d}\mathcal{K}(\sigma_t)|^2 + |\mathcal{K}(\sigma_t)|^2\right)\mu_t \\
&\leq -\frac{1}{C}\big(\inf_{\mathbb{S}^d}\mu_t\big)\left(\sigma_t(\mathbb{S}^d)^2 + \|\sigma_t\|_{\dot{H}^{1-2s}}^2\right) \\
&\leq -\frac{1}{C}\big(\inf_{\mathbb{S}^d}\mu_t\big)\left(\sigma_t(\mathbb{S}^d)^2 + \|\sigma_t\|_{\dot{H}^\gamma}^{-\frac{2s-2}{\gamma+s}}\|\sigma_t\|_{\dot{H}^{-s}}^{2+\frac{2s-2}{\gamma+s}}\right) \leq -\frac{1}{C}\big(\inf_{\mathbb{S}^d}\mu_t\big)E_\gamma(t)^{-\frac{s-1}{\gamma+s}}\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t)^{1+\frac{s-1}{\gamma+s}}.
\end{aligned} \tag{4.34}$$

记 $M := E_\gamma(0)$，并考虑

$$\tau := \sup\left\{t > 0 : E_\gamma(r) \leq 2M \ \text{ and } \ \inf_{\mathbb{S}^d}\mu_r \geq \alpha/2,\ \forall r \in [0,t)\right\} \in (0,\infty].$$

我们的目标是证明：若 $\mathscr{E}_{\mathbb{S}^d}^\nu(\bar{\mu})$ 充分小，则 $\tau = \infty$。由（4.34）以及 τ 的定义，我们有

$$\frac{d}{dt}\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t) \leq -\kappa\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t)^{1+\frac{s-1}{\gamma+s}} \qquad \forall t \in [0,\tau),$$

积分之后，我们得到

$$\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t) \leq \mathscr{E}_{\mathbb{S}^d}^\nu(\bar{\mu})\,(1+Kt)^{-\frac{\gamma+s}{s-1}} \quad \forall t \in [0,\tau), \qquad K = C\mathscr{E}_{\mathbb{S}^d}^\nu(\bar{\mu})^{\frac{s-1}{\gamma+s}}. \tag{4.35}$$

由索博列夫插值，我们推出 $\|\sigma_t\|_{\dot{H}^{\gamma+1-2s}}$的一个界：

$$\begin{aligned}
\|\sigma_t\|_{\dot{H}^{\gamma+1-2s}} &\leq \|\sigma_t\|_{\dot{H}^{-s}}^{\frac{2s-1}{\gamma+s}}\|\sigma_t\|_{\dot{H}^\gamma}^{\frac{\gamma+1-s}{\gamma+s}} \\
&\leq C\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t)^{\frac{2s-1}{2(\gamma+s)}}E_\gamma(t)^{\frac{\gamma+1-s}{2(\gamma+s)}} \leq C\mathscr{E}_{\mathbb{S}^d}^\nu(\bar{\mu})^{\frac{2s-1}{2(\gamma+s)}}(1+Kt)^{-\frac{2s-1}{2s-2}}
\end{aligned} \qquad \forall t \in [0,\tau). \tag{4.36}$$

类似地，我们可以得到 $\|\mathcal{K}(\sigma_t)\|_{C^2}$ 的一个界：令 $\eta := \gamma - d/2 > 0$ 和 $\hat{\gamma} := d/2 + 2 - 2s + \eta/2$，由 $\nabla^2(-\Delta)^{-s}$：$\dot{H}^{\hat{\gamma}} \to L^\infty$ 的有界性以及 Sobolev 插值可得

$$\begin{aligned}
\|\mathcal{K}(\sigma_t)\|_{C^2} &\leq C\|\mathcal{K}(\sigma_t)\|_{H^{\hat{\gamma}+2s}} \\
&\leq C\mathscr{E}_{\mathbb{S}^d}^\nu(\mu_t)^{\frac{2s-2+\eta/2}{2(\gamma+s)}}E_\gamma(t)^{\frac{\gamma+2-\eta/2-s}{2(\gamma+s)}} \leq C\mathscr{E}_{\mathbb{S}^d}^\nu(\bar{\mu})^{\frac{2s-2+\eta/2}{2(\gamma+s)}}(1+Kt)^{-\frac{2s-2+\eta/2}{2s-2}}
\end{aligned} \qquad \forall t \in [0,\tau). \tag{4.37}$$

在时间区间 $[0,\tau)$ 将 (4.35)、(4.36) 和 (4.37) 中的界进行积分，我们得到以下积分估计（常数与 τ 无关）：

$$\int_0^\tau \mathscr{E}^\nu_{\mathbb{S}^d}(\mu_t)dt \le C\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)^{\frac{\gamma+1}{\gamma+s}},\quad \int_0^\tau \|\sigma_t\|_{\dot H^{\gamma+1-2s}}dt \le C\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)^{\frac{1}{2(\gamma+s)}},\quad \int_0^\tau \|\mathcal{K}(\sigma_t)\|_{C^2}dt \le C\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)^{\frac{\gamma-d/2}{4(\gamma+s)}}. \tag{4.38}$$

由此，我们推出 μ_t 有一致下界。事实上，由于 $\inf \bar\mu \ge \alpha$，根据定理 4.5 中得到的解 μ 的拉格朗日表示公式，我们得到

$$\inf_{\mathbb{S}^d}\mu_t \ge \left(\inf_{\mathbb{S}^d}\bar\mu\right)\exp\left(-C\int_0^t \|\mathcal{K}(\sigma_r)\|_{C^2}dr\right) \ge \alpha\exp\left(-C\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)^{\frac{\gamma-d/2}{4(\gamma+s)}}\right) \qquad \forall t\in[0,\tau)。 \tag{4.39}$$

现在，我们通过证明：若 $\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)$ 足够小，则 $\tau=\infty$。结合式（4.35），可知式（1.20）中的估计对所有时刻 $t\in[0,\infty)$ 均成立，从而完成证明。反设$\tau<\infty$。由式（4.39），通过选取足够小的 $\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)$，我们可以确保对每个$\inf_{\mathbb{S}^d}\mu_t \ge 3\alpha/4$，都有$t\in[0,\tau)$，因此必有 $E_\gamma(\tau)=2M$。然而，由式（4.33）和式（4.38）可得

$$E_\gamma(t) \le \left(M + C\left(\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)^{\frac{\gamma+1}{\gamma+s}} + \mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)^{\frac{1}{2(\gamma+s)}}\right)\right)\exp\left(C\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)^{\frac{\gamma-d/2}{4(\gamma+s)}}\right) \le \frac{3}{2}M,$$

只要将 $\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)$ 选得足够小即可。这就产生了矛盾，证明得证。 □

注记 4.8。与定理 3.4中一样，假设 $\nu\in H^{\gamma+1}(\mathbb{S}^d)$ 是技术性的。在神经网络设定下研究分数阶索博列夫正则性时，可以将其放宽为 $\nu\in H^{\gamma+\varepsilon}(\mathbb{S}^d)$，对任意固定的 $\varepsilon>0$，通过在 $\mathbb{S}^d$ 上应用分数阶莱布尼茨法则，并将 ε 导数从（4.29）重新分配到另外两个因子上（而非使用完整导数）；例如参见 [Bui25，推论 1。2]。 ■

定理 *1.10*的证明。令 $\gamma := \beta - s - 1 \ge s-1$，并考虑在（4.3）中定义并在定理 4.1中研究的乘子算子 $\mathcal{H}$。有 $\nu = \mathcal{H}^{-1}(g+S) = \mathcal{H}^{-1}(g) + c(d)S$，其中 $c(d) > 0$是某个正的常数。因此，由索伯列夫嵌入可得：

$$\min_{\mathbb{S}^d}\nu \ge c(d)S - C(d,\beta)\|\mathcal{H}^{-1}(g)\|_{H^{\gamma+1}} \ge c(d)K(1+\|g\|_{H^\beta}) - C(d,\beta)\|g\|_{H^\beta} \ge 2$$

只要我们选取 $K = K(d,\beta) > 0$ 足够大。在此 K 的选取下，我们有$\|\nu\|_{H^{\gamma+1}} \lesssim_{d,\beta} \|g\|_{H^\beta}+1$。

接下来，注意到 $\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu) = \|f-f_{\bar\mu}\|^2_{L^2} \le \delta$。因此，记 $\eta := \gamma - d/2 > 0$，由索伯列夫嵌入与插值，

$$\begin{aligned}\|\bar\mu-\nu\|_{L^\infty} &\le C(d,\beta)\|\bar\mu-\nu\|_{H^{\frac{d+\eta}{2}}}\\ &\le C(d,\beta)\mathscr{E}^\nu_{\mathbb{S}^d}(\bar\mu)^{\frac{2\gamma-d-\eta}{4(\gamma+s)}}\|\bar\mu-\nu\|_{H^\gamma}^{\frac{d+2s+\eta}{4(\gamma+s)}} \le C(d,\beta)\delta^{\frac{2\gamma-d-\eta}{4(\gamma+s)}}\left(M+\|g\|_{H^\beta}\right)^{\frac{d+2s+\eta}{4(\gamma+s)}} \le 1,\end{aligned}$$

只要我们选取 δ足够小，且仅依赖于 d,β,M 与 $\|g\|_{H^\beta}$。在此选取下， $\bar\mu\ge 1$。至此，该推论由直接应用定理 1.7 得到，注意到 $\|f_{\mu_t}-f\|^2_{L^2} = \mathscr{E}^\nu_{\mathbb{S}^d}(\mu_t)$，以及$\|f_{\mu_t}\|_{H^{\beta-1}} \approx_{d,\beta} \|\mu_t\|_{H^\gamma}$。

□

附录

附录分为两部分。在第 A.1 节中，我们汇总了关于 Riesz 核 K_s（其中 $s\ge 1$）的一些有用估计。接着，在第 A.2 节中，我们将一些分数阶 Kato–Ponce 对易子估计推广到周期情形。

A.1 Riesz 核估计

这里，对于任意 $s \geq 1$， K_s： $\mathbb{T}^d \to \mathbb{R}$ 是式（2.1）中定义的周期 Riesz 核。我们先给出以下引理，它精确描述了 K_s 在 0 处的渐近行为（例如参见 [Fah17，定理 34]，其中给出了更一般的结果）。

引理 A.1。 设 $s \geq 1$ 和 K_s 为 (2.1) 中定义的核。设 $\varphi \in C^\infty(\mathbb{T}^d;\ [0,1])$ 为截断函数，满足 $\varphi = 1$ 在 $B_{1/8}$ 中且 $\varphi = 0$ 在 $B_{1/4}$中。则有

$$K_s = c_{d,s} K_s^{\mathrm{sing}} \varphi + K_s^{\mathrm{reg}}, \qquad K_s^{\mathrm{sing}}(x) := \begin{cases} |x|^{2s-d} & \text{for } 2s-d \notin 2\mathbb{N} \cup \{0\}, \\ |x|^{2s-d}\log(|x|) & \text{for } 2s-d \in 2\mathbb{N} \cup \{0\}, \end{cases} \qquad K_s^{\mathrm{reg}} \in C^\infty(\mathbb{T}^d). \tag{A.1}$$

特别地，以下界成立：

$$|\nabla K_s(x)| \lesssim_{d,s} \begin{cases} |x|^{2s-d-1} & \text{for } s \in \left[1, \frac{d}{2}+\frac{1}{2}\right), \\ 1 & \text{for } s \geq \frac{d}{2}+\frac{1}{2}, \end{cases} \qquad |\nabla^2 K_s(x)| \lesssim_{d,s} \begin{cases} |x|^{2s-d-2} & \text{for } s \in \left[1, \frac{d}{2}+1\right), \\ \log(2+|x|^{-1}) & \text{for } s = \frac{d}{2}+1, \\ 1 & \text{for } s > \frac{d}{2}+1. \end{cases} \tag{A.2}$$

接下来，我们给出 $\nabla K_s * \eta$的连续模估计，其中 η 属于空间 $\mathscr{X}_s(\mathbb{T}^d)$，该空间定义于 (1.5) 中。

引理 A.2。 设 $s \geq 1$ 和 $\mathscr{X}_s(\mathbb{T}^d)$ 由 (1.5)定义。则对任意 $\eta \in \mathscr{X}_s(\mathbb{T}^d)$，有

$$|\nabla K_s * \eta(x) - \nabla K_s * \eta(y)| \lesssim_{d,s} \|\eta\|_{\mathscr{X}_s} \omega_s(|x-y|) \qquad \forall x, y \in \mathbb{T}^d. \tag{A.3}$$

其中，连续模 ω_s： $[0,\infty) \to [0,\infty)$ 定义为

$$\omega_s(r) := \begin{cases} r & \text{if } s \notin \{1, d/2+1\}, \\ r\log(2+r^{-1}) & \text{if } s \in \{1, d/2+1\}. \end{cases} \tag{A.4}$$

证明。 我们根据 $s \geq 1$ 的选择来组织证明。

($s = 1$). 取 $x, y \in \mathbb{T}^d$，记 $r := |x-y|$，并且不失一般性地假设 $r < 1/8$。于是，

$$|\nabla K_1 * \eta(x) - \nabla K_1 * \eta(y)| \leq \|\eta\|_{L^\infty} \int_{\mathbb{T}^d} |\nabla K_1(x-z) - \nabla K_1(y-z)| dz.$$

我们将右侧的积分拆分到两个区域 $B_{2r}(x)$ 和 $\mathbb{T}^d \setminus B_{2r}(x)$ 上。对于第一项，有

$$\begin{aligned} \int_{B_{2r}(x)} |\nabla K_1(x-z) - \nabla K_1(y-z)| dz &\leq \int_{B_{2r}(x)} |\nabla K_1(x-z)| dz + \int_{B_{3r}(x)} |\nabla K_1(y-z)| dz \\ &\lesssim_d \int_{B_{3r}} |w|^{1-d} \lesssim_d r, \end{aligned}$$

其中使用了（ $|\nabla K_s|$ A.2）中的界。另一方面，注意到对于任意 $z \in \mathbb{T}^d \setminus B_{2r}(x)$，根据中值定理，存在 p_z 位于连接 x 与 y 的线段上，使得

$$|\nabla K_1(x-z) - \nabla K_1(y-z)| \leq |\nabla^2 K_1(p_z - z)||x-y| \lesssim_d r|p_z - z|^{-d} \lesssim r|x-z|^{-d},$$

其中使用了（ $|\nabla^2 K_s|$A.2）中的界。因此：

$$\int_{\mathbb{T}^d \setminus B_{2r}(x)} |\nabla K_1(x-z) - \nabla K_1(y-z)| dz \lesssim_d r \int_{\mathbb{T}^d \setminus B_{2r}} |w|^{-d} \lesssim_d r\log(2+r^{-1}).$$

将这两个估计合并，便得到情形下的精确结果（A.3）： $s = 1$。

($1 < s < d/2+1$) 令 $p = d/(2s-2)$ 且 $q = p' = d/(d-2s+2)$。根据洛伦兹空间中的赫尔德不等式 s,

$$\|\nabla^2 K_s * \eta\|_{L^\infty} \lesssim_{d,s} \|\nabla^2 K_s\|_{L^{q,\infty}} \|\eta\|_{L^{p,1}} \lesssim_{d,s} \|\eta\|_{L^{p,1}},$$

其中我们使用了以下界 $|\nabla^2 K_s(x)| \lesssim_{d,\ s} |x|^{2s-d-2} \in L^{q,\ \infty}$，该界来自A.2。

($s = d/2+1$)。 在此情形下，A.1给出

$$|\nabla K_{d/2+1}(x) - \nabla K_{d/2+1}(y)| \lesssim_d |x-y| \log(2 + |x-y|^{-1}) \qquad \forall x, y \in \mathbb{T}^d.$$

因此，

$$|\nabla K_{d/2+1} * \eta(x) - \nabla K_{d/2+1} * \eta(y)| \lesssim_d \|\eta\|_{\mathcal{M}} |x-y| \log(2 + |x-y|^{-1}).$$

($s > d/2+1$)。 根据（A.2），有 $\|\nabla^2 K_s * \eta\|_{L^\infty} \le \|\nabla^2 K_s\|_{L^\infty} \|\eta\|_{\mathcal{M}} \lesssim_{d,\ s} \|\eta\|_{\mathcal{M}}$。 □

在接下来的两个引理中，我们将用 Wasserstein 距离 $W_2(\eta_1,\ \eta_2)$ 以及对 η_1 和 η_2 的适当可积性控制，给出 $\nabla K_s *(\eta_1 - \eta_2)$ 在勒贝格空间中的界。

引理 A.3. 设 $s \in [1, d/2+1)$、$p := d/(2s-2) \in (1、\infty]$、$q := p' = d/(d-2s+2) \in [1,\infty)$以及$\eta_1$、$\eta_2 \in \mathscr{P} \cap L^p(\mathbb{T}^d)$。则有以下估计：

$$\|\nabla K_s * (\eta_1 - \eta_2)\|_{L^{2q}} \lesssim_{d,s} \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\}^{1/2} W_2(\eta_1, \eta_2).$$

证明。 由于 $\eta_1, \eta_2 \ll \mathscr{L}^d$，存在唯一的最优传输映射 $T\colon \mathbb{T}^d \to \mathbb{T}^d$，其对应二次代价。记 $\xi\colon \mathbb{T}^d \to \mathbb{R}^d$ 为由位移 $\xi(x) = T(x) - x$ 给出的向量场，并令$X\colon [1,2] \times \mathbb{T}^d \to \mathbb{T}^d$为相应的流映射 $X_t(x) = (2-t)x + (t-1)T(x)$，使得 $X_1 = \mathrm{id}$ 且$X_2 = T$。由此定义位移插值

$$\eta_t := (X_t)_\# \eta_1 \qquad \forall t \in [1,2].$$

众所周知，$\eta \in AC([1,2]; \mathscr{P}(\mathbb{T}^d))$ 是连接 η_1 与 η_2 的瓦瑟斯坦空间中的唯一常速测地线，并且满足连续性方程

$$\partial_t \eta_t + \mathrm{div}(\eta_t w_t) = 0, \qquad w_t = \xi \circ X_t^{-1}, \qquad \|w_t\|_{L^2(\mathbb{T}^d, \eta_t; \mathbb{R}^d)} = W_2(\eta_1, \eta_2). \tag{A.5}$$

此外，插值 η_t 的 L^p-范数由极值 η_1，η_2 的范数所界：

$$\|\eta_t\|_{L^p} \le \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\} \qquad \forall t \in [1,2]. \tag{A.6}$$

我们可以在分布的意义下写成：

$$\nabla K_s * (\eta_1 - \eta_2) = \nabla K_s * \int_1^2 \partial_t \eta_t dt = -\int_1^2 \nabla K_s * \mathrm{div}(\eta_t w_t) dt = \int_1^2 \nabla^2 K_s * (\eta_t w_t) dt. \tag{A.7}$$

现在，设 $\beta := 2q/(2q-1) \in (1,2]$ 满足 $1/\beta = 1/2p + 1/2$。结合赫尔德不等式以及（A.5）和（A.6），我们得到

$$\|\eta_r w_r\|_{L^\beta} \le \|\eta_r\|_{L^p}^{1/2} \|\eta_r^{1/2} w_r\|_{L^2} \le \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\}^{1/2} W_2(\eta_1, \eta_2) \qquad \forall r \in [1,2].$$

因此，由（A.7），

$$\begin{aligned}\|\nabla K_s * (\eta_1 - \eta_2)\|_{L^{2q}} &\le \int_1^2 \|\nabla^2 K_s * (\eta_r w_r)\|_{L^{2q}} dr \\ &\lesssim_{d,s} \int_1^2 \|\eta_r w_r\|_{L^\beta} dr \le \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\}^{1/2} W_2(\eta_1, \eta_2),\end{aligned}$$

其中，在第二个不等式中我们使用了 $\nabla^2 K_s * = (-\Delta)^{1-s} \nabla^2 K_1 *$，并将 Calderón-Zygmund估计 $\nabla^2 K_1 *\colon L^{2q} \to L^{2q}$ 与零平均函数的索伯列夫嵌入 $(-\Delta)^{1-s}\colon\ L^{2q} \to L^\beta$结合起来。这就是所需的不等式，证明至此完成。

□

引理 A.4. 设 $s \geq 1$，且 $\eta_1, \eta_2 \in \mathscr{P}(\mathbb{T}^d)$。则以下估计成立：

i) 若 $s \in [1, d/2+1)$，$p = d/(2s-2)$ 且 $\eta_1, \eta_2 \in L^p(\mathbb{T}^d)$：

$$\|\nabla K_s * (\eta^1 - \eta^2)\|_{L^\infty} \lesssim_{d,s} \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\}^{\frac{d+2-2s}{d+3-2s}} W_2(\eta^1, \eta^2)^{\frac{1}{d+3-2s}}.$$

ii) 若 $s = d/2+1$：

$$\|\nabla K_s * (\eta_1 - \eta_2)\|_{L^\infty} \lesssim_d W_2(\eta_1, \eta_2) \log\left(2 + W_2(\eta_1, \eta_2)^{-1}\right).$$

iii) 若 $s > d/2+1$：

$$\|\nabla K_s * (\eta_1 - \eta_2)\|_{L^\infty} \lesssim_{d,s} W_2(\eta_1, \eta_2).$$

证明。我们根据 $s \geq 1$ 的选择来拆分证明。

($1 \leq s < d/2+1$)。设 $\psi \in C_c^\infty(B_{1/4}; [0,1])$ 是一个截断函数，满足 $\psi = 1$ 属于 $B_{1/8}$，并设 $\varepsilon \in (0,1)$ 为稍后选定的某个参数。对于固定的 $x \in \mathbb{T}^d$，我们可以拆分

$$\begin{aligned}\nabla K_s * (\eta_1 - \eta_2)(x) = &\int_{\mathbb{T}^d} \nabla K_s(x-y)(\eta_1 - \eta_2)(y)\psi\left(\varepsilon^{-1}(x-y)\right) \\ &+ \int_{\mathbb{T}^d} \nabla K_s(x-y)(\eta_1 - \eta_2)(y)\left(1 - \psi\left(\varepsilon^{-1}(x-y)\right)\right) = (I) + (II).\end{aligned}$$

由于 $p = d/(2s-2)$ 且 $q = p' = d/(d-2s+2)$，对于第一项，我们使用赫尔德不等式以及 (A.2) 中关于 $|\nabla K_s|$ 的界：

$$\begin{aligned}|(I)| &\leq \int_{B_{\varepsilon/4}(x)} |\nabla K_s(x-y)||(\eta_1 - \eta_2)(y)|dy \lesssim \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\}\left(\int_{B_{\varepsilon/4}} |\nabla K_s|^q\right)^{1/q} \\ &\lesssim_{d,s} \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\}\left(\int_{B_{\varepsilon/4}} |x|^{q(2s-d-1)}\right)^{1/q} \lesssim_{d,s} \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\}\varepsilon,\end{aligned}$$

为了界定第二项，注意函数 $\nabla K_s(x-\cdot)\left(1 - \psi\left(\varepsilon^{-1}(\cdot - x)\right)\right)$ 的利普希茨常数为 $\lesssim_{d,s} \varepsilon^{2s-d-2}$ 这是由 (A.2) 得出的。因此：

$$|(II)| \lesssim_{d,s} \varepsilon^{2s-d-2} \sup_{\mathrm{Lip}(f)\leq 1} \int_{\mathbb{T}^d} f(\eta_1 - \eta_2) = \varepsilon^{2s-d-2} W_1(\eta_1, \eta_2) \leq \varepsilon^{2s-d-2} W_2(\eta_1, \eta_2),$$

这里我们用到了著名的康托罗维奇-鲁宾斯坦对偶公式来处理 1-瓦瑟斯坦距离 W_1，以及平凡估计 $W_1 \leq W_2$。选择 $c = c(d)$，使得

$$\varepsilon = c\left(\frac{W_2(\eta_1, \eta_2)}{\max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\}}\right)^{\frac{1}{d+3-2s}} \in (0,1),$$

并汇总 (I) 和 (II) 的估计，我们得到

$$\begin{aligned}\|\nabla K_s * (\eta_1 - \eta_2)\|_{L^\infty} &\lesssim_{d,s} \varepsilon \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\} + \varepsilon^{2s-d-2} W_2(\eta_1, \eta_2) \\ &\lesssim_d \max\{\|\eta_1\|_{L^p}, \|\eta_2\|_{L^p}\}^{\frac{d+2-2s}{d+3-2s}} W_2(\eta_1, \eta_2)^{\frac{1}{d+3-2s}},\end{aligned}$$

正如所需。

($s = d/2+1$)。采用上文相同的论证，并选择 $\varepsilon = cW_2(\eta_1, \eta_2)\log\left(2 + W_2(\eta_1, \eta_2)^{-1}\right) \in (0,1)$，可得

$$\begin{aligned}\|\nabla K_s * (\eta_1 - \eta_2)\|_{L^\infty} &\lesssim_{d,s} \varepsilon + \log(2+\varepsilon^{-1}) W_2(\eta_1, \eta_2) \\ &\lesssim_d W_2(\eta_1, \eta_2)\log\left(2 + W_2(\eta_1, \eta_2)^{-1}\right).\end{aligned}$$

($s > d/2+1$)。在这种情况下 $\|\nabla^2 K_s\|_{L^\infty} \lesssim_{d,s} 1$。因此，对于任意 $x \in \mathbb{T}^d$，

$$|\nabla K_s * (\eta_1 - \eta_2)(x)| = \left|\int_{\mathbb{T}^d} \nabla K_s(x-y)(\eta_1 - \eta_2)dy\right| \lesssim_{d,s} \sup_{\mathrm{Lip}(f)\leq 1} \int f(\eta_1 - \eta_2) \leq W_2(\eta_1, \eta_2),$$

如所需。 □

最后，我们借助一个略强范数的对数，给出临界空间中的 L^∞ 界。

引理 A.5。 设 $s \geq 1$。以下结论成立：

i) 若 $s = 1$ $\alpha \in (0, 1)$，且 $f \in C^{0,\alpha}(\mathbb{T}^d)$，则

$$\|\nabla^2 K_1 * f\|_{L^\infty} \lesssim_d \|f\|_{L^\infty} \left(1 + \log\left(1 + \frac{[f]_{C^{0,\alpha}}}{\|f\|_{L^\infty}}\right)\right).$$

ii) 若 $s \in (1, d/2+1]$，$p = d/(2s-2) \in [1, \infty)$，且 $f \in L^r(\mathbb{T}^d)$ 对某个 $r > p$ 成立，则

$$\|\nabla^2 K_s * f\|_{L^\infty} \lesssim_{d,s,r} \|f\|_{L^p} \left(1 + \log\left(1 + \frac{\|f\|_{L^r}}{\|f\|_{L^p}}\right)\right).$$

证明。 我们首先考虑 $s = 1$ 的情形。令 $\psi \in C_c^\infty(B_{1/4}; [0,1])$ 为一个径向截断函数，使得 $\psi = 1$ 属于 $B_{1/8}$，并令 $\varepsilon \in (0, 1)$ 为稍后选取的参数。对于给定的 $x \in \mathbb{T}^d$，我们进行分解

$$\nabla^2 K_1 * f(x) = \underbrace{\int_{B_{\varepsilon/4}(x)} \nabla^2 K_1(x-y)\psi((y-x)/\varepsilon) f(y)}_{(I)} + \underbrace{\int_{\mathbb{T}^d \setminus B_{\varepsilon/4}(x)} \nabla^2 K_1(x-y)\left(1 - \psi((y-x)/\varepsilon)\right) f(y)}_{(II)}.$$

由（A.2），我们可以估计 (II)，其中

$$|(II)| \lesssim_d \|f\|_{L^\infty} \int_{\mathbb{T}^d \setminus B_{\varepsilon/4}(x)} |y-x|^{-d} \lesssim_d \|f\|_{L^\infty} \log(1 + \varepsilon^{-1}).$$

为估计 (I)，注意到由对称性，$\nabla^2 K_1$ 在 ∂B_t 中的平均值为零，对所有 $t \in (0, 1/4)$ 均成立。由于 ψ 是径向的，我们推得 $\nabla^2 K_1(x-\cdot)\,\psi((\cdot - x)/\varepsilon)$ 在 $B_{\varepsilon/4}(x)$ 中的平均值为零。因此，利用（A.2），我们可以估计

$$|(I)| = \left| \int_{B_{\varepsilon/4}(x)} \nabla^2 K_1(x-y)\psi((y-x)/\varepsilon)(f(y) - f(x)) \right| \lesssim_d [f]_{C^{0,\alpha}} \int_{B_{\varepsilon/4}(x)} |y-x|^{\alpha - d} \lesssim_{d,\alpha} [f]_{C^{0,\alpha}} \varepsilon^\alpha.$$

结合这两个估计并取 $\varepsilon = \min\{1/2, (\|f\|_{L^\infty}/[f]_{C^{0,\alpha}})^{1/\alpha}\}$，我们得到所需的估计。对于 $s \in (1, d/2+1]$，可以采用相同的方法处理，利用赫尔德不等式和（A.2）得到上述两个项（I）、（II）的类似估计，其中用 $\|f\|_{L^p}$ 替代 $\|f\|_{L^\infty}$，用 $\|f\|_{L^r}$ 替代 $[f]_{C^{0,\alpha}}$，其中 $p = d/(2s-2)$ 且 $r > p$。 □

A.2 Kato–Ponce交换子估计：在 d 维环面上

本节将 Kato–Ponce 分数阶交换子估计（[KP88、Li19]）推广到周期情形。我们采用 [BOZ25], 提出的策略，通过泊松求和公式[8]将问题化约为 $\mathbb{R}^d$ 中的相应估计[8]。以下，对于任意 $\beta \in \mathbb{R}$，记 $D^\beta_{\mathbb{T}^d} := (-\Delta_{\mathbb{T}^d})^{\beta/2}$。

引理 A.6。 设 $\gamma > 0$ 和 $2 \leq p_1, q_1, p_2, q_2 \leq \infty$ 满足 $1/2 = 1/p_i + 1/q_i$, $i = 1, 2$。那么，对于每个 $f, g \in C^\infty(\mathbb{T}^d)$，我们有

$$\|D^\gamma_{\mathbb{T}^d}(fg) - f D^\gamma_{\mathbb{T}^d} g\|_{L^2(\mathbb{T}^d)} \lesssim_{d,\ \gamma,\ p_1,\ q_1,\ p_2,\ q_2} \|\nabla f\|_{L^{p_1}(\mathbb{T}^d)} \|D^{\gamma-1}_{\mathbb{T}^d} g\|_{L^{q_1}(\mathbb{T}^d)} + \|D^\gamma_{\mathbb{T}^d} f\|_{L^{p_2}(\mathbb{T}^d)} \|g\|_{L^{q_2}(\mathbb{T}^d)}. \tag{A.8}$$

[8]对于每个施瓦茨函数 $u \in \mathcal{S}(\mathbb{R}^d)$:

$$\sum_{k \in \mathbb{Z}^d} \mathcal{F}_{\mathbb{R}^d}(u)(k) e^{2\pi i k \cdot x} = \sum_{k \in \mathbb{Z}^d} u(x+k) \qquad \forall x \in \mathbb{R}^d.$$

证明。 我们将通过将结果归约为 $\mathbb{R}^d$中的相应不等式来证明该结果，该不等式已在 [Li19的推论 5 中证明。2]不失一般性，我们可以假设 $f, g \in C^\infty(\mathbb{T}^d)$ 在 $\mathbb{T}^d$中具有零平均值。我们分别称 $F, G \in C^\infty(\mathbb{R}^d)$ 为 f, g的周期延拓。我们考虑一个周期单位分解，即由以下函数构成的集合：1-周期函数 $\varphi_1, \dots, \varphi_N \in C^\infty(\mathbb{R}^d; [0,1])$ 满足

$$\sum_{j=1}^N \varphi_j(x) = 1 \quad \forall x \in \mathbb{R}^d, \qquad \operatorname{supp} \varphi_j \subset \bigcup_{k \in \mathbb{Z}^d} \left\{ Q_{1/2}(\bar{x}_j) + k \right\} \quad \forall j \in \{1, \dots, N\},$$

其中 $Q_{1/2}(\bar{x}_j)$ 是以 $\bar{x}_j \in [0,1]^d$ 为中心、边长为 $1/2$ 的开立方体。我们将 $G\varphi_j$ 的单周期局部化记为 G_j：

$$G_j := G\varphi_j \mathbb{1}_{Q_1(\bar{x}_j)} \qquad \forall j \in \{1, \dots, N\}.$$

对于每个 $k \in \mathbb{Z}^d$，我们可以写成

$$\begin{aligned}
\mathcal{F}_{\mathbb{T}^d}(g)(k) &= \int_{\mathbb{T}^d} e^{-2\pi i k \cdot x} g(x) = \int_{[0,1]^d} e^{-2\pi i k \cdot x} G(x) \\
&= \sum_{j=1}^N \int_{[0,1]^d} e^{-2\pi i k \cdot x} G(x) \varphi_j(x) = \sum_{j=1}^N \int_{Q_1(\bar{x}_j)} e^{-2\pi i k \cdot x} G(x) \varphi_j(x) = \sum_{j=1}^N \mathcal{F}_{\mathbb{R}^d}(G_j)(k),
\end{aligned}$$

其中 $\mathcal{F}_{\mathbb{T}^d}$ 和 $\mathcal{F}_{\mathbb{R}^d}$ 分别表示关于 $\mathbb{T}^d$ 和 $\mathbb{R}^d$ 的傅里叶变换。同样，

$$\mathcal{F}_{\mathbb{T}^d}(fg)(k) = \sum_{j=1}^N \mathcal{F}_{\mathbb{R}^d}(FG_j)(k).$$

特别地，交换子 $[D^\gamma_{\mathbb{T}^d}, f](g)(x)$ 可以写成如下形式：

$$\begin{aligned}
D^\gamma_{\mathbb{T}^d}(fg)(x) - f D^\gamma_{\mathbb{T}^d} g(x) &= \sum_{k \in \mathbb{Z}^d} (2\pi|k|)^\gamma \left[\mathcal{F}_{\mathbb{T}^d}(fg)(k) - f(x) \mathcal{F}_{\mathbb{T}^d}(g)(k) \right] e^{2\pi i k \cdot x} \\
&= \sum_{j=1}^N \sum_{k \in \mathbb{Z}^d} (2\pi|k|)^\gamma \left[\mathcal{F}_{\mathbb{R}^d}(FG_j)(k) - F(x) \mathcal{F}_{\mathbb{R}^d}(G_j)(k) \right] e^{2\pi i k \cdot x} \\
&= \sum_{j=1}^N \sum_{k \in \mathbb{Z}^d} \left[\mathcal{F}_{\mathbb{R}^d}\left(D^\gamma_{\mathbb{R}^d}(FG_j) \right)(k) - F(x) \mathcal{F}_{\mathbb{R}^d}\left(D^\gamma_{\mathbb{R}^d} G_j \right) \right] e^{2\pi i k \cdot x} \\
&= \sum_{j=1}^N \sum_{k \in \mathbb{Z}^d} \left[D^\gamma_{\mathbb{R}^d}(FG_j) - F D^\gamma_{\mathbb{R}^d} G_j \right](x + k),
\end{aligned}$$

其中 $D^\gamma_{\mathbb{R}^d} := (-\Delta_{\mathbb{R}^d})^{\gamma/2}$，并且我们在最后一步使用了泊松求和公式。特别地，我们可以给出上界

$$\begin{aligned}
\| D^\gamma_{\mathbb{T}^d}(fg) - f D^\gamma_{\mathbb{T}^d} g \|_{L^2(\mathbb{T}^d)} &\lesssim_d \sum_{j=1}^N \left\| \sum_{k \in \mathbb{Z}^d} \left[D^\gamma_{\mathbb{R}^d}(FG_j) - F D^\gamma_{\mathbb{R}^d} G_j \right](\cdot + k) \right\|_{L^2([0,1]^d)} \\
&\lesssim_d \sum_{j=1}^N \sum_{\substack{k \in \mathbb{Z}^d \\ |k| \le 2\sqrt{d}}} \left\| \left[D^\gamma_{\mathbb{R}^d}(FG_j) - F D^\gamma_{\mathbb{R}^d} G_j \right](\cdot + k) \right\|_{L^2([0,1]^d)} \\
&\quad + \sum_{j=1}^N \left\| \sum_{\substack{k \in \mathbb{Z}^d \\ |k| > 2\sqrt{d}}} \left[D^\gamma_{\mathbb{R}^d}(FG_j) - F D^\gamma_{\mathbb{R}^d} G_j \right](\cdot + k) \right\|_{L^2([0,1]^d)} =: (I) + (II).
\end{aligned}$$

我们来估计 (I)。对每个 $j\in\{1,\dots,N\}$，我们取一个截断函数 $\psi_j\in C_c^\infty(Q_{4/5}(\bar{x}_j);[0,1])$，使得$\psi_j=1$在 $Q_{3/4}(\bar{x}_j)$ 上成立，并记 $F_j:=F\psi_j$，从而 $FG_j=F_jG_j$。设 $k\in\mathbb{Z}^d$ 满足 $|k|\le 2\sqrt{d}$，则

$$\begin{aligned}\left\|\left[D^\gamma_{\mathbb{R}^d}(FG_j)-FD^\gamma_{\mathbb{R}^d}G_j\right](\cdot+k)\right\|_{L^2([0,1]^d)}&\le\left\|D^\gamma_{\mathbb{R}^d}(FG_j)-FD^\gamma_{\mathbb{R}^d}G_j\right\|_{L^2(\mathbb{R}^d)}\\&\le\left\|D^\gamma_{\mathbb{R}^d}(F_jG_j)-F_jD^\gamma_{\mathbb{R}^d}G_j\right\|_{L^2(\mathbb{R}^d)}+\left\|F(1-\psi_j)D^\gamma_{\mathbb{R}^d}G_j\right\|_{L^2(\mathbb{R}^d)}.\end{aligned}$$

为了估计最后一个不等式右侧的第一项，我们使用 [Li19 的欧几里得对易子估计，推论 52]:

$$\begin{aligned}\left\|D^\gamma_{\mathbb{R}^d}(F_jG_j)-F_jD^\gamma_{\mathbb{R}^d}G_j\right\|_{L^2(\mathbb{R}^d)}&\lesssim_{d,\ \gamma,\ p_1,\ q_1,\ p_2,\ q_2}\|\nabla F_j\|_{L^{p_1}(\mathbb{R}^d)}\|D^{\gamma-1}_{\mathbb{R}^d}G_j\|_{L^{q_1}(\mathbb{R}^d)}+\|D^\gamma_{\mathbb{R}^d}F_j\|_{L^{p_2}(\mathbb{R}^d)}\|G_j\|_{L^{q_2}(\mathbb{R}^d)}\\&\lesssim_{d,\gamma,p_1,q_1,p_2,q_2}\|\nabla f\|_{L^{p_1}(\mathbb{T}^d)}\|D^{\gamma-1}_{\mathbb{T}^d}g\|_{L^{q_1}(\mathbb{T}^d)}+\|D^\gamma_{\mathbb{T}^d}f\|_{L^{p_2}(\mathbb{T}^d)}\|g\|_{L^{q_2}(\mathbb{T}^d)},\end{aligned}$$

接下来，我们考虑第二项。记 $\mathcal{K}_\gamma:\mathbb{R}^d\to\mathbb{R}$ 为与算子 $D^\gamma_{\mathbb{R}^d}$[9] 相关联的核，

我们有

$$|\mathcal{K}_\gamma(x)|\lesssim_{d,\gamma}|x|^{-d-\gamma}\qquad\forall x\in\mathbb{R}^d.$$

特别地，利用赫尔德不等式，以及 $\psi_j\equiv 1$ 在 $Q_{3/4}(\bar{x}_j)$ 上和 $\operatorname{supp}G_j\subset Q_{1/2}(\bar{x}_j)$，可得

$$\begin{aligned}\|F(1-\psi_j)D^\gamma_{\mathbb{R}^d}G_j\|_{L^2}&\le\left(\sum_{\ell\in\mathbb{Z}^d}\int_{(\ell+[0,1]^d)\setminus Q_{3/4}(\bar{x}_j)}(F\mathcal{K}_\gamma*G_j)^2\right)^{1/2}\\&\le\left(\sum_{\ell\in\mathbb{Z}^d}\|F\|^2_{L^{p_2}\left((\ell+[0,1]^d)\setminus Q_{3/4}(\bar{x}_j)\right)}\|\mathcal{K}_\gamma*G_j\|^2_{L^{q_2}\left((\ell+[0,1]^d)\setminus Q_{3/4}(\bar{x}_j)\right)}\right)^{1/2}\\&\lesssim_{d,\gamma}\|f\|_{L^{p_2}(\mathbb{T}^d)}\left(\sum_{\ell\in\mathbb{Z}^d}|\ell|^{-2d-2\gamma}\|g\|^2_{L^{q_2}(\mathbb{T}^d)}\right)^{1/2}\lesssim_{d,\gamma}\|D^\gamma_{\mathbb{T}^d}f\|_{L^{p_2}(\mathbb{T}^d)}\|g\|_{L^{q_2}(\mathbb{T}^d)}.\end{aligned}$$

因此:

$$(I)\lesssim_{d,\gamma,p_1,q_1,p_2,q_2}\|\nabla f\|_{L^{p_1}(\mathbb{T}^d)}\|D^{\gamma-1}_{\mathbb{T}^d}g\|_{L^{q_1}(\mathbb{T}^d)}+\|D^\gamma_{\mathbb{T}^d}f\|_{L^{p_2}(\mathbb{T}^d)}\|g\|_{L^{q_2}(\mathbb{T}^d)}.$$

最后，我们来界定（II）。对于任意满足$|k|>2\sqrt{d}$的 $k\in\mathbb{Z}^d$、任意 $x\in[0,1]^d$，以及任意满足 $\operatorname{supp}G_j$的 $y\in$，都有 $|x+k-y|\gtrsim|k|$。因此，

$$\begin{aligned}\left|\left[D^\gamma_{\mathbb{R}^d}(FG_j)-FD^\gamma_{\mathbb{R}^d}G_j\right](x+k)\right|&=\left|\int_{\mathbb{R}^d}K_\gamma(x+k-y)\left[F(y)-F(x)\right]G_j(y)dy\right|\\&\lesssim|k|^{-d-\gamma}\int_{\mathbb{R}^d}(|F(x)|+|F(y)|)\,|G_j(y)|dy.\end{aligned}$$

对 k 求和，得到

$$\left|\sum_{\substack{k\in\mathbb{Z}^d\\|k|>2\sqrt{d}}}\left[D^\gamma_{\mathbb{R}^d}(FG_j)-FD^\gamma_{\mathbb{R}^d}G_j\right](x+k)\right|\lesssim\int_{\mathbb{R}^d}(|F(x)|+|F(y)|)\,|G_j(y)|dy\le(|F(x)|+\|f\|_{L^{p_2}(\mathbb{T}^d)})\|g\|_{L^{q_2}(\mathbb{T}^d)}.$$

我们得出

$$(II)\lesssim\|f\|_{L^{p_2}(\mathbb{T}^d)}\|g\|_{L^{q_2}(\mathbb{T}^d)}\lesssim\|D^\gamma_{\mathbb{T}^d}f\|_{L^{p_2}(\mathbb{T}^d)}\|g\|_{L^{q_2}(\mathbb{T}^d)}.$$

结合(I)和(II)的界，我们得到（A.8），证明至此完成。 □

注A.7。当 $\gamma\in(0,1]$ 和 $2\le p,\ q\le\infty$ 满足 $1/p+1/q=1/2$时，如上所述，利用A.8的对应欧几里得界，可以得到以下改进： [Li19，引理 5。2]:

$$\|D_{\mathbb{T}^d}(fg)-fD_{\mathbb{T}^d}g\|_{L^2(\mathbb{T}^d)}\lesssim_{d,p,q}\|\nabla f\|_{L^p(\mathbb{T}^d)}\|g\|_{L^q(\mathbb{T}^d)},\tag{A.9}$$

$$\|D^\gamma_{\mathbb{T}^d}(fg)-fD^\gamma_{\mathbb{T}^d}g\|_{L^2(\mathbb{T}^d)}\lesssim_{d,\gamma,p,q}\|D^\gamma_{\mathbb{T}^d}f\|_{L^p(\mathbb{T}^d)}\|g\|_{L^q(\mathbb{T}^d)}\qquad\text{for }\gamma\in(0,1).\tag{A.10}$$

■

[9] $\mathcal{K}_\gamma$ 满足：对于每个 $u\in\mathcal{S}(\mathbb{R}^d)$，都有 $D^\gamma_{\mathbb{R}^d}u=\mathcal{K}_\gamma*u$。

注记 A.8。 由定理 A.6 以及 Gagliardo–Nirenberg（Brezis–Mironescu [BM18],分数阶情形下的）插值不等式可直接推出，对于每个 $f \in C^\infty(\mathbb{T}^d)$，我们得到

$$\|D^\gamma_{\mathbb{T}^d}(\nabla f D^\beta_{\mathbb{T}^d} f) - \nabla f D^{\gamma+\beta}_{\mathbb{T}^d} f\|_{L^2} \lesssim_{d,\gamma,\beta} \|\nabla^2 f\|_{L^\infty} \|D^{\gamma+\beta-1}_{\mathbb{T}^d} f\|_{L^2}, \qquad \forall \gamma > 1, \quad \forall \beta \geq 2. \tag{A.11}$$

事实上，取

$$p = \frac{2(\gamma+\beta-3)}{\gamma-1}, \qquad q = \frac{2(\gamma+\beta-3)}{\beta-2},$$

由分数阶 Gagliardo–Nirenberg 不等式 [BM18],

$$\|D^\gamma_{\mathbb{T}^d} \nabla f\|_{L^p} \|D^\beta_{\mathbb{T}^d} f\|_{L^q} \lesssim \|\nabla^2 f\|^{\frac{2}{q}}_{L^\infty} \|D^{\gamma+\beta-1}_{\mathbb{T}^d} f\|^{\frac{2}{p}}_{L^2} \|\nabla^2 f\|^{\frac{2}{p}}_{L^\infty} \|D^{\gamma+\beta-1}_{\mathbb{T}^d} f\|^{\frac{2}{q}}_{L^2} = \|\nabla^2 f\|_{L^\infty} \|D^{\gamma+\beta-1}_{\mathbb{T}^d} f\|_{L^2}.$$

现在，(A.11) 可由 (A.8) 得出，取$(p_1,\ q_1,\ p_2,\ q_2) = (2,\ \infty,\ p,\ q)$。 ■

致谢。M.C.、R.C. 和 X.F. 获得瑞士教育、研究与创新国务秘书处（SERI）的资助，合同编号为 MB22.00034，资助项目为 TENSE。R.C. 和 X.F. 还获得瑞士国家科学基金会（SNF，资助编号 PZ00P 2_208930）的资助。X.F. 另获得 AEI 项目 PID2021-125021NA-I00（西班牙）的资助。

参考文献

[AB08] Luigi Ambrosio 和 Patrick Bernard。Osgood 向量场连续性方程的有符号测度解的唯一性。*Rend. Lincei Mat. Appl.*，19(3):237–245，2008 年。[ABS21] Luigi Ambrosio、Elia Brué 和 Daniele Semola。最优传输讲义，Unitext 系列第 130 卷。*Unitext*。施普林格，2021 年。[AGS08] Luigi Ambrosio、Nicola Gigli 和 Giuseppe Savaré。梯度流：度量空间与概率测度空间中的梯度流。《苏黎世联邦理工学院数学讲义》。Birkhäuser Basel，2008 年。[AHS23] Fabian Altekrüger、Johannes Hertrich 和 Gabriele Steidl。具有 Riesz 核差异度的神经网络 Wasserstein 梯度流。载于 第 *40* 届国际机器学习会议（*ICML*）论文集，第 202 卷，第 664–690 页。PMLR，2023 年。[AKSG19] Michael Arbel、Anna Korba、Adil Salim 和 Arthur Gretton。最大均值差异梯度流。神经信息处理系统进展，32:6484–6494，2019 年。[AS21] Shunta Akiyama 和 Taiji Suzuki。关于教师—学生设定中通过梯度法学习两层 ReLU 神经网络，2021 年。[Bac17] Francis Bach。利用凸神经网络打破维数灾难。机器学习研究杂志，18(19):1–53，2017 年。[BB21] Alberto Bietti 和 Francis Bach。在核区间内，深层 ReLU 网络等价于浅层网络。载于国际学习表征会议，2021 年。[BLL12] Andrea L. Bertozzi、Thomas Laurent 和 Flavien Léger。通过牛顿势实现聚集与扩散：斑块解的动力学。*Math. Models Methods Appl. Sci.*，22(SUPPL.1):1140005，2012 年。[BLR11] Andrea L. Bertozzi、Thomas Laurent 和 Jesús Rosado。多维聚集方程的 Lp 理论。《纯粹与应用数学通讯》，64(1):45–83，2011 年。[BM18] Haïm Brezis 和 Petru Mironescu。Gagliardo–Nirenberg 不等式与非不等式：完整论述。*Ann. Inst. H. Poincaré C Anal. Non Linéaire*，35(5):1355–1376，2018 年。

[BOZ25] Árpád Bényi、Tadahiro Oh 和 Tengfei Zhao。环面上的分数阶 Leibniz 法则。《美国数学会会刊》，153(1):207–221，2025 年。[Bui25] The Anh Bui。超越欧几里得空间的双线性与分数阶 Leibniz 法则：加权Besov空间与 Triebel–Lizorkin估计。预印本：arXiv:2510.01942，2025 年。[BV25] Siwan Boufadène 和 François-Xavier Vialard。关于库仑差异度 Wasserstein 梯度流的全局收敛。《SIAM数学分析杂志》，57(4):4556–4587，2025 年。[CB18] Lénaïc Chizat 和 Francis Bach。利用最优传输研究过参数化模型中梯度下降的全局收敛。神经信息处理系统进展，31:3040–3050，2018 年。[CB20] Lénaïc Chizat 和 Francis Bach。使用逻辑损失训练的宽两层神经网络中梯度下降的隐式偏差。载于学习理论会议（COLT），第 125 卷，第 1305–1338 页，2020 年。[Chi22] Lénaïc Chizat。利用过参数化梯度下降进行基于测度的稀疏优化。数学规划，194(1–2):487–532，2022 年。[CLM13] Gianluca Crippa 和 Magali Lécureux-Mercier。具有非局部流的连续性方程组的测度解的存在性与唯一性。Nonlinear Differ. Equ. Appl.，20:523–537，2013 年。[CMG+25] Zonghao Chen、Aratrika Mustafi、Pierre Glaser、Anna Korba、Arthur Gretton 和 Bharath K. Sriperumbudur。（去）正则化最大均值差异梯度流。机器学习研究杂志，26(235):1–77，2025 年。[CS09] Youngmin Cho 和 Lawrence K. Saul。用于深度学习的核方法。神经信息处理系统进展，22:342–350，2009 年。[CS11] Youngmin Cho 和 Lawrence K. Saul。反余弦核在大间隔分类中的分析与扩展。预印本：arXiv:1112.3712，2011 年。[DSB+25] Richard Duong、Viktor Stein、Robert Beinert、Johannes Hertrich 和 Gabriele Steidl。具有距离核的 MMD泛函的 Wasserstein梯度流及分位函数上的柯西问题。ESAIM Control Optim. Calc. Var.，31:97，2025 年。[DX13] Feng Dai 和 Yuan Xu。球面与球上的逼近论与调和分析。施普林格数学专著。施普林格，2013 年。[Fah17] Matthias Albrecht Fahrenwaldt。闭流形上伪微分算子的非对角热核渐近与从属布朗运动。Integral Equations Oper. Theory，87(3):327–347，2017 年。[FG21] Alessio Figalli 和 Federico Glaudo。最优传输、Wasserstein距离与梯度流导论。数学 EMS 教科书。EMS出版社，2021 年。[FRF22] Xavier Fernández-Real 和 Alessio Figalli。浅层神经网络作为 Wasserstein 型梯度流的连续表述。载于《大尺度分析：献给让·布尔甘的生命与工作》，第 29–57 页。施普林格，2022 年。[GBR+06] Arthur Gretton、Karsten M. Borgwardt、Malte Rasch、Bernhard Schölkopf 和 Alexander J. Smola。用于双样本问题的核方法。神经信息处理系统进展，19:513–520，2006 年。[GdBG25] Alexandre Galashov、Valentin de Bortoli 和 Arthur Gretton。无需对抗训练的深度 MMD 梯度流。载于国际学习表征会议，2025 年。

[GM17] Thomas Gallouët 和 Leonard Monsaingeon。Kantorovich–Fisher–Rao 梯度流的 JKO 分裂格式。《SIAM数学分析杂志》，49(2):1100–1130，2017年。[GW92] Paul Goodey 和 Wolfgang Weil。中心对称凸体与球面 Radon 变换。《微分几何杂志》，35(3):675–688，1992年。[HGBS24] Johannes Hertrich、Manuel Gräf、Robert Beinert 和 Gabriele Steidl。基于 Riesz 核的差异度 Wasserstein 最速下降流。《数学分析与应用杂志》，531(1):127829，2024年。[HHA+24] Paul Hagemann、Johannes Hertrich、Fabian Altekrüger、Robert Beinert、Jannis Chemseddine 和 Gabriele Steidl。基于负距离核的 MMD 梯度流进行后验采样。载于国际学习表征会议，2024年。[HWAH24] Johannes Hertrich、Christian Wald、Fabian Altekrüger 和 Paul Hagemann。基于 Riesz 核的生成式切片 MMD 流。载于国际学习表征会议，2024年。[KP88] Tosio Kato 和 Gustavo Ponce。Euler 方程与 Navier–Stokes 方程的交换子估计。《纯粹与应用数学通讯》，41(7):891–907，1988年。[Li19] Dong Li。关于 Kato–Ponce 估计与分数阶 Leibniz 法则。《伊比利亚美洲数学杂志》，35(1):23–100，2019年。[LMS23] Matthias Liero、Alexander Mielke 和 Giuseppe Savaré。Hellinger–Kantorovich 距离下测地线的精细性质与测地线 λ-凸性。《有理力学与分析档案》，247(6):112，2023年。[LMZ20] Yuanzhi Li、Tengyu Ma 和 Hongyang R. Zhang。超越神经正切核，学习过参数化双层神经网络。载于学习理论会议（COLT），第125卷，第2613–2682页，2020年。[Loe06] Grégoire Loeper。具有有界密度的 Vlasov–Poisson 系统解的唯一性。《纯粹与应用数学杂志》，86(1):68–79，2006年。[MMM19] Song Mei、Theodor Misiakiewicz 和 Andrea Montanari。双层神经网络的平均场理论：无维数依赖界与核极限。载于学习理论会议（COLT），第99卷，第2388–2464页，2019年。[MMN18] Song Mei、Andrea Montanari 和 Phan-Minh Nguyen。双层神经网络景观的平均场视角。《美国国家科学院院刊》，115(33):E7665–E7671，2018年。[MP94] Carlo Marchioro 和 Mario Pulvirenti。《不可压缩无黏流体的数学理论》，应用数学科学第96卷。施普林格，1994年。[PR14] Benedetto Piccoli 和 Francesco Rossi。广义 Wasserstein 距离及其在带源项输运方程中的应用。《有理力学与分析档案》，211(1):335–358，2014年。[RVE22] Grant Rotskoff 和 Eric Vanden-Eijnden。人工神经网络的可训练性与准确性：一种相互作用粒子系统方法。《纯粹与应用数学通讯》，75(9):1889–1935，2022年。[Ser24] Sylvia Serfaty。关于库仑气体与 Riesz 气体的讲义。预印本：arXiv:2407.21194，2024年。[SR13] Gábor J. Székely 和 Maria L. Rizzo。能量统计：一类基于距离的统计量。《统计规划与推断杂志》，143(8):1249–1272，2013年。[SS18] Itay Safran 和 Ohad Shamir。伪局部极小值在两层 ReLU 神经网络中很常见。载于国际机器学习会议（ICML），第80卷，第4433–4441页，2018年。[SS20] Justin Sirignano 和 Konstantinos Spiliopoulos。神经网络的平均场分析：大数定律。《SIAM应用数学杂志》，80(2):725–752，2020年。

[UNS+18] Thomas Unterthiner、Bernhard Nessler、Calvin Seward、Günter Klambauer、Martin Heusel、Hubert Ramsauer 和 Sepp Hochreiter。库仑生成对抗网络：通过势场实现可证明最优的纳什均衡。载于国际学习表征会议，2018年。[Wai19] Martin J. Wainwright。高维统计学：非渐近视角，剑桥统计与概率数学系列第48卷。剑桥大学出版社，2019年。[Woj20] Stephan Wojtowytsch。关于平均场状态下两层 ReLU 神经网络梯度下降训练的收敛。预印本：*arXiv:2005.13530a*，2020年。[Yud63] Victor I. Yudovich。理想不可压缩液体的非定常流。《苏联计算数学与数学物理》，3(6):1407–1456，1963年。[ZG24] Mo Zhou 和 Rong Ge。梯度下降如何学习特征——正则化双层神经网络的局部分析。神经信息处理系统进展，37:46481–46538，2024年。[ZGJ21] Mo Zhou、Rong Ge 和 Chi Jin。温和过参数化双层神经网络的局部收敛理论。载于学习理论会议（*COLT*），第134卷，第4577–4632页，2021年。