Analysis & PDE

José A. Carrillo, Matias G. Delgadino,
Laurent Desvillettes and Jeremy S.-H. Wu

THE LANDAU EQUATION AS A GRADIENT FLOW

msp

THE LANDAU EQUATION AS A GRADIENT FLOW

JOSÉ A. CARRILLO, MATIAS G. DELGADINO, LAURENT DESVILLETTES AND JEREMY S.-H. WU

We propose a gradient flow perspective to the spatially homogeneous Landau equation for soft potentials. We construct a tailored metric on the space of probability measures based on the entropy dissipation of the Landau equation. Under this metric, the Landau equation can be characterized as the gradient flow of the Boltzmann entropy. In particular, we characterize the dynamics of the PDE through a functional inequality which is usually referred as the energy dissipation inequality (EDI). Furthermore, analogous to the optimal transportation setting, we show that this interpretation can be used in a minimizing movement scheme to construct solutions to a regularized Landau equation.

1. Introduction

The Landau equation is an important partial differential equation in kinetic theory. It gives a description of colliding particles in plasma physics [Lifshitz and Pitaevskiǐ 1981], and it can be formally derived as a limit of the Boltzmann equation where grazing collisions are dominant [Degond and Lucquin-Desreux 1992; Villani 1998a]. Similar to the Boltzmann equation (see [Boblylev et al. 2013] for a consistency result and related derivation issues), the rigorous derivation of the Landau equation from particle dynamics is still a huge challenge. For a spatially homogeneous density of particles $f = f_t(v)$ for $t \in (0, \infty)$, $v \in \mathbb{R}^d$, the homogeneous Landau equation reads

$$\partial_t f(v) = \nabla_v \cdot \left(f(v) \int_{\mathbb{R}^d} |v - v_*|^{2+\gamma} \Pi[v - v_*](\nabla_v \log f(v) - \nabla_{v_*} \log f(v_*)) f(v_*) \, dv_* \right). \quad (1)$$

For notational convenience, we sometimes abbreviate $f = f_t(v)$ and $f_* = f_t(v_*)$. We also denote the differentiations by $\nabla = \nabla_v$ and $\nabla_* = \nabla_{v_*}$. The physically relevant parameters are usually $d = 2, 3$ and $\gamma \geq -d - 1$ with $\Pi[z] = I - (z \otimes z)/|z|^2$ being the projection matrix onto $\{z\}^\perp$. In this paper, for simplicity we will focus in the case $d = 3$ and vary the weight parameter γ, although most of our results are valid in arbitrary dimension. The regime $0 < \gamma < 1$ corresponds to the so-called *hard potentials*, while $\gamma < 0$ corresponds to the *soft potentials* with a further classification of $-2 \leq \gamma < 0$ as the moderately soft potentials and $-4 \leq \gamma < -2$ as the very soft potentials. The particular instances of $\gamma = 0$ and $\gamma = -d$ are known as the Maxwellian and Coulomb cases respectively.

The purpose of this work is to propose a new perspective inspired from gradient flows for weak solutions to (1), which is in analogy with the relationship of the heat equation and the 2-Wasserstein metric; see [Jordan et al. 1998; Ambrosio et al. 2008]. Our main result is inspired by and extends [Erbar 2023]. There, he establishes the gradient flow perspective for the closely related spatially homogeneous Boltzmann

equation with bounded collision kernels ($\gamma = 0$) which we perform in the case of Landau for $\gamma \in (-3, 0]$ (see Theorem 12). One of the fundamental steps is to symmetrize the right-hand side of (1). More specifically, if we consider a test function $\phi \in C_c^\infty(\mathbb{R}^d)$, we can formally characterize the equation by

$$\frac{d}{dt} \int_{\mathbb{R}^d} \phi f \, dv = -\frac{1}{2} \iint_{\mathbb{R}^{2d}} ff_* |v - v_*|^{2+\gamma} (\nabla\phi - \nabla_*\phi_*) \cdot \Pi[v - v_*](\nabla \log f - \nabla_* \log f_*) \, dv_* \, dv, \quad (2)$$

where the change of variables $v \leftrightarrow v_*$ has been exploited. Building our analogy with the heat equation and the 2-Wasserstein distance, we define an appropriate gradient

$$\tilde{\nabla}\phi := |v - v_*|^{1+\gamma/2} \Pi[v - v_*](\nabla\phi - \nabla_*\phi_*),$$

so that (2) now looks like

$$\frac{d}{dt} \int_{\mathbb{R}^d} \phi f \, dv = -\frac{1}{2} \iint_{\mathbb{R}^{2d}} ff_* \tilde{\nabla}\phi \cdot \tilde{\nabla} \log f \, dv_* \, dv,$$

noting that $\Pi^2 = \Pi$. To highlight the use of this interpretation, we notice that $\tilde{\nabla}\phi = 0$, when we choose as test functions $\phi = 1$, v_i, $|v|^2$ for $i = 1, \ldots, d$, which immediately shows that formally the equation conserves mass, momentum and energy. The action functional defining the Landau metric mimics the Benamou–Brenier formula [2000] for the 2-Wasserstein distance; see [Dolbeault et al. 2009; Erbar 2014; Erbar and Maas 2014] for other distances defined analogously for nonlinear and nonlocal mobilities. In fact, the Landau metric is built by considering a minimizing action principle over curves that are solutions to the appropriate continuity equation, that is,

$$d_L(f, g) := \min_{\substack{\partial_t \mu + \tilde{\nabla} \cdot (V\mu\mu_*)/2 = 0 \\ \mu_0 = f, \, \mu_1 = g}} \left\{ \frac{1}{2} \int_0^1 \iint_{\mathbb{R}^{2d}} |V|^2 \, d\mu(v) \, d\mu(v_*) \, dt \right\}, \quad (3)$$

where the $\tilde{\nabla} \cdot$ is the appropriate divergence; the formal adjoint to the appropriate gradient (see Section 2.1).

Also, we notice that analogously to the heat equation, written as the continuity equation $\partial_t f = \nabla \cdot (f \nabla \log f)$, the Landau equation can be formally rewritten as

$$\partial_t f = \tfrac{1}{2} \tilde{\nabla} \cdot (ff_* \tilde{\nabla} \log f),$$

equivalent to the continuity equation with nonlocal velocity field given by

$$\begin{cases} \partial_t f + \nabla \cdot (U(f)f) = 0, \\ U(f) := -\int_{\mathbb{R}^d} |v - v_*|^{2+\gamma} \Pi[v - v_*](\nabla \log f - \nabla_* \log f_*) f_* \, dv_*. \end{cases} \quad (4)$$

This is a direct way to write (1) in the form of a continuity equation. Considering the evolution of Boltzmann entropy we formally obtain

$$\frac{d}{dt} \int_{\mathbb{R}^d} f \log f \, dv =: -D(f_t) = -\frac{1}{2} \iint_{\mathbb{R}^{2d}} |\tilde{\nabla} \log f|^2 ff_* \, dv_* \, dv \leq 0. \quad (5)$$

In physical terms this is referred to as the *entropy dissipation* (referred to as entropy production in the physics literature from defining $\mathcal{H}$ with a minus sign) since it formally shows that the entropy functional

$$\mathcal{H}[f] := \int_{\mathbb{R}^d} f \log f \, dv$$

is nonincreasing along the dynamics of the Landau equation. Moreover, by integrating (5) in time one formally obtains

$$\mathcal{H}[f_t] + \int_0^t D(f_s)\, ds = \mathcal{H}[f_0]. \tag{6}$$

Villani [1998a] introduced the notion of H-solution, which captures this formal property. Motivated by the physical considerations of certain conserved quantities and entropy dissipation, H-solutions provided a step towards well-posedness of the Landau equation in the soft potential case. One advantage to this approach is that it avoids assuming that the solutions belongs to $L^p(\mathbb{R}^3)$ for $p > 1$. For moderately soft potentials, the propagation of L^p norms is proven and this is enough to make sense of classical weak solutions [Wu 2014]. In the very soft potential case, there is no longer a guarantee of L^p propagation due to the singularity of the weight. We refer to [Desvillettes 2015, Section 1.2] for a heuristic description of this difficulty.

Similar to H-solutions our approach will also be based on the entropy dissipation (6). Following De Giorgi's minimizing movement ideas [Ambrosio 1995; Ambrosio et al. 2008], we characterize the Landau equation by its associated energy dissipation inequality. More specifically, we show that weak solutions to (1) with initial data f_0 are completely determined by the functional inequality

$$\mathcal{H}[f_t] + \frac{1}{2} \int_0^t |\dot{f}|_{d_L}^2(s)\, ds + \frac{1}{2} \int_0^t D(f_s)\, ds \le \mathcal{H}[f_0] \quad \text{for a.e. every } t > 0,$$

where $|\dot{f}|_{d_L}^2(s)$ stands for the metric derivative associated to the Landau metric defined above. Our analysis is also largely inspired by Erbar's approach [2023] in viewing the Boltzmann equation as a gradient flow and recent numerical simulations of the homogeneous Landau equation in [Carrillo et al. 2020] based on a regularized version of (4). In contrast with the classical 2-Wasserstein metric, one of the main features of the Landau equation (1) and metric (3) is that they are nonlocal. To be precise, gradient flow theory has been successfully applied to the study of many nonlocal PDEs [Carrillo et al. 2010; 2012; Blanchet et al. 2008] by viewing them as gradient flows of appropriate energy functionals with respect to the 2-Wasserstein metric. The novelty in this work is the construction of the *nonlocal* metric d_L with respect to which (1) can be viewed as the gradient flow of $\mathcal{H}$. Hence, the convergence analysis usually relying on convexity and lower-semicontinuity needs to be adapted to deal with the nonlocality of this equation. In particular, our characterization Theorem 12 is based in using (expected) a priori estimates to deal with the nonlocality through appropriate bounds.

On the other hand, the state of the art related to the uniqueness for the Landau equation depends on the range of values γ may take. In the cases of hard potentials or Maxwellian, the uniqueness theory is very well understood due to Villani and the third author [Desvillettes and Villani 2000a; 2000b; Villani 1998b]. In the soft potential case, one of the first major contributions to the general theory of the spatially inhomogeneous Landau equation ($\gamma \ge -3$) was the global existence and uniqueness result of [Guo 2002]. This result was achieved in a perturbative framework with high regularity assumptions on the initial data. Through probabilistic arguments, the next major improvement to uniqueness for $\gamma \in (-3, 0)$ came from [Fournier and Guérin 2009]. Their result established uniqueness in a class of solutions that shrinks as γ decreases towards -3, as more L^p and moments assumptions are needed. In their proof, uniqueness is shown by proving stability with respect to the 2-Wasserstein metric.

Still lots of open questions for the soft potential case remain. In particular, a fundamental question like uniqueness for the Coulomb case is unresolved. To tackle this and other problems an array of novel methods have been employed. Here is an incomplete sample of the contributions made in this direction which highlight the difficulties of the soft potential case [Desvillettes and Villani 2000a; 2000b; Alexandre et al. 2015; Carrapatoso and Mischler 2017; Carrapatoso et al. 2017; Wu 2014; Gualdani and Zamponi 2017, 2018a; 2018b; Gualdani and Guillen 2016; Strain and Wang 2020; Golse et al. 2019a; 2019b; Silvestre 2017]. A brief glance at some of these references illustrates the breadth of techniques that have found partial success at answering the open questions: probability-based arguments, kinetic and parabolic theory, and many more.

The purpose of this paper is to bring in another set of techniques to help answer some of these fundamental questions. The gradient flow theory applied to PDEs has flourished in the last decades. In their seminal paper, Jordan, Kinderlehrer, and Otto [Jordan et al. 1998] proposed a variational approach (JKO scheme) extended later on to a wide class of PDEs using the optimal transportation distance of probability measures. These results and many more achievements from their contemporaries allowed for novel approaches to questions of existence, uniqueness, convergence to equilibrium, and other aspects of a large class of PDEs; we mention [Ambrosio et al. 2008; Santambrogio 2017] for a coherent exposition of these techniques and the relevant literature, even as more advances have been made since then.

The advantage of our variational characterization of the Landau equation is that it unveils new possible routes of showing convergence results for this equation. First of all, it allows for natural regularizations of the Landau equation by taking the steepest descent of regularized entropy functionals instead of the Boltzmann entropy as in [Carrillo et al. 2019]. This idea was recently developed in [Carrillo et al. 2020] leading to structure-preserving particle schemes with good accuracy. We can also consider the framework of convergence of gradient flows based on Γ-convergence introduced in [Sandier and Serfaty 2004; Serfaty 2011] to attack the convergence of these numerical methods [Carrillo et al. 2020]. Moreover, this approach is flexible enough to also study the rigorous convergence of the grazing collision limit of the Boltzmann equation to the Landau equation. The grazing collision limit was recently revisited in the gradient flow framework by three of the authors [Carrillo et al. 2022]. There, ideas from Γ-convergence were used to pass from Erbar's gradient flow description [2023] for the Boltzmann equation to the present work's description of the Landau equation. Finally, deriving uniqueness from the variational structure is classically done through convexity properties of the entropy functional with respect to the geodesics of the Landau metric. This is another important avenue of research that our work opens. Moreover, gradient flows of convex entropies typically enjoy instantaneous smoothing [Ambrosio et al. 2008]; even if the entropy at $t = 0$ is infinite, for $t > 0$, the entropy becomes finite. In the case of Landau, we are not aware if this property holds for $\mathcal{H}$.

We mention briefly the connection between (1) and the Fokker–Planck equation. For $\gamma = 0$, one can formally compute the evolution of $\int v^i v^j f(v) \, dv$ through (1). This a priori information allows one to reduce (1) to a linear Fokker–Planck equation for $\gamma = 0$. The present work proposes the alternative viewpoint that the resultant Fokker–Planck equation can be viewed as the d_L-gradient flow of $\mathcal{H}$ for $\gamma = 0$. Since many variants of the linear Fokker–Planck equation have been well-studied, this case serves as a nice benchmark to test the gradient flow theory developed here.

The plan of this paper is as follows. Section 2 introduces the prerequisites and contains the statements of the main results. We first construct and analyze in Section 3 the Landau metric based on (3). For a regularized problem, Section 4 shows the equivalence between weak solutions and gradient flows, while Section 5 shows the existence of gradient flow solutions via a minimizing movement scheme. Finally, we show in Section 6 that a gradient flow solution is equivalent to H-solutions of the Landau equation (1) under some integrability assumptions. The Appendix is devoted to some technical lemmas needed in the proof of the main theorems regarding the chain rule identity behind the definition of weak solutions for the regularized Landau equation.

2. Preliminaries and the main results

We start by introducing the necessary notation and definitions together with a quick overview of gradient flow concepts to make our main results fully self-contained.

2.1. *Notation and definitions.* We define

$$a \lesssim_{\cdots} b \iff \text{there exists } C(\cdots) > 0 \text{ such that } a \leq C(\cdots)b.$$

We adopt the Japanese angle bracket notation for a smooth alternative to absolute value

$$\langle v \rangle^2 = 1 + |v|^2, \quad v \in \mathbb{R}^d.$$

For $\epsilon > 0$, we define our regularization kernel to be an exponential distribution:

$$G^\epsilon(v) = \epsilon^{-d} G\left(\frac{v}{\epsilon}\right), \quad G(v) = C_d \exp(-\langle v \rangle), \quad C_d = \left(\int_{\mathbb{R}^d} \exp(-\langle v \rangle)\, dv\right)^{-1}.$$

Our results work for some general tail behavior in the kernels given by

$$G^{s,\epsilon}(v) = \epsilon^{-d} G^s\left(\frac{v}{\epsilon}\right), \quad G^s(v) = C_{s,d} \exp(-\langle v \rangle^s), \quad C_{s,d} = \left(\int_{\mathbb{R}^d} \exp(-\langle v \rangle^s)\, dv\right)^{-1}$$

for $s > 0$; we point out some of the limitations and restrictions on $s > 0$ in the later estimates. We shall refer to $G^{2,\epsilon}$ as the Maxwellian regularization. We denote the space of probability measures over $\mathbb{R}^d$ by $\mathscr{P}(\mathbb{R}^d)$, endowed with the weak topology against bounded continuous functions. We will mostly be dealing with the Lebesgue measure on $\mathbb{R}^d$ as our reference measure, which we denote by $\mathcal{L}$. The subset $\mathscr{P}^a(\mathbb{R}^d) \subset \mathscr{P}(\mathbb{R}^d)$ denotes the set of absolutely continuous probability measures with respect to Lebesgue measure. For $p > 0$, we also define the probability measures with finite p-moments $\mathscr{P}_p(\mathbb{R}^d)$ by

$$\mathscr{P}_p(\mathbb{R}^d) := \left\{ \mu \in \mathscr{P}(\mathbb{R}^d) \,\middle|\, m_p(\mu) := \int_{\mathbb{R}^d} \langle v \rangle^p\, d\mu(v) < \infty \right\}.$$

Finally, for $E > 0$, we consider the subset $\mathscr{P}_{p,E}(\mathbb{R}^d) \subset \mathscr{P}_p(\mathbb{R}^d)$ of probability measures with p-moments uniformly bounded by E:

$$\mathscr{P}_{p,E}(\mathbb{R}^d) := \{ \mu \in \mathscr{P}_p(\mathbb{R}^d) \mid m_p(\mu) \leq E \}.$$

We denote by $\mathcal{M}$ the space of signed Radon measures on $\mathbb{R}^d \times \mathbb{R}^d$ with the standard weak* topology against the continuous and compactly supported functions of $\mathbb{R}^d \times \mathbb{R}^d$. The space $\mathcal{M}^d$ is the space of signed d-length Radon measures. For $T > 0$, we will add the time contribution of the measures by defining $\mathcal{M}_T$ to be the space of signed Radon measures on $\mathbb{R}^d \times \mathbb{R}^d \times [0, T]$ with the usual weak* topology. Similarly, $\mathcal{M}_T^d$ will be the space of signed d-length Radon measures on $\mathbb{R}^d \times \mathbb{R}^d \times [0, T]$.

For $\mu \in \mathscr{P}(\mathbb{R}^d)$, we define a family of regularized entropies $\mathcal{H}_\epsilon[\mu]$ by

$$\mathcal{H}_\epsilon[\mu] := \int_{\mathbb{R}^d} [\mu * G^\epsilon](v) \log[\mu * G^\epsilon](v) \, dv,$$

which we shall see is well-defined provided μ has a finite moment in Lemma 30. Formally, one can calculate the first variation of this functional in $\mathscr{P}_2$ as

$$\frac{\delta \mathcal{H}_\epsilon}{\delta \mu}(v) = G^\epsilon * \log[\mu * G^\epsilon](v).$$

This can be formally obtained by calculating Fréchet derivatives in the sense of identifying the limit

$$\int_{\mathbb{R}^d} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu}(v) \phi(v) \, dv = \lim_{t \downarrow 0} \frac{\mathcal{H}_\epsilon[\mu + t\phi] - \mathcal{H}_\epsilon[\mu]}{t}$$

for arbitrary $\phi \in C_c^\infty(\mathbb{R}^d)$ with zero mean $\int_{\mathbb{R}^d} \phi = 0$. To be precise, the first variation (in an L^2 setting) would actually be $\delta \mathcal{H}_\epsilon / \delta \mu = 1 + G^\epsilon * \log[\mu * G^\epsilon]$. We drop the constant term since our functional space is $\mathscr{P}$ and the first variation typically appears with derivatives applied to it. For a functional $\mathcal{F} : \mathscr{P}^a(\mathbb{R}^d) \to \mathbb{R}$ with first variation $\delta \mathcal{F} / \delta f$, we refer to the $\mathcal{F}$ Landau equation as

$$\partial_t f = \nabla \cdot \left(f \int_{\mathbb{R}^d} f_* |v - v_*|^{2+\gamma} \Pi[v - v_*] \left(\nabla \frac{\delta \mathcal{F}}{\delta f} - \nabla_* \frac{\delta \mathcal{F}_*}{\delta f_*} \right) dv_* \right). \tag{7}$$

To clarify the meaning of $\tilde{\nabla} \cdot$, for a given test function $\phi = \phi(v) \in \mathbb{R}^d$ and vector-valued test function $A = A(v, v_*) \in \mathbb{R}^d$, we have

$$\iint_{\mathbb{R}^{2d}} [\tilde{\nabla}\phi](v, v_*) \cdot A(v, v_*) \, dv_* \, dv = - \int_{\mathbb{R}^d} \phi(v)[\tilde{\nabla} \cdot A](v) \, dv.$$

In this way, the $\mathcal{F}$ Landau equation (7) can be concisely written as

$$\partial_t f = \frac{1}{2} \tilde{\nabla} \cdot \left(f f_* \tilde{\nabla} \frac{\delta \mathcal{F}}{\delta f} \right).$$

Note, by formally testing (7) with $\phi = \delta \mathcal{F} / \delta f$, one obtains an analogy of Boltzmann's H-theorem with the functional $\mathcal{F}$:

$$\frac{d}{dt} \mathcal{F}[f_t] = -D_\mathcal{F}(f_t) := -\frac{1}{2} \iint_{\mathbb{R}^{2d}} f f_* \left| \tilde{\nabla} \frac{\delta \mathcal{F}}{\delta f} \right|^2 dv \, dv_* \leq 0.$$

We will refer to $D_\mathcal{F}$ as the $\mathcal{F}$ dissipation. This notation induces our notion of weak solutions to the $\mathcal{F}$ Landau equation (7) closely following Villani's H-solutions [1998a].

Definition 1 (weak $\mathcal{F}$ solutions). For $T > 0$, we say that a curve $f \in C([0, T]; L^1(\mathbb{R}^d))$ is a weak solution to the $\mathcal{F}$ Landau equation (7) if the following hold:

(1) $f\mathcal{L}$ is a probability measure with uniformly bounded second moment so that

$$f_t \geq 0, \qquad \int_{\mathbb{R}^d} f_t(v)\, dv = 1 \quad \text{for all } t \in [0, T], \qquad \sup_{t \in [0,T]} \int_{\mathbb{R}^d} \langle v \rangle^2 f_t(v)\, dv < \infty.$$

(2) The functional $\mathcal{F}$ evaluated along the curve is bounded by its initial value:

$$\mathcal{F}[f_t] \leq \mathcal{F}[f_0] < +\infty \quad \text{for all } t \in [0, T].$$

(3) The $\mathcal{F}$ dissipation is time integrable:

$$\int_0^T D_{\mathcal{F}}(f_t)\, dt = \frac{1}{2} \int_0^T \iint_{\mathbb{R}^{2d}} f f_* \left| \tilde{\nabla} \frac{\delta \mathcal{F}}{\delta f} \right|^2 dv\, dv_*\, dt < \infty.$$

(4) For every test function $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d)$, equation (7) is satisfied in weak form:

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \phi f_t(v)\, dv\, dt = \frac{1}{2} \int_0^T \iint_{\mathbb{R}^{2d}} f f_* \tilde{\nabla} \phi \cdot \tilde{\nabla} \frac{\delta \mathcal{F}}{\delta f}\, dv\, dv_*\, dt.$$

For $\epsilon > 0$, we will refer to the weak $\mathcal{H}_\epsilon$ solutions as *ϵ-solutions* and, recalling $\mathcal{H}$ is the Boltzmann entropy, we will refer to weak $\mathcal{H}$ solutions as just *weak solutions* or *H-solutions*. We deliberately use the terminology of H-solutions since the time integrability of $D_{\mathcal{H}}(f_t)$, as for [Villani 1998a], is essential in our analysis.

2.2. *Quick review of gradient flow theory.* We recall the basic definitions of gradient flow theory that can be found in more generality in [Ambrosio et al. 2008, Chapter 1]. Throughout, (X, d) denotes a complete (pseudo-)metric space X with (pseudo-)metric d. Points $a < b \in \mathbb{R}$ will refer to endpoints of some interval. $F : X \to (-\infty, \infty]$ will denote a proper function.

Definition 2 (absolutely continuous curve). A function $\mu : t \in (a, b) \mapsto \mu_t \in X$ is said to be an *absolutely continuous curve* if there exists $m \in L^2(a, b)$ such that for every $s \leq t \in (a, b)$

$$d(\mu_t, \mu_s) \leq \int_s^t m(r)\, dr.$$

Among all possible functions m in Definition 2, one can make the following minimal selection.

Definition 3 (metric derivative). For an absolutely continuous curve $\mu : (a, b) \to X$, we define its *metric derivative* at every $t \in (a, b)$ by

$$|\dot{\mu}|(t) := \lim_{h \to 0} \frac{d(\mu_{t+h}, \mu_t)}{|h|}.$$

Further properties of the metric derivative can be found in [Ambrosio et al. 2008, Theorem 1.1.2].

Definition 4 (strong upper gradient). The function $g : X \to [0, \infty]$ is a *strong upper gradient* with respect to F if for every absolutely continuous curve $\mu : t \in (a, b) \mapsto \mu_t \in X$ we have that $g \circ \mu : (a, b) \to [0, \infty]$ is Borel and the following inequality holds:

$$|F[\mu_t] - F[\mu_s]| \leq \int_s^t g(\mu_r)|\dot{\mu}|(r)\, dr \quad \text{for all } a < s \leq t < b.$$

Using Young's inequality and moving everything to one side, the inequality in Definition 4 implies

$$F[\mu_t] - F[\mu_s] + \frac{1}{2} \int_s^t g(\mu_r)^2 \, dr + \frac{1}{2} \int_s^t |\dot{\mu}|^2(r) \, dr \geq 0 \quad \text{for all } a < s \leq t < b.$$

If the reverse inequality also holds, one obtains the stronger energy dissipation *equality*. This leads to our notion of gradient flows.

Definition 5 (curve of maximal slope). An absolutely continuous curve $\mu : (a, b) \to X$ is said to be a *curve of maximal slope* for F with respect to its strong upper gradient $g : X \to [0, \infty]$ if $F \circ \mu : (a, b) \to [0, \infty]$ is nonincreasing and the following inequality holds:

$$F[\mu_t] - F[\mu_s] + \frac{1}{2} \int_s^t g(\mu_r)^2 \, dr + \frac{1}{2} \int_s^t |\dot{\mu}|^2(r) \, dr \leq 0 \quad \text{for all } a < s \leq t < b.$$

F has the following natural candidates for upper gradient.

Definition 6 (slopes). We define the *local slope of F* by

$$|\partial F|(\mu) := \limsup_{\nu \to \mu} \frac{(F(\nu) - F(\mu))^+}{d(\nu, \mu)}.$$

The superscript "+" refers to the positive part. The *relaxed slope of F* is given by

$$|\partial^- F|(\mu) := \inf\{\liminf_{n \to \infty} |\partial F|(\mu_n) \mid \mu_n \to \mu, \ \sup_{n \in \mathbb{N}}(d(\mu_n, \mu), F(\mu_n)) < +\infty\}.$$

2.3. *Main results.* In order to understand the Landau equation as a gradient flow, we need to clarify what type of object the corresponding metric is.

Theorem 7 (distance on $\mathscr{P}_{2,E}(\mathbb{R}^d)$). *The (pseudo-)metric d_L on $\mathscr{P}_{2,E}(\mathbb{R}^d)$ satisfies*:

- *d_L-convergent sequences are weakly convergent.*
- *d_L-bounded sets are weakly compact.*
- *The map $(\mu_0, \mu_1) \mapsto d_L(\mu_0, \mu_1)$ is weakly lower semicontinuous.*
- *For any $\tau \in \mathscr{P}_2(\mathbb{R}^d)$ the subset $\mathscr{P}_\tau(\mathbb{R}^d) := \{\mu \in \mathscr{P}_{2,m_2(\tau)}(\mathbb{R}^d) \mid d_L(\mu, \tau) < \infty\}$ is a complete geodesic space.*

The content of this theorem is essentially that our new proposed distance actually provides a meaningful topological structure on $\mathscr{P}_{2,E}(\mathbb{R}^d)$. Furthermore, the connection to ϵ-solutions of Landau is established when considering the previous notions of slope and upper gradient with respect to d_L. General conditions which guarantee $d_L(\mu_0, \mu_1) < +\infty$ are presently unknown. In Lemma 15, we will see that a necessary condition is that μ_0 and μ_1 have the same mean velocity. Moreover, for $\gamma \in [-4, -2]$, Lemma 15 asserts that they should have the same second moment. In the construction of d_L detailed in Section 3, if $\mu = \mu(t)$ for $t \in [0, T]$ is an H-solution of Landau, then it is certainly true that $d_L(\mu(t), \mu(s)) < +\infty$ for all $0 \leq t, s \leq T$.

Theorem 8 (epsilon equivalence). *Fix any $\epsilon, E > 0$, $\gamma \in [-4, 0]$. Assume that a curve $\mu : [0, T] \to \mathscr{P}_{2,E}(\mathbb{R}^d)$ has a density $\mu_t = f_t \mathcal{L}$. Then μ is a curve of maximal slope for $\mathcal{H}_\epsilon$ with respect to its upper gradient $\sqrt{D_{\mathcal{H}_\epsilon}}$ if and only if its density f is an ϵ-solution to the Landau equation.*

From the numerical perspective, we can also construct ϵ-solutions using the JKO scheme (see Section 5) which is the following:

Theorem 9 (existence of curves of maximal slope). *For any* ϵ, $E > 0$, $\gamma \in [-4, 0]$, *and initial data* $\mu_0 \in \mathscr{P}_{2,E}(\mathbb{R}^d)$, *there exists a curve of maximal slope in* $\mathscr{P}_{2,E}(\mathbb{R}^d)$ *for* $\mathcal{H}_\epsilon$ *with respect to its upper gradient* $\sqrt{D_{\mathcal{H}_\epsilon}}$.

Remark 10. The curves constructed in Theorem 9 do not necessarily have a density with respect to Lebesgue measure; the regularization allows $\mathcal{H}_\epsilon[\mu] < +\infty$ without μ being absolutely continuous with respect to Lebesgue measure. Moreover, uniqueness of such curves is beyond the scope of the present work although it would be interesting to see what convexity properties are available for $\mathcal{H}_\epsilon$ with respect to d_L. This could also shed some insight into the available convexity of $\mathcal{H}$ with respect to d_L.

Remark 11. The choice of an exponential convolution kernel G^ϵ for the regularized entropy $\mathcal{H}_\epsilon$ is perhaps unnatural compared to the Maxwellian regularization $G^{2,\epsilon}$. We discuss in more detail the estimates that fail using $G^{2,\epsilon}$ in Remark 33 as it pertains to Theorem 8. With respect to Theorem 9, the general construction of *some* curve can be done even with the Maxwellian regularization. However, due to the same lack of estimates, this curve might not be a curve of maximal slope with respect to $\sqrt{D_{\mathcal{H}_\epsilon}}$. This is discussed in Remark 37.

Motivated by recent numerical experiments [Carrillo et al. 2020], Theorems 8 and 9 provide the theoretical basis to this ϵ-approximated Landau equation. In the limit $\epsilon \to 0$, more assumptions are required.

Theorem 12 (full equivalence). *We fix* $d = 3$ *and* $\gamma \in (-3, 0]$. *Suppose that, for some* $T > 0$, *a curve* $\mu : [0, T] \to \mathscr{P}(\mathbb{R}^3)$ *has a density* $\mu_t = f_t \mathcal{L}$ *that satisfies the following set of assumptions:*

(A1) (moments and L^p) *Assume that, for some* $0 < \eta \leq \gamma + 3$, *we have*

$$\langle v \rangle^{2-\gamma} f_t(v) \in L_t^\infty(0, T; L_v^1 \cap L_v^{(3-\eta)/(3+\gamma-\eta)}(\mathbb{R}^3)).$$

(A2) (finite entropy) *We assume that the initial entropy is finite*

$$\mathcal{H}[f_0] = \int_{\mathbb{R}^3} f_0 \log f_0 < +\infty.$$

(A3) (finite entropy-dissipation) *We assume that the entropy-dissipation of* f *is integrable in time:*

$$D(f_t) = D_{\mathcal{H}}(f_t) = \frac{1}{2} \iint_{\mathbb{R}^6} f f_* \left| \tilde{\nabla} \frac{\delta \mathcal{H}}{\delta f} \right|^2 dv \, dv_*$$

$$= \frac{1}{2} \iint_{\mathbb{R}^6} f f_* |v - v_*|^{\gamma+2} |\Pi[v - v_*](\nabla \log f - \nabla_* \log f_*)|^2 \, dv \, dv_* \in L_t^1(0, T).$$

Then μ *is a curve of maximal slope for* $\mathcal{H}$ *with respect to its upper gradient* $\sqrt{D}$ *if and only if its density* f *is a weak solution of the Landau equation.*

Remark 13. When $\gamma \in [-2, 0]$, it is known that for suitable initial data (lying in weighted L^p spaces for p large enough and for a sufficient power-like weight), weak solutions of Landau equation satisfying (A1)–(A3) are known to exist (and to be strong and unique under extra conditions). We refer to [Wu 2014], and Appendix B of [Desvillettes 2022] when $\gamma > -2$, for details.

When $\gamma \in (-3, -2)$, (A1) is not known to hold for global weak solutions with large initial data. Solutions satisfying (A1)–(A3) are nevertheless known to exist for initial data close to equilibrium (see [Guo 2002] in a much larger spatially inhomogeneous context), or in the Coulomb case $\gamma = -3$ (in that case $(3 - \eta)/(3 + \gamma - \eta)$ being replaced by ∞) for large initial data, but on specific intervals of times only [Desvillettes et al. 2023; Arsenev and Peskov 1977].

The focus on the Maxwellian and soft potential regime $\gamma \leq 0$ here is motivated by building a gradient flow framework to address the open questions for Landau. The hard potential case $\gamma \in (0, 1)$ has already been studied in detail in [Desvillettes and Villani 2000b; 2000a]. We believe that our results also carry to the hard potentials. In particular, the exponents in (A1) should be modified to

$$\langle v \rangle^{2+\gamma} f_t(v) \in L_t^\infty(0, T; L_v^1(\mathbb{R}^3)), \quad f_t(v) \in L_t^\infty(0, T; L_v^{(3/(3-\gamma))+}(\mathbb{R}^3)), \quad 0 < \gamma < 1.$$

We emphasize that these conditions are guaranteed since the required moments and L^p integrability are propagated from appropriate initial data when $\gamma > 0$ [Desvillettes and Villani 2000a; 2000b]. This condition appears in [Desvillettes 2016, Corollary 2.7]. It is the hard potential version of Theorem 41, which is crucial to the proof of Theorem 12. Much of our analysis remains the same; however, the space $\mathscr{P}_2$ should be changed to $\mathscr{P}_{2+\gamma}$ cohering with the moment condition above and trivializing Lemma 43, for example.

It is an open problem to find the range of values γ under which we can show the existence of curves of maximal slope for the original Landau equation (1), or equivalently, constructing solutions of the original Landau equation passing $\epsilon \to 0$ in Theorem 9. Some of the difficulties to achieve this result are the propagation of moments for the regularized Landau equation uniformly in ϵ and the compactness of sequences with bounded in ϵ regularized entropy dissipation $D_{\mathcal{H}_\epsilon}$. The rest of this work is devoted to showing the main four theorems in the next four sections.

3. The Landau metric d_L

Our approach to defining the distance d_L mentioned in Theorem 7 closely follows the dynamic formulation of transport distances originally due to Benamou and Brenier [2000] and further extended by Dolbeault, Nazaret, and Savaré [Dolbeault et al. 2009]. We also refer the reader to [Erbar 2023] for a similar approach.

3.1. *Grazing continuity equation.* We consider for $\gamma \in [-4, 0]$ the *grazing continuity equation*

$$\partial_t \mu_t + \tfrac{1}{2} \tilde{\nabla} \cdot M_t = 0 \quad \text{in } (0, T) \times \mathbb{R}^d, \tag{8}$$

which is interpreted in the sense of distributions. For every $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d)$, we have

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \phi(t, v) \, d\mu_t(v) \, dt + \frac{1}{2} \int_0^T \iint_{\mathbb{R}^{2d}} [\tilde{\nabla}\phi](t, v, v_*) \, dM_t(v, v_*) \, dt = 0.$$

Another formulation (see Lemma 14) is the following for $\zeta \in C_c^\infty(\mathbb{R}^d)$:

$$\frac{d}{dt} \int_{\mathbb{R}^d} \zeta(v) \, d\mu_t(v) = \frac{1}{2} \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\zeta(v, v_*) \, dM_t(v, v_*). \tag{9}$$

The curves $(\mu_t)_{t\in[0,T]}$, $(M_t)_{t\in[0,T]}$ are Borel families of measures belonging to $\mathcal{M}_+$ and $\mathcal{M}^d$ respectively. We will refer to μ from the pair as a *curve* and M as a *grazing rate*. For some regularity properties, we will also need to assume the moment condition

$$\int_0^T \iint_{\mathbb{R}^{2d}} (1 + |v| + |v_*|) \, d|M_t|(v, v_*) \, dt < \infty. \tag{10}$$

We first establish some a priori properties of solutions to the grazing continuity equation.

Lemma 14 (continuous representative). *For families (μ_t), (M_t) satisfying the grazing continuity equation and the finite moment condition* (10), *there exists a unique weakly* continuous representative curve $(\tilde{\mu}_t)_{t\in[0,T]}$ such that $\tilde{\mu}_t = \mu_t$ for a.e. $t \in [0, T]$. Furthermore, for any $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d)$ and any $t_0, t_1 \in [0, T]$, we have the formula*

$$\int_{\mathbb{R}^d} \phi_{t_1} \, d\tilde{\mu}_{t_1} - \int_{\mathbb{R}^d} \phi_{t_0} \, d\tilde{\mu}_{t_0} = \int_{t_0}^{t_1} \int_{\mathbb{R}^d} \partial_t \phi \, d\mu_t \, dt + \frac{1}{2} \int_{t_0}^{t_1} \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\phi \, dM_t \, dt.$$

Proof. This proof is nearly identical to [Ambrosio et al. 2008, Lemma 8.1.2]. There, it was crucial to estimate the distributional time derivative of $t \mapsto \mu_t$. We perform the analogous estimate here to highlight the difference in our context. Fix $\zeta \in C_c^\infty(\mathbb{R}^d)$ and consider the map

$$t \in (0, T) \mapsto \mu_t(\zeta) = \int_{\mathbb{R}^d} \zeta(v) \, d\mu_t(v) \in \mathbb{R}.$$

According to (9), the distributional time derivative is

$$\dot{\mu}_t(\zeta) = \frac{1}{2} \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\zeta \, dM_t(v, v_*) = \frac{1}{2} \iint_{\mathbb{R}^{2d}} |v - v_*|^{1+\gamma/2}\Pi[v - v_*](\nabla\zeta - \nabla_*\zeta_*) \, dM_t(v, v_*).$$

Depending on the values of γ above or below -2, the integrand can be estimated:

$$\left| |v - v_*|^{1+\gamma/2}\Pi[v-v_*](\nabla\zeta - \nabla_*\zeta_*) \right| \le \begin{cases} 2^{1+\gamma/2}\sup_{w\in\mathbb{R}^d} |\nabla\zeta(w)|(|v|^{1+\gamma/2} + |v_*|^{1+\gamma/2}), & \gamma \in [-2, 0], \\ \sup_{w\in\mathbb{R}^d} |D^2\zeta(w)| \, |v - v_*|^{2+\gamma/2}, & \gamma \in [-4, -2). \end{cases}$$

Consequently, using the moment condition (10), we have the following estimates depending on $\gamma \in [-4, 0]$:

$$|\dot{\mu}_t(\zeta)| \lesssim \begin{cases} \sup_{w\in\mathbb{R}^d} |\nabla\zeta(w)| \iint_{\mathbb{R}^{2d}}(1 + |v| + |v_*|)d|M_t|(v, v_*), & \gamma \in [-2, 0], \\ \sup_{w\in\mathbb{R}^d} |D^2\zeta(w)| \iint_{\mathbb{R}^{2d}}(1 + |v| + |v_*|)d|M_t|(v, v_*), & \gamma \in [-4, -2). \end{cases}$$

The rest of the proof proceeds as in [Ambrosio et al. 2008, Lemma 8.1.2] using the C^2-norm of ζ for the soft potentials $\gamma \in [-4, -2)$ as opposed to their C^1 control of ζ. □

Lemma 15 (conservation lemma). *Fix $\gamma \in [-4, 0]$ and let $(\mu_t)_{t\in[0,T]}$, $(M_t)_{t\in[0,T]}$ be Borel families of measures in $\mathcal{M}_+$, $\mathcal{M}^d$ respectively satisfying* (8) *and the moment condition* (10). *Assume further that*

$(\mu_t)_{t \in [0,T]}$ *is weakly* continuous with respect to t. We have that mass and momentum are conserved*:

$$\mu_t(\mathbb{R}^d) = \mu_0(\mathbb{R}^d), \quad \int_{\mathbb{R}^d} v \, d\mu_t(v) = \int_{\mathbb{R}^d} v \, d\mu_0(v) \quad \text{for all } t \in [0,T].$$

In the case $\gamma \in [-4, -2]$ we have that the energy is conserved:

$$\int_{\mathbb{R}^d} |v|^2 \, d\mu_t(v) = \int_{\mathbb{R}^d} |v|^2 \, d\mu_0(v) \quad \text{for all } t \in [0,T].$$

Proof. To minimize clutter, we introduce $w = |v - v_*|^{1+\gamma/2}$. We show the proof of the conservation of energy for $\gamma \in [-4, -2]$. We consider a fixed $\varphi \in C_c^\infty(B_2)$ which satisfies

$$0 \leq \varphi \leq 1 \quad \text{and} \quad \varphi(v) = 1 \quad \text{in } B_1.$$

We define

$$\varphi_R(v) = \varphi\left(\frac{v}{R}\right).$$

Using the grazing continuity equation, we have

$$\int_{\mathbb{R}^d} |v|^2 \varphi_R(v) \, d\mu_t(v) - \int_{\mathbb{R}^d} |v|^2 \varphi_R(v) \, d\mu_0(v)$$
$$= \int_0^t \iint_{\mathbb{R}^{2d}} w \Pi \left(v\varphi_R(v) + |v|^2 \frac{\nabla\varphi(v/R)}{R} - v_*\varphi_R(v_*) - |v_*|^2 \frac{\nabla\varphi(v_*/R)}{R} \right) dM_s(v, v_*) \, ds. \quad (11)$$

We estimate the contribution of $v\phi_R(v) - v_*\phi_R(v_*)$ from the integral in (11) using the cancellation from the projection $\Pi[v - v_*]$ to obtain

$$\left| \int_0^t \iint_{\mathbb{R}^{2d}} w\Pi(v\varphi_R(v) - v_*\varphi_R(v_*)) \, dM_s \right| \leq \int_0^t \iint_{(B_R \times B_R)^c} w|v\varphi_R(v) - v_*\varphi_R(v_*)| \, d|M_s|$$
$$\lesssim \int_0^t \iint_{(B_R \times B_R)^c} (1 + |v| + |v_*|) \, d|M_s|,$$

where we have used $\gamma \in [-4, -2]$ to bound

$$w|v\varphi_R(v) - v_*\varphi_R(v_*)| \lesssim \begin{cases} 1, & |v - v_*| \leq 1, \\ |v| + |v_*|, & |v - v_*| \geq 1. \end{cases}$$

Similarly, using that $\nabla\phi_R$ is supported in $B_{2R} \setminus B_R$ and that $|\partial_{v^i}\{|v|^2 \partial_{v^j}\varphi(v/R)/R\}| \lesssim 1$ for every index $i, j \in \{1, \ldots, d\}$, we obtain

$$\left| \int_0^t \iint_{\mathbb{R}^{2d}} w\Pi\left(|v|^2 \frac{\nabla\varphi(v/R)}{R} - |v_*|^2 \frac{\nabla\varphi(v_*/R)}{R} \right) dM_s \right| \lesssim \iint_{(B_R \times B_R)^c} 1 + |v| + |v_*| \, d|M_s|,$$

where we have controlled the difference with a mean-value-type estimate. From the previous bounds, we can use hypothesis (10) to take $R \to \infty$ in (11) and obtain the conservation of energy

$$\lim_{R \to \infty} \int_{\mathbb{R}^d} |v|^2 \varphi_R(v) \, d\mu_t(v) = \lim_{R \to \infty} \int_{\mathbb{R}^d} |v|^2 \varphi_R(v) \, d\mu_0(v).$$

The proofs for conservation of mass and momentum involve testing the grazing continuity equation against ϕ_R and $v_i\phi_R$ respectively, where v_i is the i-th component of v. For these statements, the case $\gamma \in [-4, -2]$ follows in the same way. For $\gamma \in [-2, 0]$, the estimates can be more blunt since the weight is no longer singular. $\qquad \square$

Remark 16. Note that as γ increases into the range $(-2, 0]$, the weight function w starts adding growth so the mean-value-type argument in Lemma 15 no longer helps unless more moments of M are assumed than (10). Due to the conservation of mass, the unique weakly* continuous representative $(\tilde{\mu}_t)$ of Lemma 14 has the additional property of being weakly continuous in the context of $\mathscr{P}(\mathbb{R}^d)$.

Based on the previous results, we propose the following definition.

Definition 17 (grazing continuity equation). For some terminal time $T > 0$, we define $\mathcal{GCE}_T$ to be the set of pairs of measures $(\mu_t, M_t)_{t \in [0,T]}$ satisfying the following:

(1) $\mu_t \in \mathscr{P}(\mathbb{R}^d)$ is weakly continuous with respect to $t \in [0, T]$. $(M_t)_{t \in [0,T]}$ is a family of Borel measures belonging to $\mathcal{M}^d$.

(2) We have the moment bound

$$\int_0^T \iint_{\mathbb{R}^{2d}} (1 + |v| + |v_*|) \, d|M_t|(v, v_*) \, dt < \infty.$$

(3) The grazing continuity equation (8) is satisfied in the distributional sense. That is, for every $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d)$,

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \phi \, d\mu_t \, dt + \frac{1}{2} \int_0^T \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \phi \, dM_t \, dt = 0,$$

or equivalently, for every $\zeta \in C_c^\infty(\mathbb{R}^d)$,

$$\frac{d}{dt} \int_{\mathbb{R}^d} \zeta(v) \, d\mu_t(v) = \frac{1}{2} \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \zeta(v, v_*) \, dM_t(v, v_*).$$

For fixed probability measures λ, ν, we may also specify the subset $\mathcal{GCE}(\lambda, \nu)$ as those pairs $(\mu, M) \in \mathcal{GCE}_T$ such that $\mu_0 = \lambda$, $\mu_T = \nu$. For $E > 0$, we will speak of curves $(\mu, M) \in \mathcal{GCE}_T^{2,E}$ such that

$$\int_{\mathbb{R}^d} |v|^2 \, d\mu_t(v) \leq E \quad \text{for all } t \in [0, T].$$

3.2. *Action of a curve.* In this section, we construct the action of a curve under the grazing continuity equation. We introduce the function $\alpha : \mathbb{R}^d \times \mathbb{R}_{\geq 0} \to [0, \infty]$ defined by

$$\alpha(u, s) := \begin{cases} |u|^2/(2s), & s \neq 0, \\ 0, & s = 0, u = 0, \\ \infty, & s = 0, u \neq 0. \end{cases}$$

Remark 18. The function α is lower semicontinuous (lsc), convex, and positively 1-homogeneous.

For fixed $\mu \in \mathscr{P}(\mathbb{R}^d)$, $M \in \mathcal{M}^d$, we consider the tensorized probability measure $\mu \otimes \mu \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d)$ given by $\mu \otimes \mu(dv, dv_*) = \mu(dv)\mu(dv_*)$. Define $\tau \in \mathcal{M}$ given by $\tau = \mu \otimes \mu + |M|$ and the decompositions $\mu \otimes \mu = f^1 \tau$ and $M = N\tau$. We define the action functional as

$$\mathcal{A}(\mu, M) := \iint_{\mathbb{R}^{2d}} \alpha(N, f^1) d\tau. \tag{12}$$

This is well-defined by the 1-homogeneity of α. The following lemma establishes a more concrete expression for the action functional.

Lemma 19. *Let $\mu \in \mathscr{P}(\mathbb{R}^d)$ be absolutely continuous with respect to $\mathcal{L}$ and $\mu = f\mathcal{L}$. Let $M \in \mathcal{M}^d$ be given such that $\mathcal{A}(\mu, M) < \infty$. Then, M is absolutely continuous with respect to $ff_* dv\, dv_*$ given by some density $U : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ such that $M = ff_*U\, dv\, dv_* = m\, dv\, dv_*$ and*

$$\mathcal{A}(\mu, M) = \frac{1}{2} \iint_{\mathbb{R}^{2d}} ff_*|U|^2\, dv\, dv_* = \frac{1}{2} \iint_{\mathbb{R}^{2d}} \frac{|m|^2}{ff_*}\, dv\, dv_*.$$

Proof. The proof is identical to [Erbar 2023, Lemma 3.6] up to appropriate modifications. Define $\tau \in \mathcal{M}$ by $\tau = \mu \otimes \mu + |M|$ and label the corresponding densities (which may be infinite) $\mu \otimes \mu = g\tau$ and $M = N\tau$. It suffices to show that M is absolutely continuous with respect to $\mu \otimes \mu$, which is the goal of this proof.

Suppose $S \subset \mathbb{R}^{2d}$ is a measurable set such that $\mu \otimes \mu(S) = 0$. This is equivalent to saying $g = 0$ τ-almost everywhere in S. Since α is positive, the assumption $\mathcal{A}(\mu, M) < +\infty$ certainly implies $\alpha(N, g) < +\infty$ τ-almost everywhere in S. By the definition of α, we must also have $N = 0$ τ-almost everywhere in S, which is equivalent to saying $M(S) = 0$. $\qquad\square$

Lemma 20 (lower semicontinuity of action functional). *The action functional $\mathcal{A}$ as defined in (12) is lower semicontinuous in both arguments. Specifically, if $\mu_n \rightharpoonup \mu$ weakly in $\mathscr{P}(\mathbb{R}^d)$ and $M_n \overset{*}{\rightharpoonup} M$ weakly* in $\mathcal{M}^d$, we have*

$$\mathcal{A}(\mu, M) \leq \liminf_{n\to\infty} \mathcal{A}(\mu_n, M_n).$$

Proof. This result is an application of the general lsc result in [Buttazzo 1989, Theorem 3.4.3] since α satisfies the required convexity, lsc, and homogeneity assumptions by Remark 18. $\qquad\square$

Another useful property of the action functional is the compactness provided by bounded action. We first state:

Lemma 21. *Let $F : \mathbb{R}^{2d} \to [0, \infty]$ be measurable and fix any $\mu \in \mathscr{P}(\mathbb{R}^d)$, $M \in \mathcal{M}^d$. We have the following bound*:

$$\iint_{\mathbb{R}^{2d}} F(v, v_*)d|M|(v, v_*) \leq \sqrt{2}\mathcal{A}(\mu, M)^{1/2}\left(\iint_{\mathbb{R}^{2d}} F(v, v_*)^2 d\mu(v)d\mu(v_*)\right)^{1/2}. \qquad (13)$$

Proof. This proof follows [Erbar 2023, Lemma 3.8]. We assume $\mathcal{A}(\mu, M) < +\infty$ or else (13) holds automatically. This implies that whenever $A \subset \mathbb{R}^{2d}$ is a measurable set, $\mu \otimes \mu(A) = 0$ if and only if $|M|(A) = 0$. Therefore, in the following computations we are implicitly integrating away from sets of zero $\mu \otimes \mu$-measure. We provide the simple argument by Cauchy–Schwarz for completeness. By considering $\tau = \mu \otimes \mu + |M|$, we estimate

$$\iint_{\mathbb{R}^{2d}} F d|M|(v, v_*) \leq \iint_{\mathbb{R}^{2d}} F\left|\frac{dM}{d\tau}\right|d\tau(v, v_*) = \iint_{\mathbb{R}^{2d}} F\left(\left|\frac{dM}{d\tau}\right| \middle/ \sqrt{2\frac{d\mu \otimes \mu}{d\tau}}\right)\sqrt{2\frac{d\mu \otimes \mu}{d\tau}}d\tau$$

$$\leq \left(\iint_{\mathbb{R}^{2d}} \alpha\left(\frac{dM}{d\tau}, \frac{d\mu \otimes \mu}{d\tau}\right)d\tau\right)^{1/2}\left(\iint_{\mathbb{R}^{2d}} 2F^2 d\mu \otimes \mu\right)^{1/2}$$

$$= \sqrt{2}\mathcal{A}(\mu, M)^{1/2}\left(\iint_{\mathbb{R}^{2d}} F(v, v_*)^2 d\mu(v)d\mu(v_*)\right)^{1/2}. \qquad\square$$

Remark 22. Suppose we have $\mu_t \in \mathscr{P}(\mathbb{R}^d)$ such that

$$\int_0^T m_2(\mu_t)\, dt = \int_0^T \int_{\mathbb{R}^d} \langle v \rangle^2 \, d\mu_t(v)\, dt < \infty.$$

Then for $M \in \mathcal{M}_T^d$ the previous estimate (13) yields

$$\int_0^T \iint_{\mathbb{R}^{2d}} (1 + |v| + |v_*|)\, d|M_t|(v, v_*)\, dt \lesssim \int_0^T \mathcal{A}(\mu_t, M_t)^{1/2} \Big(1 + 2 \int_{\mathbb{R}^d} |v|^2 \, d\mu_t\Big)^{1/2} dt. \quad (14)$$

Therefore, if the integral in time of the second moment of μ is bounded, then M satisfies the moments conditions (10) and the energy is conserved (Lemma 15). In the sequel, we will be considering curves that have bounded second moment which guarantee (14).

Proposition 23. *Let $(\mu_t^n, M_t^n)_n$ be a sequence in $\mathcal{GCE}_T$ such that $(\mu_0^n)_n$ is tight and we have the uniform bounds*

$$\sup_{n \in \mathbb{N}} \int_0^T \int_{\mathbb{R}^d} |v|^2 \, d\mu_t^n \, dt < \infty \quad and \quad \sup_{n \in \mathbb{N}} \int_0^T \mathcal{A}(\mu_t^n, M_t^n)\, dt < \infty. \quad (15)$$

Then, there exists $(\mu_t, M_t) \in \mathcal{GCE}_T$ such that, possibly after extracting a subsequence, we have the convergences

$$\mu_t^n \rightharpoonup \mu_t \qquad \text{weakly in } \mathscr{P}(\mathbb{R}^d) \text{ for all } t \in [0, T],$$
$$M_t^n \, dt \overset{*}{\rightharpoonup} M_t dt \quad \text{weakly* in } \mathcal{M}_T^d.$$

Furthermore, along this subsequence we have the lower semicontinuity

$$\int_0^T \mathcal{A}(\mu_t, M_t)\, dt \leq \liminf_{n \to \infty} \int_0^T \mathcal{A}(\mu_t^n, M_t^n)\, dt.$$

Sketch of the proof. This result follows from a proof similar to that of [Dolbeault et al. 2009, Lemma 4.5] and [Erbar 2023, Proposition 3.11], which we sketch. The second moment bound for μ^n in (15) produces a limit μ. Recalling the application of Lemma 21 in Remark 22, the bounded action in (15) and the estimate (14) produce a limit $M_t \, dt$ for a subsequence of $M_t^n \, dt$. The lower semicontinuity follows from Fatou's lemma and Lemma 20. $\qquad \square$

3.3. *Properties of the Landau metric.* We define the distance, d_L induced by the action functional on $\mathscr{P}_{2,E}(\mathbb{R}^d)$. Throughout, we will be working in the grazing continuity equation space defined earlier by $\mathcal{GCE}_T^{2,E}$ for $T > 0$ some terminal time and $E > 0$ any second moment bound.

Definition 24. For $\lambda, \nu \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ we define the (square of the) Landau distance by

$$d_L^2(\lambda, \nu) := \inf \Big\{ T \int_0^T \mathcal{A}(\mu_t, M_t)\, dt \ \Big| \ (\mu, M) \in \mathcal{GCE}_T^{2,E}(\lambda, \nu) \Big\}. \quad (16)$$

Notice this definition is independent of $T > 0$ considering the scaling of the grazing collision equation and the 1-homogeneity of $\mathcal{A}$. We have an equivalent characterization of d_L which can be seen in other PDE contexts such as [Erbar 2023; Dolbeault et al. 2009].

Lemma 25. *Given* $\lambda, \nu \in \mathscr{P}_{2,E}(\mathbb{R}^d)$, *we have*

$$d_L(\lambda, \nu) = \inf\left\{ \int_0^T \sqrt{\mathcal{A}(\mu_t, M_t)} \, dt \;\middle|\; (\mu, M) \in \mathcal{GCE}_T^{2,E}(\lambda, \nu) \right\}. \tag{17}$$

Proof. This proof uses the same reparametrization technique in [Dolbeault et al. 2009, Theorem 5.4]. □

Proposition 26 (minimizing curve). *Suppose that* $\mu_0, \mu_1 \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ *are probability measures such that* $d_L(\mu_0, \mu_1) < \infty$. *Then there exists a curve* $(\mu, M) \in \mathcal{GCE}_1^{2,E}(\mu_0, \mu_1)$ *attaining the infimum of* (16) (*equivalently, also* (17)) *and* $\mathcal{A}(\mu_t, M_t) = d_L^2(\mu_0, \mu_1)$ *for almost every* $t \in [0, 1]$.

Proof. This result follows from the direct method of calculus of variations where the lower semicontinuity comes from Proposition 23. □

Proof of Theorem 7. We prove the statements in exactly the order they are presented in the theorem, starting with the properties of the proposed Landau distance as a metric. The positivity of d_L follows from the positivity of α. We now check that d_L satisfies the properties of a metric.

<u>d_L distinguishes points</u>: Fix $\mu_0, \mu_1 \in \mathscr{P}_{2,E}(\mathbb{R}^d)$. We check that $d_L(\mu_0, \mu_1) = 0 \iff \mu_0 = \mu_1$. Suppose that $d_L(\mu_0, \mu_1) = 0$. By Proposition 26 we can find $(\mu, M) \in \mathcal{GCE}_1^{2,E}(\mu_0, \mu_1)$, which is a minimizing curve and moreover $0 = d_L(\mu_0, \mu_1) = \mathcal{A}(\mu_t, M_t)$ implies $M = 0$. The grazing continuity equation reduces to $\partial_t \mu_t = 0$, which implies μ_t is constant in time.

The converse statement follows similarly by pairing the constant curve $\mu : t \mapsto \mu_0 = \mu_1$ with the zero measure so that $(\mu, 0) \in \mathcal{GCE}_1^{2,E}(\mu_0, \mu_1)$.

<u>Symmetry</u>: Symmetry follows because time can be reversed for every curve. For instance, if $(\mu, M) \in \mathcal{GCE}_T^{2,E}(\mu_0, \mu_1)$, then one can check that the pair

$$\mu^r : t \mapsto \mu(T - t), \quad M^r : t \mapsto -M(T - t)$$

belongs to $\mathcal{GCE}_T^{2,E}(\mu_1, \mu_0)$ with the same action.

<u>Triangle inequality</u>: We sketch the argument using a gluing lemma as in [Dolbeault et al. 2009, Lemma 4.4]. Let $\mu^0, \mu^1, \mu^2 \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ be such that $d_L(\mu^0, \mu^1) < \infty$ and $d_L(\mu^1, \mu^2) < \infty$. If not, $d_L(\mu^0, \mu^2) \le d_L(\mu^0, \mu^1) + d_L(\mu^1, \mu^2)$ holds trivially. By Proposition 26, we can find minimizing curves connecting these probability measures

$$\left\{ \begin{array}{ll} (\mu^{0 \to 1}, M^{0 \to 1}) & \in \mathcal{GCE}_1^{2,E}(\mu^0, \mu^1), \\ (\mu^{1 \to 2}, M^{1 \to 2}) & \in \mathcal{GCE}_1^{2,E}(\mu^1, \mu^2) \end{array} \right\}.$$

Their concatenation from time 0 to 1 is given by

$$\mu_t := \begin{cases} \mu_{2t}^{0 \to 1}, & 0 \le t \le \frac{1}{2}, \\ \mu_{2(t-1/2)}^{1 \to 2}, & \frac{1}{2} \le t \le 1, \end{cases} \quad M_t := \begin{cases} 2M_{2t}^{0 \to 1}, & 0 \le t \le \frac{1}{2}, \\ 2M_{2t-1/2}^{1 \to 2}, & \frac{1}{2} < t \le 1. \end{cases}$$

One can check that $(\mu, M) \in \mathcal{GCE}_1^{2,E}(\mu^0, \mu^2)$, so it is an admissible competitor in the computation of $d_L(\mu^0, \mu^2)$. By looking at the action on the different time pieces, we obtain

$$d_L(\mu^0, \mu^2) \le \int_0^1 \sqrt{\mathcal{A}(\mu_t, M_t)} \, dt = d_L(\mu^0, \mu^1) + d_L(\mu^1, \mu^2).$$

d_L-convergence/boundedness implies weak convergence/compactness: Fix $\mu^n, \mu^\infty \in \mathscr{P}_{2,E}$ for $n \in \mathbb{N}$ such that $d_L(\mu^\infty, \mu^n) \to 0$ as $n \to \infty$. By Proposition 26, take minimizing curves $(v^n, M^n) \in \mathcal{GCE}_1^{2,E}(\mu^\infty, \mu^n)$ such that

$$d_L(\mu^\infty, \mu^n) = \sqrt{\mathcal{A}(v_t^n, M_t^n)} \quad \text{for a.e. } t \in [0, 1].$$

By compactness in Proposition 23, there are limits $(v, M) \in \mathcal{GCE}_1^{2,E}$ such that $v^n \rightharpoonup v$ and $M^n \overset{*}{\rightharpoonup} M$ up to a subsequence. Moreover, the lower semicontinuity in Proposition 23 gives

$$\mathcal{A}(v_t, M_t) \leq \liminf_{n \to \infty} \mathcal{A}(v_t^n, M_t^n) = 0;$$

hence $M = 0$ so that v is a constant in time. Since $v(0) = \mu^\infty$, this implies $\mu^\infty = v(1) = \lim_{n \to \infty} \mu^n$, which establishes the weak convergence.

$(\mathscr{P}_\tau, d_L)$ is a complete geodesic space: We start with the geodesic property from completely analogous arguments to [Erbar 2023]; the remaining statement that $\mathscr{P}_\tau$ equipped with d_L is a complete geodesic space follows. Fix $\tau \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ with $\mu_0, \mu_1 \in \mathscr{P}_\tau$. The triangle inequality ensures $d_L(\mu_0, \mu_1) < \infty$ so Proposition 26 guarantees the existence of a minimizing curve $(\mu, M) \in \mathcal{GCE}_1^{2,E}(\mu_0, \mu_1)$. One easily sees that this also induces a minimizing curve for intermediate times. More precisely, for every $0 \leq r \leq s \leq 1$, we have that $(t \mapsto \mu_{t+r}, t \mapsto M_{t+r}) \in \mathcal{GCE}_{s-r}^{2,E}(\mu_r, \mu_s)$ also minimizes $d_L(\mu_r, \mu_s)$.

To show completeness, let $(\mu^n)_{n \in \mathbb{N}}$ be a Cauchy sequence in $\mathscr{P}_\tau$. The sequence is certainly d_L-bounded so by Proposition 23, we can find, up to extraction of a weakly convergent subsequence, $\mu^\infty \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ such that $\mu^n \rightharpoonup \mu^\infty$ in $\mathscr{P}_{2,E}(\mathbb{R}^d)$. Lower semicontinuity of d_L and the Cauchy property of the subsequence give

$$d_L(\mu^n, \mu^\infty) \leq \liminf_{m \to \infty} d_L(\mu^n, \mu^m) \to 0 \quad \text{as } n \to \infty.$$

For any $n \in \mathbb{N}$ the triangle inequality gives

$$d_L(\mu^\infty, \tau) \leq d_L(\mu^\infty, \mu^n) + d_L(\mu^n, \tau) < \infty,$$

So $\mu^\infty \in \mathscr{P}_\tau$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

Proposition 27 (metric derivative). *A curve $(\mu_t)_{t \in [0,T]} \subset \mathscr{P}_{2,E}(\mathbb{R}^d)$ is absolutely continuous with respect to d_L if and only if there exists a Borel family $(M_t)_{t \in [0.T]}$ belonging to $\mathcal{M}_T^d$ such that $(\mu, M) \in \mathcal{GCE}_T^{2,E}$ with the property that*

$$\int_0^T \sqrt{\mathcal{A}(\mu_t, M_t)} \, dt < \infty.$$

In this equivalence, we have a bound on the metric derivative

$$\lim_{h \downarrow 0} \frac{d_L^2(\mu_{t+h}, \mu_t)}{h^2} =: |\dot{\mu}|^2(t) \leq \mathcal{A}(\mu_t, M_t) \quad \text{for a.e. } t \in (0, T).$$

Furthermore, there exists a unique Borel family $(\widetilde{M}_t)_{t \in [0,T]}$ belonging to $\mathcal{M}^d$ which is characterized by

$$M_t = U \mu_t \otimes \mu_t \quad \text{and} \quad U \in T_\mu := \overline{\{\widetilde{\nabla}\phi \mid \phi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu_t \otimes \mu_t)}$$

such that $(\mu, \widetilde{M}) \in \mathcal{GCE}_T^E(\mu_0, \mu_T)$ where we have equality

$$|\dot{\mu}|^2(t) = \mathcal{A}(\mu_t, \widetilde{M}_t) \quad \text{for a.e. } t \in (0, T).$$

Proof. The argument follows exactly as in [Dolbeault et al. 2009, Theorem 5.17]. $\qquad\qquad\qquad\qquad\quad$ $\square$

4. Energy dissipation equality

The goal in this section is to prove Theorem 8, which states that the notions of gradient flow solutions coincide with ϵ-solutions to the Landau equation. To fix ideas, we recall the regularized entropy functionals acting on probability measures

$$\mathcal{H}_\epsilon[\mu] = \int_{\mathbb{R}^d} (\mu * G^\epsilon)(v) \log(\mu * G^\epsilon)(v)\, dv,$$

with $G^\epsilon(v)$ given by

$$G^\epsilon(v) = \epsilon^{-d} C_d \exp\left\{-\left\langle \frac{v}{\epsilon} \right\rangle\right\}.$$

The crucial ingredient to prove Theorem 8 is the following:

Proposition 28 (chain rule ϵ). *Fix* $\gamma \in [-4, 0]$ *and suppose* $(\mu, M) \in \mathcal{GCE}_T^{2,E}$ *and*

$$\int_0^T \mathcal{A}(\mu_t, M_t)\, dt < \infty.$$

Then, $\sup_{t \in [0,T]} \mathcal{H}_\epsilon[\mu_t] < \infty$ *and the "chain rule" holds*:

$$\mathcal{H}_\epsilon[\mu_r] - \mathcal{H}_\epsilon[\mu_s] = \frac{1}{2} \int_s^r \iint_{\mathbb{R}^{2d}} \tilde\nabla\left[\frac{\delta\mathcal{H}_\epsilon}{\delta\mu}\right] \cdot dM_t\, dt \quad \text{for all } 0 \le s \le r \le T. \tag{18}$$

Remark 29. Recall the expression for the dissipation

$$D_\epsilon[\mu] = \frac{1}{2} \iint_{\mathbb{R}^{2d}} \left| \tilde\nabla\left[\frac{\delta\mathcal{H}_\epsilon}{\delta\mu}\right] \right|^2 d\mu(v) d\mu(v_*).$$

Using a time integrated version of Lemma 21, we have the estimate

$$\frac{1}{2} \int_s^r \iint_{\mathbb{R}^{2d}} \left| \tilde\nabla\left[\frac{\delta\mathcal{H}_\epsilon}{\delta\mu}\right] \right| \cdot d|M_t|(v, v_*)\, dt \le \int_s^r \mathcal{A}(\mu_t, M_t)^{1/2} D_\epsilon[\mu_t]^{1/2}\, dt.$$

Therefore, under the hypothesis of Proposition 28, we have that

$$|\mathcal{H}_\epsilon(\mu_r) - \mathcal{H}_\epsilon(\mu_r)| \le \int_s^r |\dot\mu|(t) D_\epsilon[\mu_t]^{1/2}\, dt,$$

which implies that $D_\epsilon[\mu_t]^{1/2}$ is a strong upper gradient of $\mathcal{H}_\epsilon$; see Definition 4.

Taking Proposition 28 for granted, we can prove Theorem 8.

Proof of Theorem 8. Throughout, $\mu = f\mathcal{L}$ is a curve of probability measures with uniformly bounded second moment.

<u>Weak ϵ-solution $\Longrightarrow$ curve of maximal slope</u>: Consider f an ϵ-solution to the Landau equation. Define $m = -ff_* \tilde\nabla(\delta\mathcal{H}_\epsilon/\delta f)$ so that the pair of measures $(\mu = f\mathcal{L}, M = m\mathcal{L} \otimes \mathcal{L})$ therefore belong to $\mathcal{GCE}_T^E$. Indeed, the distributional grazing continuity equation from Definition 17 is precisely the weak ϵ-Landau equation. Based on the definition of M and the finite $\mathcal{H}_\epsilon$ dissipation, we have the bound

$$\int_0^T \mathcal{A}(\mu_t, M_t)\, dt = \int_0^T D_\epsilon(f_t)\, dt < \infty,$$

which implies the weak continuity of μ. By Proposition 27, we have

$$|\dot{\mu}|^2(t) = \mathcal{A}(\mu_t, M_t) = D_\epsilon(f_t) < \infty \quad \text{for a.e. } t \in [0, T].$$

Using Proposition 28, we have, for any $0 \leq s \leq r \leq T$,

$$\mathcal{H}_\epsilon[\mu_r] - \mathcal{H}_\epsilon[\mu_s] + \frac{1}{2}\int_s^r D_\epsilon(\mu_t)\,dt + \frac{1}{2}\int_s^r |\dot{\mu}|^2(t)\,dt \leq 0.$$

According to Definition 5, this is the curve of maximal slope property.

Curve of maximal slope $\Longrightarrow$ weak ϵ-solution: Assume that $\mu = f\mathcal{L}$ is a curve of maximal slope for $\mathcal{H}_\epsilon$ with respect to the upper gradient $\sqrt{D_\epsilon}$. Since μ is absolutely continuous with respect to d_L, Proposition 27 guarantees existence of a unique curve $M : t \in [0, T] \mapsto M_t \in \mathcal{M}^d$ such that $\int_0^T \sqrt{\mathcal{A}(\mu_t, M_t)}\,dt < \infty$ and $|\dot{\mu}|^2(t) = \mathcal{A}(\mu_t, M_t)$ for a.e. $t \in [0, T]$. Furthermore, $(\mu, M) \in \mathcal{GCE}_T^E$. According to Lemma 19, let $M = m\mathcal{L} \otimes \mathcal{L}$ for some measurable function m. We apply the chain rule (18) with Cauchy–Schwarz and Young's inequalities with minus signs in the following computations:

$$\begin{aligned}
\mathcal{H}_\epsilon[f_T] - \mathcal{H}_\epsilon[f_0] &= \frac{1}{2}\int_0^T \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta f} \cdot m\,dv\,dv_*\,dt \\
&\geq -\frac{1}{2}\int_0^T \left(\iint_{\mathbb{R}^{2d}} ff_*\left|\tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta f}\right|^2 dv\,dv_*\right)^{1/2}\left(\iint_{\mathbb{R}^{2d}} \frac{|m|^2}{ff_*}\,dv\,dv_*\right)^{1/2}dt \\
&\geq -\frac{1}{2}\int_0^T \left(\frac{1}{2}\iint_{\mathbb{R}^{2d}} ff_*\left|\tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta f}\right|^2 dv\,dv_*\right)dt - \frac{1}{2}\int_0^T \left(\frac{1}{2}\iint_{\mathbb{R}^{2d}} \frac{|m|^2}{ff_*}\,dv\,dv_*\right)dt \\
&= -\frac{1}{2}\int_0^T D_\epsilon(f_t)\,dt - \frac{1}{2}\int_0^T |\dot{f}|^2(t)\,dt.
\end{aligned}$$

All the inequalities in the calculations above are actually equalities owing to the fact that μ is a curve of maximal slope. In particular, since we have the equality in the Young's inequality, this implies

$$\frac{m}{\sqrt{ff_*}} = -\sqrt{ff_*}\tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta f}.$$

As in the previous direction, the weak ϵ Landau equation coincides with the grazing continuity equation when m is equal to $-ff_*\tilde{\nabla}(\delta\mathcal{H}_\epsilon/\delta f)$. $\qquad\square$

The rest of this section is devoted to proving Proposition 28. We need some lemmas to establish crucial estimates. The following result is a variation of [Carlen and Carvalho 1992, Lemma 2.6].

Lemma 30 [Carlen and Carvalho 1992]. *Let μ be a probability measure on $\mathbb{R}^d$ with finite second moment/energy, $m_2(\mu) \leq E$ for $E > 0$. Then, for every $\epsilon > 0$, there exists a constant $C = C(\epsilon, E) > 0$ such that*

$$|\log(\mu * G^\epsilon)(v)| \leq C\left\langle\frac{v}{\epsilon}\right\rangle.$$

Proof. Starting with an upper bound, we easily see

$$\mu * G^\epsilon(v) = \int_{\mathbb{R}^d} G^\epsilon(v - v')\,d\mu(v') \lesssim_\epsilon 1.$$

Turning to the lower bound, we cut off the integration domain to $|v'| \leq R$ for some $R > 0$ to be chosen later. We estimate, for $\epsilon > 0$ small enough,

$$\left\langle \frac{v - v'}{\epsilon} \right\rangle = \sqrt{1 + \left| \frac{v - v'}{\epsilon} \right|^2} \leq \sqrt{1 + 2 \left| \frac{v}{\epsilon} \right|^2 + 2 \left(\frac{R}{\epsilon} \right)^2} \leq \sqrt{2} \left(\left\langle \frac{v}{\epsilon} \right\rangle + \left\langle \frac{R}{\epsilon} \right\rangle \right).$$

This is substituted into $G^\epsilon(v - v')$ to obtain

$$\mu * G^\epsilon(v) \geq \int_{|v'| \leq R} G^\epsilon(v - v') \, d\mu(v') \gtrsim_\epsilon \exp \left\{ -\sqrt{2} \left(\left\langle \frac{v}{\epsilon} \right\rangle + \left\langle \frac{R}{\epsilon} \right\rangle \right) \right\} \int_{|v'| \leq R} d\mu(v').$$

At this point, we appeal to Chebyshev's inequality to see

$$\int_{|v'| \leq R} d\mu(v') = 1 - \int_{|v'| \geq R} d\mu(v') \geq 1 - \frac{1}{R^2} \int_{|v'| \geq R} |v'|^2 \, d\mu(v').$$

We can now choose, for example, large R such that $1 - E/R^2 \geq \frac{1}{2}$ to uniformly lower bound the integral $\int_{|v'| \leq R} d\mu(v')$ away from 0 and then conclude the result after applying logarithms. $\square$

Lemma 31 (log-derivative estimates). *For fixed $\epsilon > 0$ we have the formula*

$$\nabla G^\epsilon(v) = -\frac{1}{\epsilon} \left\langle \frac{v}{\epsilon} \right\rangle^{-1} G^\epsilon(v) \frac{v}{\epsilon}. \tag{19}$$

For $\mu \in \mathscr{P}(\mathbb{R}^d)$, defining $\partial^i = \partial / \partial v^i$ and $\partial^{ij} = \partial^2 / (\partial v^i \partial v^j)$, we obtain

$$|\nabla \log(\mu * G^\epsilon)(v)| \leq \frac{1}{\epsilon}, \quad |\partial^{ij} \log(\mu * G^\epsilon)(v)| \leq \frac{4}{\epsilon^2}. \tag{20}$$

Proof. Equation (19) is a direct computation after noticing

$$\frac{\nabla G^\epsilon}{G^\epsilon} = \nabla \log G^\epsilon = \nabla \left(-\left\langle \frac{v}{\epsilon} \right\rangle + \text{const.} \right) = -\frac{1}{\epsilon} \left\langle \frac{v}{\epsilon} \right\rangle^{-1} \frac{v}{\epsilon}.$$

The first order log-derivative estimate of (20) is calculated using formula (19) to obtain

$$|\nabla(\mu * G^\epsilon)(v)| = |\mu * \nabla G^\epsilon(v)| \leq \frac{1}{\epsilon} \int_{\mathbb{R}^d} \left\langle \frac{v - v'}{\epsilon} \right\rangle^{-1} \left| \frac{v - v'}{\epsilon} \right| G^\epsilon(v - v') \, d\mu(v')$$

$$\leq \frac{1}{\epsilon} \int_{\mathbb{R}^d} G^\epsilon(v - v') \, d\mu(v') = \frac{1}{\epsilon}(\mu * G^\epsilon)(v).$$

For the second order, we first look at $\partial^{ij} \mu * G^\epsilon$ which can be computed with the help of (19):

$$|\partial^{ij} \mu * G^\epsilon(v)|$$

$$= \left| \partial^i \left(-\frac{1}{\epsilon} \int_{\mathbb{R}^d} \left\langle \frac{v - v'}{\epsilon} \right\rangle^{-1} \frac{v^j - v'^j}{\epsilon} G^\epsilon(v - v') \, d\mu(v') \right) \right|$$

$$= \left| \frac{1}{\epsilon^2} \int_{\mathbb{R}^d} \left(\left\langle \frac{v - v'}{\epsilon} \right\rangle^{-3} \frac{v^i - v'^i}{\epsilon} \frac{v^j - v'^j}{\epsilon} + \delta^{ij} \left\langle \frac{v - v'}{\epsilon} \right\rangle^{-1} - \left\langle \frac{v - v'}{\epsilon} \right\rangle^{-2} \frac{v^i - v'^i}{\epsilon} \frac{v^j - v'^j}{\epsilon} \right) G^\epsilon(v - v') \, d\mu(v') \right|$$

$$\leq \frac{3}{\epsilon^2} \mu * G^\epsilon(v).$$

Combining this estimate with the previous first-order one, we have

$$|\partial^{ij} \log(\mu * G^\epsilon)(v)| = \left| \frac{\partial^{ij} \mu * G^\epsilon}{\mu * G^\epsilon} - \frac{(\partial^i \mu * G^\epsilon)(\partial^j \mu * G^\epsilon)}{(\mu * G^\epsilon)^2} \right| \le \frac{4}{\epsilon^2}. \qquad \square$$

Lemma 32. *Fix $\epsilon > 0$ and $\gamma \in [-4, 0]$, with $\mu \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ for some $E > 0$. We have*

(1) *Moderately soft case $\gamma \in [-2, 0]$:*

$$\left| \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu} \right| = \left| \tilde{\nabla}[G^\epsilon * \log(\mu * G^\epsilon)](v, v_*) \right| \lesssim_\epsilon |v|^{1+\gamma/2} + |v_*|^{1+\gamma/2}.$$

(2) *Very soft case $\gamma \in [-4, -2]$:*

$$\left| \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu} \right| \lesssim_\epsilon 1.$$

In particular, it holds

$$\iint_{\mathbb{R}^{2d}} \left| \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu} \right|^2 d\mu(v) d\mu(v_*) \le E.$$

Proof. We develop the expression for $\tilde{\nabla}(\delta \mathcal{H}_\epsilon / \delta \mu)$ in integral form to be used throughout this proof:

$$\begin{aligned}
\tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu} &= \tilde{\nabla} G^\epsilon * \log(\mu * G^\epsilon)(v, v_*) \\
&= |v - v_*|^{1+\gamma/2} \Pi[v - v_*](\nabla_v G^\epsilon * \log(\mu * G^\epsilon)(v) - \nabla_{v_*} G^\epsilon * \log(\mu * G^\epsilon)(v_*)) \\
&= |v - v_*|^{1+\gamma/2} \Pi[v - v_*] \int_{\mathbb{R}^d} G^\epsilon(v') \left(\frac{\nabla \mu * G^\epsilon}{\mu * G^\epsilon}(v - v') - \frac{\nabla \mu * G^\epsilon}{\mu * G^\epsilon}(v_* - v') \right) dv'. \qquad (21)
\end{aligned}$$

(1) <u>Moderately soft case $\gamma \in [-2, 0]$</u>: We use (a concave version of) the triangle inequality (valid since $1 + \gamma/2 \ge 0$) and the first estimate of (20) to bound the last line of (21):

$$\left| \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu} \right| \le 2^{1+\gamma/2}(|v|^{1+\gamma/2} + |v_*|^{1+\gamma/2}) \frac{2}{\epsilon} \int_{\mathbb{R}^d} G^\epsilon(v') \, dv' \lesssim_\epsilon |v|^{1+\gamma/2} + |v_*|^{1+\gamma/2}.$$

(2) <u>Very soft case $\gamma \in [-4, -2]$</u>: We perform estimates in two cases, the far field $|v - v_*| \ge 1$ and near field $|v - v_*| \le 1$.

$|v - v_*| \ge 1$: In the far field, we have $|v - v_*|^{1+\gamma/2} \le 1$; hence we can brutally estimate (21) using again the first estimate of (20) to obtain, similar to the moderately soft case, the estimate

$$\left| \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu} \right| \le \frac{2}{\epsilon}.$$

$|v - v_*| \le 1$: We can remove the singularity from the weight with a mean-value estimate and the second estimate of (20):

$$\left| \frac{\nabla \mu * G^\epsilon}{\mu * G^\epsilon}(v - v') - \frac{\nabla \mu * G^\epsilon}{\mu * G^\epsilon}(v_* - v') \right| \le \sup_{i,j=1,\dots,d} \left\| \partial^i \left(\frac{\partial^j \mu * G^\epsilon}{\mu * G^\epsilon} \right) \right\|_{L^\infty} |v - v_*| \le \frac{4}{\epsilon^2} |v - v_*|.$$

Inserting this into (21), we have

$$\left|\tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta\mu}\right| \leq \frac{4}{\epsilon^2}|v-v_*|^{2+\gamma/2}\int_{\mathbb{R}^d}G^\epsilon(v')\,dv' \leq \frac{4}{\epsilon^2}. \qquad \square$$

Remark 33. Originally, we considered the general family of convolution kernels $G^{s,\epsilon}$ described in Section 2.1. Besides the context of the Landau equation, Lemma 31 (excluding the second-order log-derivative estimate) can be generalized to this family of s-order tailed exponential distributions with additional moment assumptions on μ. In particular, (19) and (20) (for $s \geq 1$) become

$$\frac{\nabla G^{s,\epsilon}}{G^{s,\epsilon}}(v) = -\frac{s}{\epsilon}\left\langle\frac{v}{\epsilon}\right\rangle^{s-2}\frac{v}{\epsilon}, \quad \frac{|\nabla(\mu*G^{s,\epsilon})|}{\mu*G^{s,\epsilon}}(v) \lesssim \frac{1}{\epsilon^s}\langle v\rangle^{s-1}.$$

Since Maxwellians are known to be stationary solutions for the Landau equation, we wanted to perform the regularization with $s = 2$. However, the analogous estimates of Lemma 31 for $s = 2$ are not sufficient for Lemma 32 in the $\mathscr{P}_2$ framework. For example, in the moderately soft potential case, the estimate reads

$$\left|\tilde{\nabla}\frac{\delta\mathcal{H}_{2,\epsilon}}{\delta\mu}\right| \lesssim_\epsilon \langle v\rangle^{2+\gamma/2} + \langle v_*\rangle^{2+\gamma/2} \notin L^2(\mu\otimes\mu).$$

However, there is one value of $\gamma = -2$ for which the estimates hold when using a Maxwellian regularization kernel $G^{2,\epsilon}$. A restriction to $\mathscr{P}_4$ resolves the issue mentioned above for the moderately soft potential case, but then a fourth moment propagation is needed, which we did not pursue. A similar issue is present in the very soft potential case.

Proof of Proposition 28. To prove (18), our strategy is to regularize the pair (μ, M) in time with parameter $\delta > 0$ and differentiate the regularization. Then we obtain uniform bounds in δ needed to take the limit $\delta \to 0$.

<u>Finite regularized entropy:</u> We have the following chain of inequalities:

$$\mathcal{H}_\epsilon[\mu_t] = \int_{\mathbb{R}^d}(\mu_t*G^\epsilon)(v)\log(\mu_t*G^\epsilon)(v)\,dv \lesssim_{\epsilon,E} \int_{\mathbb{R}^d}(\mu_t*G^\epsilon)(v)\langle v\rangle\,dv \lesssim_\epsilon 1+E.$$

The first inequality comes from Lemma 30 because $\log(\mu_t*G^\epsilon)$ has linear growth (uniform in time) while in the second inequality, one realizes that μ_t*G^ϵ has as many moments as μ_t with computable constants.

<u>Time regularization with $\delta > 0$:</u> Without loss of generality, let μ be the weakly time continuous representative (Lemma 14) and M be the optimal grazing rate (Proposition 27) achieving the finite distance d_L. We first regularize the pair (μ, M) in time for a fixed parameter $\delta > 0$ as follows. Take $\eta \in C_0^\infty(\mathbb{R})$ with the following properties:

$$\text{supp }\eta \subset (-1,1), \quad \eta \geq 0, \quad \eta(t) = \eta(-t), \quad \int_{-1}^1\eta(t)\,dt = 1.$$

We define the following measures for $t \in [0, T]$, by taking convex combinations:

$$\mu_t^\delta := \int_{-1}^1\eta(t')\mu_{t-\delta t'}\,dt', \quad M_t^\delta := \int_{-1}^1\eta(t')M_{t-\delta t'}\,dt'.$$

Here, we constantly extend the measures in time. That is, if $t - \delta t' \in [-\delta, 0]$, we treat $\mu_{t-\delta t'} = \mu_0$, $M_{t-\delta t'} = 0$. For the other end point, if $t - \delta t' \in [T, T + \delta]$, we set $\mu_{t-\delta t'} = \mu_T$, $M_{t-\delta t'} = 0$. This transformation is stable so that $(\mu^\delta, M^\delta) \in \mathcal{GCE}_T$ and in particular, the distributional grazing continuity equation holds:

$$\partial_t \mu_t^\delta + \tfrac{1}{2} \tilde{\nabla} \cdot M_t^\delta = 0.$$

We derive (18) using this regularized grazing continuity equation. Consider

$$\mathcal{H}_\epsilon[\mu_t^\delta] = \int_{\mathbb{R}^d} (\mu_t^\delta * G^\epsilon)(v) \log(\mu_t^\delta * G^\epsilon)(v) \, dv,$$

which we differentiate with respect to t by appealing to the dominated convergence theorem. Firstly, due to the time regularization, we have

$$\partial_t \{(\mu_t^\delta * G^\epsilon) \log(\mu_t^\delta * G^\epsilon)\} = [(\partial_t \mu_t^\delta) * G^\epsilon](\log(\mu_t^\delta * G^\epsilon) + 1).$$

The L_v^1 bound is obtained on the following difference quotient for a fixed time step $h > 0$:

$$\left| \frac{1}{h} [(\mu_{t+h}^\delta * G^\epsilon) \log(\mu_{t+h}^\delta * G^\epsilon) - (\mu_t^\delta * G^\epsilon) \log(\mu_t^\delta * G^\epsilon)] \right|$$
$$\leq \frac{1}{h} |(\mu_{t+h}^\delta * G^\epsilon) - (\mu_t^\delta * G^\epsilon)| \sup_{s \in [t, t+h]} |\log(\mu_s^\delta * G^\epsilon) + 1|,$$

where we have used the mean value theorem with the chain rule. Applying Lemma 30, we obtain

$$\left| \frac{1}{h} [(\mu_{t+h}^\delta * G^\epsilon) \log(\mu_{t+h}^\delta * G^\epsilon) - (\mu_t^\delta * G^\epsilon) \log(\mu_t^\delta * G^\epsilon)] \right| \lesssim_{\epsilon, E} \frac{1}{h} |(\mu_{t+h}^\delta * G^\epsilon) - (\mu_t^\delta * G^\epsilon)| \langle v \rangle.$$

We apply the mean value theorem on the difference quotient again to get

$$\left| \frac{1}{h} [(\mu_{t+h}^\delta * G^\epsilon) \log(\mu_{t+h}^\delta * G^\epsilon) - (\mu_t^\delta * G^\epsilon) \log(\mu_t^\delta * G^\epsilon)] \right| \lesssim_{\delta, \epsilon} \|\eta'\|_{L^\infty} \left(\mu_0 * G^\epsilon + \int_0^T \mu_t * G^\epsilon \, dt \right) \langle v \rangle.$$

Since μ has finite second-order moments, this last expression belongs to L_v^1. By the dominated convergence theorem,

$$\frac{d}{dt} \mathcal{H}_\epsilon[\mu_t^\delta] = \int_{\mathbb{R}^d} [(\partial_t \mu_t^\delta) * G^\epsilon](\log(\mu_t^\delta * G^\epsilon) + 1) \, dv = \int_{\mathbb{R}^d} (\partial_t \mu_t^\delta) \cdot [G^\epsilon * \log(\mu_t^\delta * G^\epsilon)] \, dv.$$

The last line is achieved by the self-adjointness of convolution with G^ϵ and eliminating the constant term due to the conserved mass of μ^δ. Integrating in t, we obtain

$$\mathcal{H}_\epsilon[\mu_r^\delta] - \mathcal{H}_\epsilon[\mu_s^\delta] = \int_s^r \int_{\mathbb{R}^d} (\partial_t \mu_t^\delta) \cdot [G^\epsilon * \log(\mu_t^\delta * G^\epsilon)] \, dv \, dt$$
$$= \frac{1}{2} \int_s^r \iint_{\mathbb{R}^{2d}} [\tilde{\nabla} G^\epsilon * \log(\mu_t^\delta * G^\epsilon)] \cdot dM_t^\delta \, dt$$
$$= \frac{1}{2} \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t^\delta} \cdot dM_t^\delta \, dt. \tag{22}$$

We now turn to establishing estimates independent of $\delta > 0$ to pass to the limit.

Estimates on the right-hand side of (22): According to Lemma 32, we have the estimate

$$\left|\tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta\mu^\delta}\right| \lesssim_{\epsilon,E} |v|^p + |v_*|^p,$$

where $p \leq 1$. By the first moment assumption of M_t, we have

$$\int_0^T \iint_{\mathbb{R}^{2d}} \left|\tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta\mu_t^\delta}\right| d|M_t|(v,v_*)\, dt \lesssim_{\epsilon,E} \int_0^T \iint_{\mathbb{R}^{2d}} |v| + |v_*|\, d|M_t|(v,v_*)\, dt < \infty.$$

This estimate also extends to M_t^δ

$$\int_0^T \iint_{\mathbb{R}^{2d}} \left|\tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta\mu_t^\delta}\right| d|M_t^\delta|(v,v_*)\, dt < \infty.$$

Note that these estimates are independent of $\delta > 0$.

Convergence $\delta \to 0$: Firstly, we establish the following identity which will be useful later. For fixed functions f^1, f^2 we have

$$\tilde{\nabla}[G^\epsilon * f^1] - \tilde{\nabla}[G^\epsilon * f^2]$$
$$= |v-v_*|^{1+\gamma/2}\Pi[v-v_*](\nabla[G^\epsilon * f^1] - \nabla[G^\epsilon * f^2] - (\nabla_*[G^\epsilon * f^1]_* - \nabla_*[G^\epsilon * f^2]_*))$$
$$= |v-v_*|^{1+\gamma/2}\Pi[v-v_*]\int_{\mathbb{R}^d}(\nabla G^\epsilon(v-v') - \nabla G^\epsilon(v_*-v'))(f^1(v') - f^2(v'))\, dv'. \tag{23}$$

Using the weak in time continuity of μ, we can consider

$$|\mu_t^\delta * G^\epsilon(v') - \mu_t * G^\epsilon(v')| \leq \int_{-1}^1 \eta(t')|\langle \mu_{t-\delta t'}, G^\epsilon(v'-\cdot)\rangle - \langle \mu_t, G^\epsilon(v'-\cdot)\rangle|\, dt'.$$

The "$\cdot$" stands for the convoluted variable. Since t belongs to a compact set, the function $t \mapsto \langle \mu_t, G^\epsilon(v'-\cdot)\rangle$ is *uniformly* continuous from the weak continuity of μ. In particular, using the continuity in v' and the lower bound from Lemma 30 we conclude that for any $R > 0$

$$|\log(\mu_t^\delta * G^\epsilon) - \log(\mu_t * G^\epsilon)| \to 0 \quad \text{uniformly on } B_R. \tag{24}$$

Therefore by Lemma 30, defining $w = |v-v_*|^{1+\gamma/2}$, and using (23) with $f^1 = \log(\mu_t^\delta * G^\epsilon)$ and $f^2 = \log(\mu_t * G^\epsilon)$, we have

$$\left|\tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta\mu_t^\delta} - \tilde{\nabla}\frac{\delta\mathcal{H}_\epsilon}{\delta\mu_t}\right| = |\tilde{\nabla}G^\epsilon * \log(\mu_t^\delta * G^\epsilon)(v,v_*) - \tilde{\nabla}G^\epsilon * \log(\mu_t * G^\epsilon)(v,v_*)|$$

$$\leq \int_{\mathbb{R}^d} w|\nabla G^\epsilon(v-v') - \nabla G^\epsilon(v_*-v')||\log(\mu_t^\delta * G^\epsilon(v')) - \log(\mu_t * G^\epsilon(v'))|\, dv'$$

$$\leq \int_{B_{R_0}^c} w|\nabla G^\epsilon(v-v') - \nabla G^\epsilon(v_*-v')|C_\epsilon\langle v'\rangle\, dv'$$
$$+ \sup_{B_{R_0}}|\log(\mu_t^\delta * G^\epsilon) - \log(\mu_t * G^\epsilon)|\int_{B_{R_0}} w|\nabla G^\epsilon(v-v') - \nabla G^\epsilon(v_*-v')|\, dv'.$$

For a fixed (v,v_*), we obtain the convergence to zero by taking $\delta \to 0$ and $R_0 \to \infty$ in the previous estimate. This holds for all $\gamma \in [-4,0]$ by taking advantage of the regularity of G^ϵ. Using continuity,

we obtain that for any $R > 0$

$$\left| \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t^\delta}(v, v_*) - \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t}(v, v_*) \right| \to 0 \quad \text{uniformly on } [0, T] \times B_R \times B_R. \tag{25}$$

We turn to the limit estimate for the right-hand side of (22). For any $R > 0$, we have

$$\left| \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t^\delta} \cdot dM_t^\delta \, dt - \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t} \cdot dM_t \, dt \right|$$

$$\leq \left| \int_s^r \iint_{\mathbb{R}^{2d}} \left(\tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t^\delta} - \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t} \right) \cdot dM_t^\delta \, dt \right| + \left| \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t} \cdot dM_t^\delta \, dt - \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t} \cdot dM_t \, dt \right|$$

$$\leq \int_s^r \iint_{B_R \times B_R} \left| \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t^\delta} - \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t} \right| d|M_t^\delta| \, dt + \int_s^r \iint_{(B_R \times B_R)^C} \left| \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t^\delta} - \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t} \right| d|M_t^\delta| \, dt + o(1).$$

The last term is $o(1)$ as $\delta \to 0$ due to similar estimates from the previous step. By sending $\delta \to 0$ (the first term vanishes due to (25)) and then sending $R \to \infty$ (the second term vanishes again due to the estimate from the previous step), we obtain the convergence

$$\lim_{\delta \to 0} \frac{1}{2} \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t^\delta} \cdot dM_t^\delta \, dt = \frac{1}{2} \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t} \cdot dM_t^\delta \, dt. \tag{26}$$

<u>Convergence of the left-hand side of (22):</u> By (24), Lemma 30 and the uniform bound on the second moment, we have

$$|\mathcal{H}_\epsilon[\mu_t^\delta] - \mathcal{H}_\epsilon[\mu_t]| \leq \int_{\mathbb{R}^d} |(\mu_t^\delta * G^\epsilon) \log(\mu_t^\delta * G^\epsilon)(v) - (\mu_t * G^\epsilon) \log(\mu_t * G^\epsilon)(v)| \, dv$$

$$\to 0 \quad \text{as } \delta \to 0.$$

Therefore, by the previous equation and (26) we can take $\delta \to 0$ in (22) to obtain

$$\mathcal{H}_\epsilon[\mu_r] - \mathcal{H}_\epsilon[\mu_s] = \frac{1}{2} \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu_t} \cdot dM_t(v, v_*) \, dt,$$

which is the desired result. $\qquad \square$

5. JKO scheme for ϵ-Landau equation

This section is devoted to the proof of Theorem 9 after a series of preliminary lemmas. Our construction of curves of maximal slope in Theorem 9 uses the basic minimizing movement/variational approximation scheme of [Jordan et al. 1998]. Fix a small time step $\tau > 0$ and initial datum $\mu_0 \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ and consider the recursive minimization procedure for $n \in \mathbb{N}$

$$\nu_0^\tau := \mu_0, \quad \nu_n^\tau \in \operatorname{argmin}_{\lambda \in \mathscr{P}_{2,E}} \left[\mathcal{H}_\epsilon(\lambda) + \frac{1}{2\tau} d_L^2(\nu_{n-1}^\tau, \lambda) \right]. \tag{27}$$

Then, we concatenate these minimizers into a curve by setting

$$\mu_0^\tau := \mu_0, \qquad \mu_t^\tau := \nu_n^\tau \quad \text{for } t \in ((n-1)\tau, n\tau]. \tag{28}$$

The scheme given by (27) and (28) satisfies the abstract formulation in [Ambrosio et al. 2008] giving:

Proposition 34 (Landau JKO scheme). *For any $\tau > 0$ and $\mu_0 \in \mathscr{P}_{2,E}(\mathbb{R}^d)$, there exists $\nu_n^\tau \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ for every $n \in \mathbb{N}$ as described in* (27). *Furthermore, up to a subsequence of μ_t^τ described in* (28) *as $\tau \to 0$, there exists a locally absolutely continuous curve $(\mu_t)_{t \geq 0}$ such that*

$$\mu_t^\tau \rightharpoonup \mu_t \quad \text{for all } t \in [0, \infty).$$

Proof. Our metric setting is $(\mathscr{P}_{\mu_0}, d_L)$ (see Theorem 7) with the weak topology σ. This space is essentially $\mathscr{P}_{2,E}(\mathbb{R}^d)$ except we need to make sure that d_L is a proper metric; hence we remove the probability measures with infinite Landau distance. We follow the proof of [Erbar 2023], which consists in verifying [Ambrosio et al. 2008, Assumptions 2.1(a)–(c)]. These assumptions are listed and verified now.

(1) $\mathcal{H}_\epsilon$ is sequentially σ-lsc on d_L-bounded sets: Suppose $\mu_n \in \mathscr{P}_{2,E}(\mathbb{R}^d) \rightharpoonup \mu \in \mathscr{P}_{2,E}(\mathbb{R}^d)$. This implies $\mu_n * G^\epsilon \rightharpoonup \mu * G^\epsilon$ in $\mathscr{P}_2(\mathbb{R}^d)$. It is known that

$$\mathcal{H}(\mu) = \begin{cases} \int_{\mathbb{R}^d} f(v) \log f(v) \, dv, & \mu = f\mathcal{L}, \\ +\infty, & \text{else} \end{cases}$$

is σ-lsc and since $\mathcal{H}_\epsilon(\mu) = \mathcal{H}(\mu * G^\epsilon)$, we achieve the first property.

(2) $\mathcal{H}_\epsilon$ is lower bounded: By Lemma 30 for fixed $\epsilon > 0$, $\log(\mu * G^\epsilon)$ is uniformly lower bounded by a linearly growing term. For fixed $\mu \in \mathscr{P}_{2,E}(\mathbb{R}^d)$, we have, with Cauchy–Schwarz,

$$\mathcal{H}_\epsilon(\mu) \gtrsim_\epsilon - \int_{\mathbb{R}^d} \langle v \rangle \mu * G^\epsilon(v) \, dv \geq -\left(\int_{\mathbb{R}^d} \langle v \rangle^2 \mu * G^\epsilon(v) \, dv \right)^{1/2} \geq -(\mathcal{O}(\epsilon) + E)^{1/2} > -\infty.$$

(3) d_L-bounded sets are relatively sequentially σ-compact: This is one of the consequences from Theorem 7.

The existence of minimizers, ν_n^τ, to (27) and limits, μ_t, to (28) is guaranteed from [Ambrosio et al. 2008, Corollary 2.2.2 and Proposition 2.2.3], respectively. □

At the abstract level, the limit curve constructed in Proposition 34 has no relation to $\sqrt{D_\epsilon}$. The following lemmas bridge this gap.

Lemma 35. *For any $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$, we have*

$$\sqrt{D_\epsilon(\mu_0)} \leq |\partial^- \mathcal{H}_\epsilon|(\mu_0).$$

Proof. For fixed $\epsilon, R_1, R_2 > 0$ and $\gamma \in \mathbb{R}$, take $T > 0$ from Theorem 48 in the Appendix and the unique weak solution $\mu \in C([0, T]; \mathscr{P}_2(\mathbb{R}^d))$ to

$$\begin{cases} \partial_t \mu = \nabla \cdot \{\mu \phi_{R_1} \int_{\mathbb{R}^d} \phi_{R_{1*}} \psi_{R_2}(v - v_*) |v - v_*|^{\gamma+2} \Pi[v - v_*](J_0^\epsilon - J_{0*}^\epsilon) \, d\mu(v_*)\}, \\ \mu(0) = \mu_0. \end{cases}$$

The functions $0 \leq \phi_{R_1}, \psi_{R_2} \leq 1$ are smooth cut-off functions with the following properties:

$$\phi_{R_1}(v) = \begin{cases} 1, & |v| \leq R_1, \\ 0, & |v| \geq R_1 + 1, \end{cases} \qquad \psi_{R_2}(z) = \begin{cases} 0, & |z| \leq 1/R_2, \\ 1, & |z| \geq 2/R_2. \end{cases}$$

The notation J_0^ϵ from the Appendix means

$$J_0^\epsilon = \nabla G^\epsilon * \log[\mu_0 * G^\epsilon] \in C^\infty(\mathbb{R}^d; \mathbb{R}^d).$$

For this proof alone, we define the reduced ϵ-entropy-dissipation

$$D_\epsilon^{R_1,R_2}(\mu_0) := \frac{1}{2} \iint_{\mathbb{R}^{2d}} \phi_{R_1} \phi_{R_1*} \psi_{R_2} (v - v_*) |v - v_*|^{\gamma+2} |\Pi[v - v_*](J_0^\epsilon - J_{0*}^\epsilon)|^2 d\mu_0(v) d\mu_0(v_*).$$

On the other hand, as the ϵ-entropy dissipation comes from the negative time derivative of entropy, we have

$$\begin{aligned}
D_\epsilon^{R_1,R_2}(\mu_0) &= \lim_{t\downarrow 0} \frac{\mathcal{H}_\epsilon(\mu_0) - \mathcal{H}_\epsilon(\mu_t)}{t} = \lim_{t\downarrow 0} \frac{\mathcal{H}_\epsilon(\mu_0) - \mathcal{H}_\epsilon(\mu_t)}{d_L(\mu_0, \mu_t)} \frac{d_L(\mu_0, \mu_t)}{t} \\
&\leq \lim_{t\downarrow 0} \left\{ \frac{\mathcal{H}_\epsilon(\mu_0) - \mathcal{H}_\epsilon(\mu_t)}{d_L(\mu_0, \mu_t)} \times \frac{1}{t} \right. \\
&\qquad \left. \times \left(\int_0^t \sqrt{\frac{1}{2} \iint_{\mathbb{R}^{2d}} \phi_{R_1}^2 \phi_{R_1*}^2 \psi_{R_2}^2 |v - v_*|^{\gamma+2} |\Pi[v - v_*](J_0^\epsilon - J_{0*}^\epsilon)|^2 d\mu_s(v) d\mu_s(v_*) ds} \right) \right\} \\
&\leq |\partial\mathcal{H}_\epsilon|(\mu_0) \sqrt{D_\epsilon^{R_1,R_2}(\mu_0)}.
\end{aligned}$$

In the first inequality, we estimated $d_L(\mu_0, \mu_t)$ by considering the PDE in this lemma as the grazing collision equation with $M = -(\mu \otimes \mu)\tilde{\nabla} \log \mu_0$. In the last inequality, we have used the Lebesgue differentiation theorem with strong-weak convergence since μ is continuous in time as well as the fact that $\phi_{R_1}^2 \leq \phi_{R_1}$ and $\psi_{R_2}^2 \leq \psi_{R_2}$ since $0 \leq \phi_{R_1}, \psi_{R_2} \leq 1$. We are left with the inequality

$$\sqrt{D_\epsilon^{R_1,R_2}(\mu_0)} \leq |\partial\mathcal{H}_\epsilon|(\mu_0) \quad \text{for all } R_1, R_2 > 0.$$

Owing to the many regularizations applied, the ϵ-entropy-dissipation $\mu \mapsto D_\epsilon^{R_1,R_2}(\mu)$ is continuous with respect to weak convergence of probability measures. By considering weakly convergent sequences and passing to the limit inferior, we deduce the same inequality with the relaxed slope

$$\sqrt{D_\epsilon^{R_1,R_2}(\mu_0)} \leq |\partial^-\mathcal{H}_\epsilon|(\mu_0) \quad \text{for all } R_1, R_2 > 0.$$

As functions of R_1, R_2 individually, $D_\epsilon^{R_1,R_2}(\mu_0)$ is nondecreasing. Furthermore, the integrand of $D_\epsilon^{R_1,R_2}(\mu_0)$ converges to the integrand of $D_\epsilon(\mu_0)$ pointwise μ_0-almost every v, v_*. Thus, an application of the monotone convergence theorem in the limit $R_1, R_2 \to \infty$ on the above inequality completes the proof. $\quad\square$

Lemma 36. $|\partial^-\mathcal{H}_\epsilon|$ *is a strong upper gradient for* $\mathcal{H}_\epsilon$ *in* $\mathscr{P}_{\mu_0}(\mathbb{R}^d)$, *where* $\mu_0 \in \mathscr{P}_{2,E}(\mathbb{R}^d)$.

Proof. Fix $\lambda, \nu \in \mathscr{P}_{\mu_0}(\mathbb{R}^d)$ so that by the triangle inequality of Theorem 7, we have $d_L(\lambda, \nu) < \infty$. Now by Proposition 26, there exists a pair of curves $(\mu, M) \in \mathcal{GCE}_1^E$ connecting λ, ν and $\mathcal{A}(\mu_t, M_t) = d_L^2(\lambda, \nu)$ for almost every $t \in [0, 1]$. Using Remark 29 and Lemma 35, we have

$$|\mathcal{H}_\epsilon(\lambda) - \mathcal{H}_\epsilon(\nu)| \leq \int_0^1 \sqrt{D_\epsilon(\mu_t)} |\dot{\mu}|(t)\, dt \leq \int_0^1 |\partial^-\mathcal{H}_\epsilon|(\mu_t) |\dot{\mu}|(t)\, dt. \qquad \square$$

We now have all the ingredients to prove Theorem 9 so that we can relate curves of maximal slope to weak solutions of the ϵ-Landau equation.

Proof of Theorem 9. Take a limit curve μ_t constructed in Proposition 34. By the previous Lemma 36, the assumptions of [Ambrosio et al. 2008, Theorem 2.3.3] are fulfilled so the curve is of maximal slope with respect to $|\partial^- \mathcal{H}_\epsilon|$ and satisfies the associated energy dissipation inequality

$$\mathcal{H}_\epsilon(\mu_r) - \mathcal{H}_\epsilon(\mu_s) + \frac{1}{2} \int_s^r |\partial^- \mathcal{H}_\epsilon(\mu_t)|^2 \, dt + \frac{1}{2} \int_s^r |\dot{\mu}|^2(t) \, dt \leq 0.$$

The inequality of Lemma 35 gives

$$\mathcal{H}_\epsilon(\mu_r) - \mathcal{H}_\epsilon(\mu_s) + \frac{1}{2} \int_s^r D_\epsilon(\mu_t) \, dt + \frac{1}{2} \int_s^r |\dot{\mu}|^2(t) \, dt \leq 0,$$

which is precisely the statement that the limit curve μ_t is a curve of maximal slope with respect to $\sqrt{D_\epsilon}$. □

Remark 37. The results of Proposition 34 and Lemma 35 can be generalized to other regularization kernels $G^{s,\epsilon}$, in particular, the Maxwellian regularization. However, this is not the case for Lemma 36 since the proof relies on Proposition 28; see Remark 33.

6. Recovering the full Landau equation as $\epsilon \to 0$

Theorems 8 and 9 provide the basic existence theory for the $\epsilon > 0$ approximation of the Landau equation. In this section, we prove the $\epsilon \downarrow 0$ analogue of Theorem 8, which is Theorem 12. By definition, both H-solutions and curves of maximal slope to the full Landau equation dissipate the entropy. Therefore, the assumption of finite initial entropy (A2) automatically ensures

$$\sup_{t \in [0,T]} \mathcal{H}[f_t] = \sup_{t \in [0,T]} \int_{\mathbb{R}^3} f_t \log f_t < +\infty.$$

In the sequel, every quotation of (A2) will refer to this bound.

Sketch of the proof of Theorem 12. By repeating the proof of Theorem 8, we see that the crucial ingredient is the chain rule (18) in Proposition 28. For now assume the following:

Claim 38. *Assume* (A1), (A2), (A3) *and let M be any grazing rate such that $(\mu, M) \in \mathcal{GCE}_T^E$ and*

$$\int_0^T \mathcal{A}(\mu_t, M_t) \, dt < \infty.$$

Then we have the chain rule

$$\mathcal{H}[\mu_r] - \mathcal{H}[\mu_s] = \frac{1}{2} \int_s^r \iint_{\mathbb{R}^6} \tilde{\nabla} \left[\frac{\delta \mathcal{H}}{\delta \mu} \right] \cdot dM_t \, dt. \tag{29}$$

By following the steps of the proof of Theorem 8 and using (29) instead of (18), one completes the proof of Theorem 12. We dedicate this section to proving Claim 38.

Equation (29) is clearly the $\epsilon \downarrow 0$ limit of (18). The left-hand side of (29) can be obtained from the left-hand side of (18) using the finite entropy (A2) and the fact that $\epsilon \mapsto \mathcal{H}_\epsilon[\mu_t]$ is nonincreasing for every t. We refer to [Erbar 2023, Proof of Proposition 4.2, Step 4(d)] for more details on a similar argument.

The difficulty remains in deducing that the right-hand side of (18) converges to the right-hand side of (29) as $\epsilon \downarrow 0$ given by

$$\int_0^T \iint_{\mathbb{R}^6} \tilde{\nabla} \frac{\delta \mathcal{H}_\epsilon}{\delta \mu} \cdot dM_t \, dt \to \int_0^T \iint_{\mathbb{R}^6} \tilde{\nabla} \frac{\delta \mathcal{H}}{\delta \mu} \cdot dM_t \, dt, \quad \epsilon \downarrow 0, \tag{30}$$

under the additional assumptions (A1), (A2), (A3) on f. The key result which we will use repeatedly in this section is the following theorem which is a specific case of the result in [Royden 1963, Chapter 4, Theorem 17].

Theorem 39 (extended dominated convergence theorem (EDCT)). *Let* $(H_\epsilon)_{\epsilon>0}$ *and* $(I_\epsilon)_{\epsilon>0}$ *be sequences of measurable functions on X satisfying* $I_\epsilon \geq 0$ *and suppose there exists measurable functions H, I satisfying*:

(1) $|H_\epsilon| \leq I_\epsilon$ *for every* $\epsilon > 0$ *and pointwise a.e.*

(2) H_ϵ *and* I_ϵ *converge pointwise a.e. to H and I, respectively.*

(3) $\lim_{\epsilon \downarrow 0} \int_X I_\epsilon = \int_X I < \infty$.

Then, we have the convergence

$$\lim_{\epsilon \downarrow 0} \int_X H_\epsilon = \int_X H.$$

Setting $M = m\mathcal{L} \otimes \mathcal{L}$ (valid by Lemma 19) and using Young's inequality on the right-hand side of (18), we obtain the majorants

$$\tilde{\nabla}\left[\frac{\delta \mathcal{H}_\epsilon}{\delta \mu}\right] \cdot m_t \leq \frac{1}{2} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}_\epsilon}{\delta \mu}\right]\right|^2 + \frac{1}{2} \frac{|m_t|^2}{f f_*}.$$

Notice that the first term is precisely the integrand of D_ϵ, while the second term is the integrand of the action functional $\mathcal{A}(\mu_t, M_t)$, which has no dependence on ϵ and is henceforth ignored. We can apply the EDCT (Theorem 39) with $X = (0, T) \times \mathbb{R}^6$ to prove (30) once we show

$$\int_0^T \iint_{\mathbb{R}^6} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}_\epsilon}{\delta \mu}\right]\right|^2 dv_* \, dv \, dt \to \int_0^T \iint_{\mathbb{R}^6} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}}{\delta \mu}\right]\right|^2 dv_* \, dv \, dt, \quad \epsilon \downarrow 0. \tag{31}$$

The pointwise a.e. convergence hypothesis of Theorem 39 is straightforward based on the regularization of $\mathcal{H}_\epsilon$ through G^ϵ. Focusing on (31), we will use a standard dominated convergence theorem (DCT) for the integration in the t-variable, by proving

$$\iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}_\epsilon}{\delta \mu}\right]\right|^2 dv_* \, dv \to \iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}}{\delta \mu}\right]\right|^2 dv_* \, dv \quad \text{for a.e. } t,$$

$$\iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}_\epsilon}{\delta \mu}\right]\right|^2 dv_* \, dv \leq C \iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}}{\delta \mu}\right]\right|^2 dv_* \, dv \quad \text{for a.e. } t, \text{ for all } \epsilon > 0,$$

$$\tag{32}$$

where $C > 0$ is a constant independent of $\epsilon > 0$. The estimate of (32) guarantees the L_t^1 majorization due to the finite entropy-dissipation (A3). $\qquad \square$

Our estimates in this section accomplish both the convergence and the estimate of (32) by nested application of Theorem 39. The significance of all three assumptions (A1), (A2), and (A3) will be apparent in proving the convergence in (32).

Remark 40. In this section, the only properties of G^ϵ we use are that it is a nonnegative radial approximate identity with sufficiently many moments. As in the construction of minimizing movement curves in Section 5, the results of this section can be achieved with other radial approximate identities.

6.1. *Outline of technical strategy to prove* **(32).** The need to apply Theorem 39 instead of the more classical Lebesgue DCT is that we are unable to prove pointwise estimates in v for the function $v \rightarrow f \int_{\mathbb{R}^3} f_* |\tilde{\nabla}[\partial \mathcal{H}_\epsilon / \partial f]|^2 \, dv_*$. Instead, our estimates in this section rely on the self-adjointness of convolution against radial exponentials (SACRE) to construct a convergent majorant in ϵ.

Step 1: finding majorants and appealing to Theorem 39. We seek to find pointwise a.e. majorants in the v-variable:

$$f \int_{\mathbb{R}^3} f_* \left| \tilde{\nabla}\left[\frac{\delta \mathcal{H}_\epsilon}{\delta \mu} \right] \right|^2 dv_* \le I_\epsilon^1(v),$$

where $I_\epsilon^1(v)$ satisfies the hypothesis for the majorant in Theorem 39. We show that I_ϵ^1 converges pointwise to some I^1, since I_ϵ^1 depends on ϵ only through convolutions against G^ϵ, which is an approximation of the identity. Hence, we are left with showing the integral convergence of Theorem 39(3)

$$\int_{\mathbb{R}^3} I_\epsilon^1(v) \, dv \, dt \rightarrow \int_{\mathbb{R}^3} I^1(v) \, dv, \quad \epsilon \rightarrow 0.$$

Step 2: use SACRE with G^ϵ. To show the integral convergence for I_ϵ^1, we find functions A^1 and B^1 such that

$$I_\epsilon^1(v) \le A^1(v)(G^\epsilon * B^1)(v)$$

and apply Theorem 39. As in the previous step, the pointwise convergence is easily proved. Hence, we are left to show the integral convergence

$$\int_{\mathbb{R}^3} A^1(G^\epsilon * B^1) \, dv \rightarrow \int_{\mathbb{R}^3} A^1 B^1, \quad \epsilon \rightarrow 0.$$

The key observation is applying SACRE to obtain

$$\int_{\mathbb{R}^3} A^1(G^\epsilon * B^1) = \int_{\mathbb{R}^3} \overbrace{(G^\epsilon * A^1)B^1}^{=:I_\epsilon^2}.$$

Therefore, we have reduced the problem to showing integral convergence of Theorem 39(3) for I_ϵ^2 (as the pointwise convergence is easily proved).

Step 3: repeat Step 2. We repeat the process outlined in Step 2 by finding functions A^2 and B^2 such that we have the pointwise bound

$$I_\epsilon^2(v) \le A^2(v)(G^\epsilon * B^2)(v).$$

Again the pointwise convergence for the majorant follows easily; hence we only need to check the integral convergence of Theorem 39(3) given by

$$\int_{\mathbb{R}^3} A^2(G^\epsilon * B^2) \to \int_{\mathbb{R}^3} A^2 B^2.$$

Using SACRE, we study instead the integral convergence of

$$I_\epsilon^3(v) = (G^\epsilon * A^2)B^2.$$

Eventually, after a finite number of times of finding majorants and applying SACRE, we will obtain a majorant I_ϵ^i for which the estimates and the convergence as $\epsilon \to 0$ follow from the standard Lebesgue DCT, using the bound of the weighted Fisher information in terms of the entropy-dissipation (see Theorem 41) and (A3).

6.2. *Preparatory results.* As mentioned in the previous section, for the final step of the proof we need a bound on the weighted Fisher information and a closely related variant in terms of the entropy-dissipation originally discovered by the third author in [Desvillettes and Fellner 2006].

Theorem 41. *Suppose $\gamma \in (-4, 0]$ and let $f \geq 0$ be a probability density belong to $L_{2-\gamma}^1 \cap L \log L(\mathbb{R}^3)$. We have*

$$\int_{\mathbb{R}^3} f(v)\langle v \rangle^\gamma \left| \nabla \frac{\delta \mathcal{H}}{\delta f} \right|^2 dv + \int_{\mathbb{R}^3} f(v)\langle v \rangle^\gamma \left| v \times \nabla \frac{\delta \mathcal{H}}{\delta f} \right|^2 dv \leq C(1 + D_{w,\mathcal{H}}(f)),$$

where $C > 0$ is a constant depending only on the bounds of $m_{2-\gamma}(f)$ and the Boltzmann entropy, $\mathcal{H}[f]$, of f.

The estimate in this precise form can be found in [Desvillettes 2022, Proposition 4, p. 10]. We will refer to the second term on the left-hand side as a "cross Fisher information". We mention here that (A2) enters in the sequel since the constant $C > 0$ in Theorem 41 depends on bounds for $\mathcal{H}[f]$.

To decompose the entropy-dissipation in a manageable way that makes the cross Fisher term more apparent, we have the following linear algebra fact.

Lemma 42. *For $x, y \in \mathbb{R}^3$, we have*

$$|x|^2(y \cdot \Pi[x]y) = |x \times y|^2.$$

Proof. Without loss of generality, we assume neither $x, y = 0$ or else the statement holds trivially. Let θ be an oriented angle between x and y. We expand the definition of $\Pi[x]$ and observe

$$\begin{aligned} |x|^2(y \cdot \Pi[x]y) = y \cdot (|x|^2 I - x \otimes x)y &= |x|^2|y|^2 - |x \cdot y|^2 \\ &= |x|^2|y|^2(1 - \cos^2 \theta) = |x|^2|y|^2 \sin^2 \theta = |x \times y|^2. \qquad \square \end{aligned}$$

The following lemma shows how we use (A1) to control the singularity of the weight.

Lemma 43. *Given $\gamma \in (-3, 0]$, assume that f satisfies (A1) for some $0 < \eta \leq \gamma + 3$. Then we have for a.e. t*

$$\int_{\mathbb{R}^3} f_*(t)|v - v_*|^\gamma \, dv_* \leq C_1(t)\langle v \rangle^\gamma, \qquad \int_{\mathbb{R}^3} f_*(t)|v_*|^2|v - v_*|^\gamma \, dv_* \leq C_2(t)\langle v \rangle^\gamma, \tag{33}$$

where

$$\|C_1\|_{L^\infty(0,T)} \lesssim_{\gamma,\eta} \|\langle \cdot \rangle^{-\gamma} f(t)\|_{L^\infty(0,T;L^1 \cap L^{(3-\eta)/(3+\gamma-\eta)}(\mathbb{R}^3))},$$

$$\|C_2\|_{L^\infty(0,T)} \lesssim_{\gamma,\eta} \|\langle \cdot \rangle^{2-\gamma} f(t)\|_{L^\infty(0,T;L^1 \cap L^{(3-\eta)/(3+\gamma-\eta)}(\mathbb{R}^3))}.$$

Proof. We will only prove the first inequality of (33) since the second inequality uses the same procedure. We split the estimation for local $|v| \le 1$ and far-field $|v| \ge 1$.

Case 1: $|v| \le 1$. We split the integral over v_* into two regions

$$\int_{\mathbb{R}^3} f_*|v - v_*|^\gamma \, dv_* = \int_{|v-v_*|\ge 1} f_*|v-v_*|^\gamma \, dv_* + \int_{|v-v_*|\le 1} f_*|v-v_*|^\gamma \, dv_*$$

$$\le 1 + \int_{|v-v_*|\le 1} f_*|v-v_*|^\gamma \, dv_*,$$

where we have used that $\int_{\mathbb{R}^3} f = 1$ and $\gamma \le 0$. For the integral with the singularity, we apply Young's convolution inequality with conjugate exponents $((3-\eta)/(3+\gamma-\eta), (-3+\eta)/\gamma)$

$$\int_{|v-v_*|\le 1} f_*|v-v_*|^\gamma \, dv_* \le \|f * (\chi_{B_1}|\cdot|^\gamma)\|_{L^\infty}$$

$$\le \|f\|_{L^{(3-\eta)/(3+\gamma-\eta)}} \|\chi_{B_1}|\cdot|^\gamma\|_{L^{(-3+\eta)/\gamma}} \le \left(\frac{\omega_2}{\eta}\right)^{(-3+\eta)/\gamma} \|f\|_{L^{(3-\eta)/(3+\gamma-\eta)}}.$$

Here, ω_2 is the volume of the unit sphere in $\mathbb{R}^3$.

Case 2: $|v| \ge 1$. Once again, we split the integral into two parts

$$\int_{\mathbb{R}^3} f_*|v-v_*|^\gamma \, dv_* = \int_{|v_*|\le |v|/2} f_*|v-v_*|^\gamma \, dv_* + \int_{|v_*|\ge |v|/2} f_*|v-v_*|^\gamma \, dv_*$$

$$\le 2^{-\gamma}|v|^\gamma \int_{|v_*|\le |v|/2} f_* \, dv_* + 2^{-\gamma}|v|^\gamma \int_{|v_*|\ge |v|/2} f_*|v_*|^{-\gamma}|v-v_*|^\gamma \, dv_*.$$

The first term and second term come from the following inequalities based on their respective integration regions:

$$|v-v_*| \ge |v| - |v_*| \ge \tfrac{1}{2}|v|, \quad 1 \le 2^{-\gamma}|v|^\gamma |v_*|^{-\gamma}.$$

We estimate the first integral using the unit mass of f, while the second integral is more delicate but again uses the splitting of the previous step to obtain

$$\int_{\mathbb{R}^3} f_*|v-v_*|^\gamma \, dv_*$$

$$\le 2^{-\gamma}|v|^\gamma + 2^{-\gamma}|v|^\gamma \left(\int_{|v-v_*|\ge 1} f_*|v_*|^{-\gamma}|v-v_*|^\gamma \, dv_* + \int_{|v-v_*|\le 1} f_*|v_*|^{-\gamma}|v-v_*|^\gamma \, dv_* \right).$$

In the large brackets, the first integral can be estimated by $m_{-\gamma}(f)$. Now we use the same Young's inequality argument for the remaining integral to obtain

$$\int_{\mathbb{R}^3} f_*|v-v_*|^\gamma \, dv_* \le 2^{-\gamma}|v|^\gamma + 2^{-\gamma}|v|^\gamma \left(m_{-\gamma}(f) + \left(\frac{\omega_2}{\eta}\right)^{(-3+\eta)/\gamma} \||\cdot|^{-\gamma} f\|_{L^{(3-\eta)/(3+\gamma-\eta)}(\mathbb{R}^3)} \right).$$

The proof is complete by combining the estimates for $|v| \le 1$ and $|v| \ge 1$. $\qquad\square$

Lemma 44 (Peetre). *For any $p \in \mathbb{R}$ and $x, y \in \mathbb{R}^d$, we have*

$$\frac{\langle x \rangle^p}{\langle y \rangle^p} \leq 2^{|p|/2} \langle x - y \rangle^{|p|}.$$

Proof. Our proof follows [Barros-Neto 1973]. Starting with the case $p = 2$, for fixed vectors $a, b \in \mathbb{R}^d$ we have, with the help of Young's inequality,

$$1 + |a - b|^2 \leq 1 + |a|^2 + 2|a||b| + |b|^2 \leq 1 + 2|a|^2 + 2|b|^2$$
$$\leq 2 + 2|a|^2 + 2|a|^2|b|^2 + 2|b|^2 = 2(1 + |a|^2)(1 + |b|^2).$$

Dividing by $\langle b \rangle^2$ and setting $a = x - y$, $b = -y$, we obtain the inequality for $p = 2$

$$\frac{\langle x \rangle^2}{\langle y \rangle^2} \leq 2 \langle x - y \rangle^2.$$

By taking nonnegative powers, this proves the inequality for $p \geq 0$. On the other hand, when we divided by $\langle b \rangle^2$ we could have also set $a = x - y$, $b = x$ to obtain

$$\frac{\langle y \rangle^2}{\langle x \rangle^2} \leq 2 \langle x - y \rangle^2.$$

Taking strictly nonnegative powers here proves the inequality for $p < 0$. $\qquad\square$

Next, we prove an estimate for algebraic functions (growing or decaying) convoluted against G^ϵ with respect to the original function.

Lemma 45. *For any $p \in \mathbb{R}$, we have*

$$\int_{\mathbb{R}^d} \langle w \rangle^p G^\epsilon (v - w) \, dw \leq C \langle v \rangle^p,$$

where $C > 0$ is a constant depending only on $|p|$ and $m_{|p|}(G)$.

Proof. We use Peetre's inequality in Lemma 44 to introduce $v - w$ into the angle brackets

$$\int_{\mathbb{R}^d} \langle w \rangle^p G^\epsilon (v - w) \, dw \leq 2^{|p|/2} \langle v \rangle^p \int_{\mathbb{R}^d} \langle v - w \rangle^{|p|} G^\epsilon (v - w) \, dw$$
$$= 2^{|p|/2} \langle v \rangle^p \int_{\mathbb{R}^d} (1 + |w|^2)^{|p|/2} \epsilon^{-d} G\left(\frac{w}{\epsilon}\right) dw$$
$$= 2^{|p|/2} \langle v \rangle^p \int_{\mathbb{R}^d} (1 + \epsilon^2 |w|^2)^{|p|/2} G(w) \, dw$$
$$\leq C_{|p|} \langle v \rangle^p \left[1 + \epsilon^{|p|} \int_{\mathbb{R}^d} |w|^{|p|} G(w) \, dw \right] \leq C_{|p|} [1 + \epsilon^{|p|} m_{|p|}(G)] \langle v \rangle^p. \quad \square$$

We stress that Peetre's inequality in Lemma 44 is necessary for the estimate of Lemma 45 with *nonpositive* powers p which we apply in the sequel. Finally, the last result we will need is an integration by parts formula for the differential operator associated to the cross Fisher information.

Lemma 46 (twisted integration by parts). *Let f, g be smooth scalar functions of $\mathbb{R}^3$ which are sufficiently integrable. Then, we have the formula*

$$\int_{\mathbb{R}^3} (v \times \nabla_v g(v)) f(v) \, dv = -\int_{\mathbb{R}^3} g(v)(v \times \nabla_v f(v)) \, dv.$$

Here, the meaning of $v \times \nabla_v$ is

$$v \times \nabla_v f(v) = (v^2 \partial^3 f(v) - v^3 \partial^2 f(v), v^3 \partial^1 f(v) - v^1 \partial^3 f(v), v^1 \partial^2 f(v) - v^2 \partial^1 f(v)).$$

6.3. *Proof of* (32) *using Theorem 39.* We start by decomposing and estimating the integrand of D_ϵ. With the help of Lemma 42, we expand the square term of the integrand to see

$$\left| \tilde{\nabla} \left[\frac{\delta \mathcal{H}_\epsilon}{\delta \mu} \right] \right|^2$$

$$= |v - v_*|^{2+\gamma} |\Pi[v - v_*](b^\epsilon * a^\epsilon - b^\epsilon * a_*^\epsilon)|^2$$

$$\leq |v - v_*|^\gamma (4|v \times (b^\epsilon * a^\epsilon)|^2 + 4|v_* \times (b^\epsilon * a_*^\epsilon)|^2 + 4|v \times (b^\epsilon * a_*^\epsilon)|^2 + 4|v_* \times (b^\epsilon * a^\epsilon)|^2)$$

$$\leq 4|v - v_*|^\gamma \underbrace{|v \times (b^\epsilon * a^\epsilon)|^2}_{\textcircled{1}} + 4|v - v_*|^\gamma \underbrace{|v_* \times (b^\epsilon * a_*^\epsilon)|^2}_{\textcircled{2}} + 4|v|^2 |v - v_*|^\gamma \underbrace{|b^\epsilon * a_*^\epsilon|^2}_{\textcircled{3}} + 4|v_*|^2 |v - v_*|^\gamma \underbrace{|b^\epsilon * a^\epsilon|^2}_{\textcircled{4}},$$

where we use the shorthand notation

$$b^\epsilon = G^\epsilon \quad \text{and} \quad a^\epsilon = \nabla \log(G^\epsilon * f). \tag{34}$$

By using that G^ϵ is an approximation of the identity, we know that the integrand of D_ϵ converges pointwise a.e. to the integrand of D as $\epsilon \downarrow 0$. As well, each $\textcircled{i}$ for $i = 1, 2, 3, 4$ converges pointwise a.e. to

$$\textcircled{1} \to \frac{|v \times \nabla f|^2}{f^2}, \quad \textcircled{2} \to \frac{|v_* \times \nabla_* f_*|^2}{f_*^2}, \quad \textcircled{3} \to \frac{|\nabla_* f_*|^2}{f_*^2}, \quad \textcircled{4} \to \frac{|\nabla f|^2}{f^2}.$$

By Theorem 39, to show the integral convergence in (32), it suffices to show, for example,

$$\iint_{\mathbb{R}^6} f f_* |v - v_*|^\gamma \textcircled{1} \, dv \, dv_* \to \iint_{\mathbb{R}^6} f f_* |v - v_*|^\gamma \frac{|v \times \nabla f|^2}{f^2} \, dv \, dv_*,$$

and similarly for each $\textcircled{i}$ for $i = 2, 3, 4$. By symmetry considerations when swapping the variables $v \leftrightarrow v_*$, the convergence for the terms $\textcircled{1}$ and $\textcircled{4}$ controls the convergence for $\textcircled{2}$ and $\textcircled{3}$, respectively. Hence we will focus on the term $\textcircled{4}$ first and then on term $\textcircled{1}$.

6.3.1. *Term* $\textcircled{4}$. We seek to show in the limit $\epsilon \downarrow 0$

$$\iint_{\mathbb{R}^6} f f_* |v_*|^2 |v - v_*|^\gamma |b^\epsilon * a^\epsilon|^2 \, dv_* \, dv = \int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} f_* |v_*|^2 |v - v_*|^\gamma \, dv_* \right) f |b^\epsilon * a^\epsilon|^2 \, dv$$

$$\to \int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} f_* |v_*|^2 |v - v_*|^\gamma \, dv_* \right) \frac{|\nabla f|^2}{f} \, dv. \tag{35}$$

By the reordering of integrations written above, we now think of the double integral over v, v_* of $f f_* |v_*|^2 |v - v_*|^\gamma |b^\epsilon * a^\epsilon|^2$ as a single integral of the function $\left(\int_{\mathbb{R}^3} f_* |v_*|^2 |v - v_*|^\gamma \, dv_* \right) f |b^\epsilon * a^\epsilon|^2$ over v.

To be precise, we wish to apply Theorem 39 with $X = \mathbb{R}^3$ with $H_\epsilon = \left(\int_{\mathbb{R}^3} f_* |v_*|^2 |v - v_*|^\gamma \, dv_* \right) f |b^\epsilon * a^\epsilon|^2$. We can use Cauchy–Schwarz on the convolution integral to absorb the power term as follows:

$$|b^\epsilon * a^\epsilon|^2 = \left| \int_{\mathbb{R}^3} b^\epsilon(v-w) a^\epsilon(w) \, dw \right|^2 \le \left(\int_{\mathbb{R}^3} \langle w \rangle^{-\gamma} b^\epsilon(v-w) \, dw \right) \left(\int_{\mathbb{R}^3} b^\epsilon(v-w) \langle w \rangle^\gamma |a^\epsilon(w)|^2 \, dw \right)$$

$$\le C \langle v \rangle^{-\gamma} b^\epsilon * [\langle \cdot \rangle^\gamma |a^\epsilon(\cdot)|^2],$$

where the last inequality comes from Lemma 45. Continuing with Lemma 43, we have

$$\left(\int_{\mathbb{R}^3} f_* |v_*|^2 |v - v_*|^\gamma \, dv_* \right) f |b^\epsilon * a^\epsilon|^2 \le C f b^\epsilon * [\langle \cdot \rangle^\gamma |a^\epsilon|^2].$$

By Theorem 39, we reduce the problem to showing in the limit $\epsilon \downarrow 0$

$$\int_{\mathbb{R}^3} f b^\epsilon * [\langle \cdot \rangle^\gamma |a^\epsilon|^2] \, dv \to \int_{\mathbb{R}^3} \langle v \rangle^\gamma \frac{|\nabla f|^2}{f} \, dv.$$

This is were we use SACRE, Step 2 of our general strategy in Section 6.1. Application of SACRE and further simplification using the specific forms of a^ϵ and b^ϵ (see (34)) yields

$$\int_{\mathbb{R}^3} f b^\epsilon * [\langle \cdot \rangle^\gamma |a^\epsilon|^2] \, dv = \int_{\mathbb{R}^3} [b^\epsilon * f] \langle v \rangle^\gamma |a^\epsilon|^2 \, dv = \int_{\mathbb{R}^3} \langle v \rangle^\gamma \frac{|b^\epsilon * \nabla f|^2}{b^\epsilon * f} \, dv. \tag{36}$$

We work with this simplified expression and note that pointwise convergence is still valid

$$\frac{|b^\epsilon * \nabla f|^2}{b^\epsilon * f} \to \frac{|\nabla f|^2}{f}.$$

Next, we notice that the function $\beta : (F, f) \mapsto |F|^2/f$ is jointly convex in $F \in \mathbb{R}^3$ and $f > 0$, so we can use Jensen's inequality with $b^\epsilon = G^\epsilon$ as the reference probability measure to obtain a further pointwise majorant for the integrand of (36)

$$\frac{|b^\epsilon * \nabla f|^2}{b^\epsilon * f}(v) = \beta(b^\epsilon * \nabla f, b^\epsilon * f)(v) = \beta \left(\int_{\mathbb{R}^3} \nabla f(v-y) b^\epsilon(y) \, dy, \int_{\mathbb{R}^3} f(v-y) b^\epsilon(y) \, dy \right)$$

$$\le \int_{\mathbb{R}^3} \beta(\nabla f(v-y), f(v-y)) b^\epsilon(y) \, dy = \int_{\mathbb{R}^3} \frac{|\nabla f(v-y)|^2}{f(v-y)} b^\epsilon(y) \, dy = b^\epsilon * \left[\frac{|\nabla f|^2}{f} \right](v).$$

Using Theorem 39 again, we reduce the problem to showing in the limit $\epsilon \downarrow 0$

$$\int_{\mathbb{R}^3} \langle v \rangle^\gamma b^\epsilon * \left[\frac{|\nabla f|^2}{f} \right] dv \to \int_{\mathbb{R}^3} \langle v \rangle^\gamma \frac{|\nabla f|^2}{f} \, dv.$$

We use SACRE once more and place the convolution onto the weight term

$$\int_{\mathbb{R}^3} \langle v \rangle^\gamma b^\epsilon * \left[\frac{|\nabla f|^2}{f} \right] dv = \int_{\mathbb{R}^3} [b^\epsilon * \langle \cdot \rangle^\gamma] \frac{|\nabla f|^2}{f} \, dv.$$

Now, we are in a position to apply the classical dominated convergence theorem. We notice that we have the pointwise convergence

$$[b^\epsilon * \langle \cdot \rangle^\gamma] \to \langle v \rangle^\gamma.$$

Furthermore, using Lemma 45, we can estimate $b^\epsilon * \langle \cdot \rangle^\gamma$ uniformly in ϵ to find the domination

$$[b^\epsilon * \langle \cdot \rangle^\gamma] \frac{|\nabla f|^2}{f} \leq C \langle v \rangle^\gamma \frac{|\nabla f|^2}{f}.$$

Using Theorem 41, the finite entropy-dissipation (A3), and uniformly bounded entropy (A2) (remember the constant in Theorem 41 depends also on bounds for the entropy), we know that the right-hand side belongs to L^1_v for a.e. $t \in (0, T)$. Therefore, for a.e. $t \in (0, T)$ the conditions of the dominated convergence theorem are satisfied so we have the integral convergence

$$\int_{\mathbb{R}^3} [b^\epsilon * \langle \cdot \rangle^\gamma] \frac{|\nabla f|^2}{f} \, dv \to \int_{\mathbb{R}^3} \langle v \rangle^\gamma \frac{|\nabla f|^2}{f} \, dv.$$

We have closed the argument for the convergence of (35) after retracing the previous estimates with Theorem 39.

6.3.2. *Term* ①. We seek to show in the limit $\epsilon \downarrow 0$

$$\iint_{\mathbb{R}^6} f f_* |v - v_*|^\gamma |v \times (b^\epsilon * a^\epsilon)|^2 \, dv_* \, dv = \int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} f_* |v - v_*|^\gamma \, dv_* \right) f |v \times (b^\epsilon * a^\epsilon)|^2 \, dv$$

$$\to \int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} f_* |v - v_*|^\gamma \, dv_* \right) \frac{|v \times \nabla f|^2}{f} \, dv \qquad (37)$$

using the same strategy of nested applications of Theorem 39 like in Section 6.3.1. We will encounter difficulty when trying to use Jensen's inequality due to the cross Fisher information term. As in Section 6.3.1, we have written this double integral over v, v_* as a single integral over v. By Theorem 39 and Lemma 43, it suffices to show the integral convergence of

$$\int_{\mathbb{R}^3} \langle v \rangle^\gamma f |v \times (b^\epsilon * a^\epsilon)|^2 \, dv \to \int_{\mathbb{R}^3} \langle v \rangle^\gamma \frac{|v \times \nabla f|^2}{f} \qquad (38)$$

to obtain the integral convergence of (37). Pointwise, we can make the following manipulations:

$$v \times (b^\epsilon * a^\epsilon) = v \times \left(\int_{\mathbb{R}^3} G^\epsilon(v - w) \nabla \log(f * G^\epsilon(w)) \, dw \right) = v \times \left(\int_{\mathbb{R}^3} \nabla G^\epsilon(v - w) \log(f * G^\epsilon(w)) \, dw \right)$$

$$= \int_{\mathbb{R}^3} w \times \nabla G^\epsilon(v - w) \log(f * G^\epsilon(w)) \, dw = \int_{\mathbb{R}^3} G^\epsilon(v - w) w \times \nabla \log(f * G^\epsilon(w)) \, dw, \quad (39)$$

where we have used the radial symmetry of G^ϵ to get the cancellation $(v - w) \times \nabla G^\epsilon(v - w) = 0$ and the twisted integration by parts Lemma 46 (we note that we do not pick up any signs in the integration by parts, as the variable w appears with a minus sign in the argument of G^ϵ).

We apply Cauchy–Schwarz, multiply and divide by $\langle w \rangle^\gamma$, and use Lemma 45 to obtain

$$|v \times (b^\epsilon * a^\epsilon)|^2 \leq \left(\int_{\mathbb{R}^3} G^\epsilon(v - w) \langle w \rangle^{-\gamma} \, dw \right) \left(\int_{\mathbb{R}^3} G^\epsilon(v - w) \langle w \rangle^\gamma \left| w \times \frac{\nabla f * G^\epsilon(w)}{f * G^\epsilon(w)} \right|^2 \, dw \right)$$

$$\lesssim_\gamma \langle v \rangle^{-\gamma} \left(\int_{\mathbb{R}^3} G^\epsilon(v - w) \langle w \rangle^\gamma \left| w \times \frac{\nabla f * G^\epsilon(w)}{f * G^\epsilon(w)} \right|^2 \, dw \right).$$

Remembering that this majorant holds pointwise on the integrand of (38), we multiply by $\langle v \rangle^\gamma f(v)$ and obtain

$$\langle v \rangle^\gamma f(v)|v \times (b^\epsilon * a^\epsilon)|^2 \lesssim f\left(\int_{\mathbb{R}^3} G^\epsilon(v-w)\langle w \rangle^\gamma \left| w \times \frac{\nabla f * G^\epsilon(w)}{f * G^\epsilon(w)} \right|^2 dw\right).$$

Now, we recognize a convolution inside the brackets. Hence, using SACRE we can rewrite

$$\int_{\mathbb{R}^3} f\left(\int_{\mathbb{R}^3} G^\epsilon(v-w)\langle w \rangle^\gamma \left| w \times \frac{\nabla f * G^\epsilon(w)}{f * G^\epsilon(w)} \right|^2 dw\right) dv = \int_{\mathbb{R}^3} \langle v \rangle^\gamma \frac{|v \times \nabla f * G^\epsilon(v)|^2}{f * G^\epsilon(v)} dv.$$

Using Theorem 39, we need to show the convergence of the right-hand side. Here, it is now possible to use Jensen's inequality after some more manipulations.

Claim 47.
$$\frac{|v \times \nabla f * G^\epsilon(v)|^2}{f * G^\epsilon(v)} \leq \int_{\mathbb{R}^3} G^\epsilon(v-w)\frac{|w \times \nabla f(w)|^2}{f(w)} dw. \tag{40}$$

Proof of Claim 47. We start by repeating an argument similar to (39). Using that G^ϵ is radially symmetric and the twisted integration by parts Lemma 46, we obtain

$$v \times \nabla f * G^\epsilon(v) = v \times \left(\int_{\mathbb{R}^3} \nabla G^\epsilon(v-w)f(w) dw\right)$$

$$= \int_{\mathbb{R}^3} w \times \nabla G^\epsilon(v-w)f(w) dw = \int_{\mathbb{R}^3} G^\epsilon(v-w)\overbrace{(w \times \nabla_w f(w))}^{=:F(w)} dw.$$

Therefore, since $\beta : (F, f) \mapsto |F|^2/f$ is jointly convex in $F \in \mathbb{R}^3$ and $f > 0$, we apply Jensen's inequality with G^ϵ as the reference probability measure to the left-hand side of (40) to see

$$\frac{|v \times \nabla f * G^\epsilon(v)|^2}{f * G^\epsilon(v)} = \frac{|F * G^\epsilon|^2}{f * G^\epsilon}(v) = \beta(G^\epsilon * F, G^\epsilon * f)(v)$$

$$= \beta\left(\int_{\mathbb{R}^3} F(v-w)G^\epsilon(w) dw, \int_{\mathbb{R}^3} f(v-w)G^\epsilon(w) dw\right)$$

$$\leq \int_{\mathbb{R}^3} \beta(F(v-w), f(v-w))G^\epsilon(w) dw = \int_{\mathbb{R}^3} \frac{|(v-w) \times \nabla F(v-w)|^2}{f(v-w)} G^\epsilon(w) dw,$$

which proves the claim. $\qquad\square$

Continuing, by Theorem 39, we seek to establish the integral convergence of

$$\int_{\mathbb{R}^3} \langle v \rangle^\gamma \left[\frac{|F|^2}{f} * G^\epsilon\right](v) dv = \int_{\mathbb{R}^3} [\langle \cdot \rangle^\gamma * G^\epsilon](v)\frac{|v \times \nabla f(v)|^2}{f(v)} dv.$$

Finally, the integrand of the right-hand side has a majorant due to Lemma 45

$$[\langle \cdot \rangle^\gamma * G^\epsilon](v)\frac{|v \times \nabla f(v)|^2}{f(v)} \lesssim \langle v \rangle^\gamma \frac{|v \times \nabla f(v)|^2}{f(v)}.$$

Once again, using Theorem 41 and Assumptions (A3) and (A2), we obtain that for a.e. $t \in (0, T)$ the right-hand side belongs to $L_v^1(\mathbb{R}^3)$. Using dominated convergence theorem, we see that the integral converges. Tracing back the estimates, this takes care of the convergence of the term ① and establishes the convergence in (38).

We note that the estimates in the previous subsections not only establish the a.e. pointwise convergence of (32), but also the majorization

$$\iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left| \tilde{\nabla} \left[\frac{\delta \mathcal{H}_\epsilon}{\delta \mu} \right] \right|^2 dv_* \, dv \le C \iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left| \tilde{\nabla} \left[\frac{\delta \mathcal{H}}{\delta \mu} \right] \right|^2 dv_* \, dv \quad \text{for a.e. } t, \text{ for all } \epsilon > 0,$$

where

$$C \lesssim \| \langle \cdot \rangle^{-\gamma} f(t) \|_{L^\infty(0,T;L^1 \cap L^{(3-\eta)/(3+\gamma-\eta)}(\mathbb{R}^3))} + \| \langle \cdot \rangle^{2-\gamma} f(t) \|_{L^\infty(0,T;L^1 \cap L^{(3-\eta)/(3+\gamma-\eta)}(\mathbb{R}^3))}$$

by Lemma 43. Hence, using (A3) and (32) we can apply Lebesgue DCT to pass to the limit in the time integral and show the desired chain rule Claim 38.

Appendix: An auxiliary PDE for Lemma 35

In this section, we fix $\epsilon > 0$ throughout and study weak solutions to the PDE

$$\begin{cases} \partial_t \mu = \nabla \cdot \{ \mu \phi_{R_1} \int_{\mathbb{R}^d} \phi_{R_1*} \psi_{R_2}(v - v_*) |v - v_*|^{\gamma+2} \Pi[v - v_*] (J_0^\epsilon - J_{0*}^\epsilon) \, d\mu(v_*) \}, \\ \mu(0) = \mu_0. \end{cases} \tag{41}$$

We assume the initial data μ_0 belongs to $\mathscr{P}_2(\mathbb{R}^d)$. For $R_1, R_2 > 0$, the functions $0 \le \phi_{R_1}, \psi_{R_2} \le 1$ are smooth cut-off functions used to approximate the identity function in different ways:

$$\phi_{R_1}(v) = \begin{cases} 1, & |v| \le R_1, \\ 0, & |v| \ge R_1 + 1, \end{cases} \qquad \psi_{R_2}(z) = \begin{cases} 0, & |z| \le 1/R_2, \\ 1, & |z| \ge 2/R_2. \end{cases}$$

For $\epsilon > 0$, J_0^ϵ is the gradient of first variation of $\mathcal{H}_\epsilon$ applied to μ_0, meaning

$$J_0^\epsilon = \nabla G^\epsilon * \log[\mu_0 * G^\epsilon] \in C^\infty(\mathbb{R}^d; \mathbb{R}^d).$$

The main result of this section is:

Theorem 48. *Fix $\epsilon, R_1, R_2 > 0$, $\gamma \in \mathbb{R}$, and $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$. Then, there is a global unique weak solution $\mu \in C([0, +\infty); \mathscr{P}_2(\mathbb{R}^d))$ to (41).*

By Lemma 31, we know that J_0^ϵ is uniformly bounded (with constant depending on ϵ and μ_0 only through bounds on its second moment). The purpose of ϕ_{R_1}, ϕ_{R_1*} is to cut off the growth of $J_0^\epsilon, J_{0*}^\epsilon$ to ensure that the "velocity field" in the right-hand side of (41) is globally Lipschitz (it is, in fact, smooth and compactly supported). The $\psi_{R_2}(v - v_*)$-term avoids the possible singularities coming from the weight $|v - v_*|^{\gamma+2}$ for soft potentials $\gamma < 0$.

The construction of the solution in Theorem 48 is given in two steps. Firstly, a local well-posedness theory established to some finite time interval $T > 0$ which depends on $\epsilon, \gamma, R_1, R_2$ and μ_0. Secondly, the time of existence (and uniqueness) is extended to $+\infty$ since T depends on μ_0 only through its second moment, which is conserved by the evolution of (41).

We fix $T > 0$ to be determined explicitly later. Our strategy is to employ a fixed-point argument in the space $C([0, T]; \mathscr{P}_2(\mathbb{R}^d))$, which we will equip with the metric

$$d(\mu, \nu) := \sup_{t \in [0, T]} W_2(\mu(t), \nu(t)), \quad \mu, \nu \in C([0, T]; \mathscr{P}_2(\mathbb{R}^d)),$$

where W_2 is the 2-Wasserstein distance on $\mathscr{P}_2(\mathbb{R}^d)$. We have closely followed the procedure in [Cañizo et al. 2011] with appropriate modifications for this setting.

Remark 49. Since we are cutting off the "velocity" field at radius R_1, R_2, the growth of J_0^ϵ is inconsequential. Hence the results of this section can be applied when replacing the convolution kernel of J_0^ϵ with general tailed exponential distributions $G^{s,\epsilon}(v)$ for $s > 0$.

For $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, we will denote by $U[\mu](v)$ the function

$$U[\mu](v) := -\phi_{R_1} \int_{\mathbb{R}^d} \phi_{R_1 *} \psi_{R_2}(v - v_*)|v - v_*|^{\gamma+2} \Pi[v - v_*](J_0^\epsilon - J_{0*}^\epsilon) \, d\mu(v_*),$$

so that the PDE in (41) can be written as a nonlinear transport/continuity equation:

$$\partial_t \mu(t) = -\nabla \cdot \{\mu(t)U[\mu(t)]\}.$$

To fix ideas, the weak formulation of (41) is such that the following equality holds for all test functions $\tau \in C_c^\infty(\mathbb{R}^d)$ and times $t \in [0, T]$

$$\int_{\mathbb{R}^d} \tau(v) \, d\mu_r(v) - \int_{\mathbb{R}^d} \tau(v) \, d\mu_0(v)$$
$$= \int_0^t \int_{\mathbb{R}^d} \phi_{R_1} \nabla \tau(v) \cdot \int_{\mathbb{R}^d} \phi_{R_1 *} \psi_{R_2}(v - v_*)|v - v_*|^{\gamma+2} \Pi[v - v_*](J_0^\epsilon - J_{0*}^\epsilon) d\mu_s(v_*) \, d\mu_s(v) \, ds.$$

Thanks to all the smooth cutoffs from ϕ_{R_1}, $\phi_{R_1 *}$, and ψ_{R_2} and $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$, we can enlarge the class of test functions to smooth functions with quadratic growth. In particular, by choosing $\tau(v) = |v|^2$ and symmetrizing the right-hand side by swapping $v \leftrightarrow v_*$, we see that the second moment of μ_0 is conserved along the evolution of (41).

Our first step is to look at the level of the characteristic equation associated to (41).

Lemma 50 (characteristic equation). *For any $T > 0$, $\mu \in C([0, T]; \mathscr{P}_2(\mathbb{R}^d))$ and $v_0 \in \mathbb{R}^d$, there exists a unique solution $v \in C^1((0, T); \mathbb{R}^d) \cap C([0, T]; \mathbb{R}^d)$ to the ODE*

$$\frac{dv}{dt} = U[\mu(t)](v), \quad v(0) = v_0.$$

Furthermore, the growth rate satisfies

$$|v(t)| \leq \max\{|v_0|, R_1 + 1\} \quad \text{for all } t \in [0, T].$$

Proof. $U[\mu(t)](\cdot)$ is smooth and compactly supported uniformly in t, so classical Cauchy–Lipschitz theory gives existence and uniqueness of solution v with the promised regularity.

For the estimate on the growth rate, note that $U[\mu]$ has support contained in B_{R_1+1}. Points outside this ball do not change in time according to this ODE. $\square$

We will denote by Φ_μ^t the flow map associated to this ODE, so that

$$\frac{d}{dt} \Phi_\mu^t(v_0) = U[\mu(t)](\Phi_\mu^t(v_0)), \quad \Phi_\mu^0(v_0) = v_0.$$

It is known that, given $v \in C([0, T]; \mathscr{P}_2(\mathbb{R}^d))$, the curve of probability measures $\mu(t) = \Phi_v^t \# \mu_0$ is a weak solution to

$$\partial_t \mu(t) = -\nabla \cdot \{\mu(t) U[v(t)]\}, \quad \mu(0) = \mu_0.$$

Here, $\Phi_v^t \# \mu_0$ is the push-forward measure of μ_0 defined in duality with $\tau \in C_b(\mathbb{R}^d)$ by

$$\int_{\mathbb{R}^d} \tau(v) \, d(\Phi_v^t \# \mu_0)(v) = \int_{\mathbb{R}^d} \tau(\Phi_v^t(v)) \, d\mu_0(v).$$

We seek to find a fixed point to the map $\mu \mapsto \Phi_\mu^t \# \mu_0$ as it would weakly solve (41). To better understand the properties of this map, we need to establish estimates on the flow map through U as a function of time and measures.

Lemma 51 (L^∞ estimate for velocity field). *There exists a constant $C = C(\epsilon, \gamma, R_1, R_2, \mu_0) > 0$ such that, for every $T > 0$ and $v \in C([0, T]; \mathscr{P}_2(\mathbb{R}^d))$, we have*

$$|U[v(t)](v)| \leq C \quad \text{for all } t \in [0, T], v \in \mathbb{R}^d.$$

Proof. Estimate for $\gamma \geq -2$: We have the three inequalities

$$|v - v_*|^{\gamma+2} \lesssim_\gamma |v|^{\gamma+2} + |v_*|^{\gamma+2}, \quad \|\Pi[v - v_*]\| \leq 1, \quad J_0^\epsilon \lesssim_{\epsilon, \mu_0} 1$$

due to the range of γ, boundedness of Π, and Lemma 31, respectively. These three inequalities provide the estimate

$$|U[v(t)](v)| \lesssim_{\gamma, \epsilon, \mu_0} \phi_{R_1}(v) \int_{\mathbb{R}^d} \phi_{R_1}(v_*)(|v|^{\gamma+2} + |v_*|^{\gamma+2}) \, dv_t(v_*),$$

where we have dropped ψ_{R_2} altogether. For the integral term, we apply Hölder's inequality taking advantage of the compact support of ϕ_{R_1} and the unit mass of v_t to further obtain

$$|U[v(t)](v)| \lesssim_{\gamma, \epsilon, \mu_0} \phi_{R_1}(v)(R_1^{2+\gamma} + \langle v \rangle^{2+\gamma}) \int_{\mathbb{R}^d} dv_t(v_*) \lesssim_{R_1} \phi_{R_1}(v) \langle v \rangle^{2+\gamma}.$$

Again, since ϕ_{R_1} has compact support, we can brutally estimate the polynomial to conclude.

Estimate for $\gamma < -2$: Unlike the previous case, we change one of the inequalities due to the unavailability of a triangle inequality and use

$$\psi_{R_2}(v - v_*)|v - v_*|^{\gamma+2} \lesssim 1/R_2^{\gamma+2}, \quad \|\Pi[v - v_*]\| \leq 1, \quad J_0^\epsilon \lesssim_{\epsilon, \mu_0} 1.$$

From these inequalities and the compact support of ϕ_{R_1}, we have

$$|U[v(t)](v)| \lesssim_{\gamma, \epsilon, \mu_0, R_2} \phi_{R_1}(v) \int_{\mathbb{R}^d} \phi_{R_1}(v_*) \, dv_t(v_*) \leq 1. \qquad \square$$

The next result follows exactly as in [Cañizo et al. 2011].

Lemma 52 (time continuity of flow map). *Let $C = C(\epsilon, \gamma, R_1, R_2, \mu_0) > 0$ be the same constant from Lemma 51. Then, for any $T > 0$, and $v \in C([0, T]; \mathscr{P}_2(\mathbb{R}^d))$ we have*

$$\|\Phi_v^t - \Phi_v^s\|_{L^\infty(\mathbb{R}^d)} \leq C|t - s|.$$

Our next objective is to establish the regularity of the flow map with respect to the measures in the subscript. To simplify the subsequent lemmas, let us use the notation in the following:

Lemma 53. *Define*

$$F : (v, w) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto \phi_{R_1}(v)\phi_{R_1}(w)\psi_{R_2}(v - w)|v - w|^{\gamma+2}\Pi[v - w](J_0^\epsilon(v) - J_0^\epsilon(w)).$$

The function F is smooth and compactly supported. In particular, for every $k, l \in \mathbb{N}$, there is a constant $C = C(\epsilon, \gamma, R_1, R_2, \mu_0, k, l) > 0$ such that

$$\|D_v^k D_w^l F\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \leq C.$$

More precisely, the constant C depends on μ_0 only through bounds on its second moment as in Lemma 31.

Proof. The compact support property comes from the factor of $\phi_{R_1}(v)\phi_{R_1}(w)$ in the definition. The regularity comes from the avoidance of $v = w$ due to the factor $\psi_{R_2}(v - w)$. $\square$

Corollary 54 (pointwise and measurewise regularity of U). *Consider the constant C from Lemma 53 above. We have the following:*

(1) *Take $C_1 = C(\epsilon, \gamma, R_1, R_2, \mu_0, 0, 1) > 0$. For every $T > 0$, $v^1, v^2 \in C([0, T]; \mathscr{P}_2(\mathbb{R}^d))$, $t \in [0, T]$, $v \in \mathbb{R}^d$ we have the estimate*

$$|U[v^1(t)](v) - U[v^2(t)](v)| \leq C_1 W_2(v_t^1, v_t^2).$$

(2) *Take $C_2 = C(\epsilon, \gamma, R_1, R_2, \mu_0, 1, 0) > 0$. For every $T > 0$, $v \in C([0, T]; \mathscr{P}_2(\mathbb{R}^d))$, $t \in [0, T]$, $v_1, v_2 \in \mathbb{R}^d$ we have the estimate*

$$|U[v(t)](v_1) - U[v(t)](v_2)| \leq C_2 |v_1 - v_2|.$$

Remark 55. By considering the antisymmetric property of F when swapping variables $v \leftrightarrow w$, one really obtains $C_1 = C_2$. Their distinction in this corollary is artificial.

Proof. (1) Firstly, for every $t \in [0, T]$ take $\pi(t) \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$, the 2-Wasserstein optimal transportation plan connecting $v^1(t)$ and $v^2(t)$ which exists; see [Villani 2009]. We estimate the difference with notation from Lemma 53:

$$
\begin{aligned}
|U[v^1(t)](v) - U[v^2(t)](v)| &= \left| \int_{\mathbb{R}^d} F(v, w)\, dv_t^1(w) - \int_{\mathbb{R}^d} F(v, \bar{w})\, dv_t^2(\bar{w}) \right| \\
&= \left| \iint_{\mathbb{R}^{2d}} F(v, w) - F(v, \bar{w}) d\pi_t(w, \bar{w}) \right| \\
&\leq C_1 \iint_{\mathbb{R}^{2d}} |w - \bar{w}| d\pi_t(w, \bar{w}) \leq C_1 W_2(v_t^1, v_t^2).
\end{aligned}
$$

The first inequality uses a mean-value-type estimate (in the second variable of F) and the second inequality uses Cauchy–Schwarz, or equivalently, that W_2 is stronger than W_1.

(2) As with item (1), we estimate the difference using F to find

$$|U[\nu(t)](v_1) - U[\nu(t)](v_2)| = \left| \int_{\mathbb{R}^d} F(v_1, w) - F(v_2, w) \, d\nu_t(w) \right|$$

$$\leq \int_{\mathbb{R}^d} |F(v_1, w) - F(v_2, w)| \, d\nu_t(w) \leq C_2 |v_1 - v_2|.$$

Once more, a mean-value-type estimate is applied (in the first variable of F) and we recall ν_t is a probability measure. $\qquad\square$

The next result combines both items of Corollary 54 to estimate the regularity of the flow map with respect to measures and follows exactly as in [Cañizo et al. 2011].

Lemma 56 (continuity of flow map with respect to measures). *For $T > 0$ fix any $\nu^1, \nu^2 \in C([0, T]; \mathscr{P}_2(\mathbb{R}^d))$ and $t \in [0, T]$. With $C := C_1 = C_2$ the same constants in Corollary 54, we have the estimate*

$$\|\Phi_{\nu^1}^t - \Phi_{\nu^2}^t\|_{L^\infty(\mathbb{R}^d)} \leq (e^{Ct} - 1) d(\nu^1, \nu^2),$$

recalling that $d(\nu^1, \nu^2) = \sup_{t \in [0,T]} W_2(\nu_t^1, \nu_t^2)$.

It is by now classical how to obtain Theorem 48 from Corollary 54 and Lemma 56; see [Cañizo et al. 2011; Carrillo et al. 2014; Golse 2016] for instance. The time of existence can be given by any $0 < T < (1/C) \log 2$, where $C > 0$ is chosen as in Lemma 56 and the result follows by a fixed-point argument. The extension to all times is owed to the fact that $C > 0$ depends on the initial data μ_0 only through its second moment. This quantity is conserved through by the evolution of (41) and so the maximal time of existence is $+\infty$.

Acknowledgements

References

[Alexandre et al. 2015] R. Alexandre, J. Liao, and C. Lin, "Some a priori estimates for the homogeneous Landau equation with soft potentials", *Kinet. Relat. Models* **8**:4 (2015), 617–650. MR Zbl

[Ambrosio 1995] L. Ambrosio, "Movimenti minimizzanti", *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl.* (5) **19** (1995), 191–246. MR Zbl

[Ambrosio et al. 2008] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Birkhäuser, Basel, 2008. MR Zbl

[Arsenev and Peskov 1977] A. A. Arsenev and N. V. Peskov, "The existence of a generalized solution of Landau's equation", *Zh. Vychisl. Mat. Mat. Fiz.* **17**:4 (1977), 1063–1068. In Russian; translated in *USSR Comput. Math. and Math. Phys.* **17**:4 (1977), 241–246. MR Zbl

[Barros-Neto 1973] J. Barros-Neto, *An introduction to the theory of distributions*, Pure Appl. Math. **14**, Dekker, New York, 1973. MR Zbl

[Benamou and Brenier 2000] J.-D. Benamou and Y. Brenier, "A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem", *Numer. Math.* **84**:3 (2000), 375–393. MR Zbl

[Blanchet et al. 2008] A. Blanchet, V. Calvez, and J. A. Carrillo, "Convergence of the mass-transport steepest descent scheme for the subcritical Patlak–Keller–Segel model", *SIAM J. Numer. Anal.* **46**:2 (2008), 691–721. MR Zbl

[Boblylev et al. 2013] A. V. Boblylev, M. Pulvirenti, and C. Saffirio, "From particle systems to the Landau equation: a consistency result", *Comm. Math. Phys.* **319**:3 (2013), 683–702. MR Zbl

[Buttazzo 1989] G. Buttazzo, *Semicontinuity, relaxation and integral representation in the calculus of variations*, Pitman Res. Notes Math. Ser. **207**, Longman Sci. Tech., Harlow, England, 1989. MR Zbl

[Cañizo et al. 2011] J. A. Cañizo, J. A. Carrillo, and J. Rosado, "A well-posedness theory in measures for some kinetic models of collective motion", *Math. Models Methods Appl. Sci.* **21**:3 (2011), 515–539. MR Zbl

[Carlen and Carvalho 1992] E. A. Carlen and M. C. Carvalho, "Strict entropy production bounds and stability of the rate of convergence to equilibrium for the Boltzmann equation", *J. Statist. Phys.* **67**:3-4 (1992), 575–608. MR Zbl

[Carrapatoso and Mischler 2017] K. Carrapatoso and S. Mischler, "Landau equation for very soft and Coulomb potentials near Maxwellians", *Ann. PDE* **3**:1 (2017), art. id. 1. MR Zbl

[Carrapatoso et al. 2017] K. Carrapatoso, L. Desvillettes, and L. He, "Estimates for the large time behavior of the Landau equation in the Coulomb case", *Arch. Ration. Mech. Anal.* **224**:2 (2017), 381–420. MR Zbl

[Carrillo et al. 2010] J. A. Carrillo, S. Lisini, G. Savaré, and D. Slepčev, "Nonlinear mobility continuity equations and generalized displacement convexity", *J. Funct. Anal.* **258**:4 (2010), 1273–1309. MR Zbl

[Carrillo et al. 2012] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepčev, "Confinement in nonlocal interaction equations", *Nonlinear Anal.* **75**:2 (2012), 550–558. MR Zbl

[Carrillo et al. 2014] J. A. Carrillo, Y.-P. Choi, and M. Hauray, "The derivation of swarming models: mean-field limit and Wasserstein distances", pp. 1–46 in *Collective dynamics from bacteria to crowds*, edited by A. Muntean and F. Toschi, CISM Courses and Lect. **553**, Springer, 2014. MR

[Carrillo et al. 2019] J. A. Carrillo, K. Craig, and F. S. Patacchini, "A blob method for diffusion", *Calc. Var. Partial Differential Equations* **58**:2 (2019), art. id. 53. MR Zbl

[Carrillo et al. 2020] J. A. Carrillo, J. Hu, L. Wang, and J. Wu, "A particle method for the homogeneous Landau equation", *J. Comput. Phys. X* **7** (2020), art. id. 100066. MR

[Carrillo et al. 2022] J. A. Carrillo, M. G. Delgadino, and J. Wu, "Boltzmann to Landau from the gradient flow perspective", *Nonlinear Anal.* **219** (2022), art. id. 112824. MR Zbl

[Degond and Lucquin-Desreux 1992] P. Degond and B. Lucquin-Desreux, "The Fokker–Planck asymptotics of the Boltzmann collision operator in the Coulomb case", *Math. Models Methods Appl. Sci.* **2**:2 (1992), 167–182. MR Zbl

[Desvillettes 2015] L. Desvillettes, "Entropy dissipation estimates for the Landau equation in the Coulomb case and applications", *J. Funct. Anal.* **269**:5 (2015), 1359–1403. MR Zbl

[Desvillettes 2016] L. Desvillettes, "Entropy dissipation estimates for the Landau equation: general cross sections", pp. 121–143 in *From particle systems to partial differential equations*, *III* (Braga, Portugal, 2014), edited by P. Gonçalves and A. J. Soares, Springer Proc. Math. Stat. **162**, Springer, 2016. MR Zbl

[Desvillettes 2022] L. Desvillettes, "Autour du thèorème *H* de Boltzmann", art. id. 9 in *Séminaire Laurent Schwartz*: *EDP et applications* 2019/2020, Ècole Polytech., Palaiseau, France, 2022.

[Desvillettes and Fellner 2006] L. Desvillettes and K. Fellner, "Exponential decay toward equilibrium via entropy methods for reaction-diffusion equations", *J. Math. Anal. Appl.* **319**:1 (2006), 157–176. MR Zbl

[Desvillettes and Villani 2000a] L. Desvillettes and C. Villani, "On the spatially homogeneous Landau equation for hard potentials, I: Existence, uniqueness and smoothness", *Comm. Partial Differential Equations* **25**:1-2 (2000), 179–259. MR Zbl

[Desvillettes and Villani 2000b] L. Desvillettes and C. Villani, "On the spatially homogeneous Landau equation for hard potentials, II: *H*-theorem and applications", *Comm. Partial Differential Equations* **25**:1-2 (2000), 261–298. MR Zbl

[Desvillettes et al. 2023] L. Desvillettes, L.-B. He, and J.-C. Jiang, "A new monotonicity formula for the spatially homogeneous Landau equation with Coulomb potential and its applications", *J. Eur. Math. Soc.* (online publication January 2023).

[Dolbeault et al. 2009] J. Dolbeault, B. Nazaret, and G. Savaré, "A new class of transport distances between measures", *Calc. Var. Partial Differential Equations* **34**:2 (2009), 193–231. MR Zbl

[Erbar 2014] M. Erbar, "Gradient flows of the entropy for jump processes", *Ann. Inst. Henri Poincaré Probab. Stat.* **50**:3 (2014), 920–945. MR Zbl

[Erbar 2023] M. Erbar, "A gradient flow approach to the Boltzmann equation", *J. Eur. Math. Soc.* (online publication June 2023).

[Erbar and Maas 2014] M. Erbar and J. Maas, "Gradient flow structures for discrete porous medium equations", *Discrete Contin. Dyn. Syst.* **34**:4 (2014), 1355–1374. MR Zbl

[Fournier and Guérin 2009] N. Fournier and H. Guérin, "Well-posedness of the spatially homogeneous Landau equation for soft potentials", *J. Funct. Anal.* **256**:8 (2009), 2542–2560. MR Zbl

[Golse 2016] F. Golse, "On the dynamics of large particle systems in the mean field limit", pp. 1–144 in *Macroscopic and large scale phenomena*: *coarse graining*, *mean field limits and ergodicity* (Enschede, Netherlands, 2012), edited by A. Muntean et al., Lect. Notes Appl. Math. Mech. **3**, Springer, 2016. MR

[Golse et al. 2019a] F. Golse, M. P. Gualdani, C. Imbert, and A. Vasseur, "Partial regularity in time for the space homogeneous Landau equation with Coulomb potential", preprint, 2019. arXiv 1906.02841

[Golse et al. 2019b] F. Golse, C. Imbert, C. Mouhot, and A. F. Vasseur, "Harnack inequality for kinetic Fokker–Planck equations with rough coefficients and application to the Landau equation", *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **19**:1 (2019), 253–295. MR Zbl

[Gualdani and Guillen 2016] M. P. Gualdani and N. Guillen, "Estimates for radial solutions of the homogeneous Landau equation with Coulomb potential", *Anal. PDE* **9**:8 (2016), 1772–1809. MR Zbl

[Gualdani and Zamponi 2017] M. P. Gualdani and N. Zamponi, "Spectral gap and exponential convergence to equilibrium for a multi-species Landau system", *Bull. Sci. Math.* **141**:6 (2017), 509–538. MR Zbl

[Gualdani and Zamponi 2018a] M. Gualdani and N. Zamponi, "A review for an isotropic Landau model", pp. 115–144 in *PDE models for multi-agent phenomena* (Rome, 2016), edited by P. Cardaliaguet and F. Salvarani, Springer INdAM Ser. **28**, Springer, 2018. MR Zbl

[Gualdani and Zamponi 2018b] M. Gualdani and N. Zamponi, "Global existence of weak even solutions for an isotropic Landau equation with Coulomb potential", *SIAM J. Math. Anal.* **50**:4 (2018), 3676–3714. MR Zbl

[Guo 2002] Y. Guo, "The Landau equation in a periodic box", *Comm. Math. Phys.* **231**:3 (2002), 391–434. MR Zbl

[Jordan et al. 1998] R. Jordan, D. Kinderlehrer, and F. Otto, "The variational formulation of the Fokker–Planck equation", *SIAM J. Math. Anal.* **29**:1 (1998), 1–17. MR Zbl

[Lifshitz and Pitaevskiĭ 1981] E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical kinetics*, Course of Theoret. Phys. **10**, Pergamon, Oxford, 1981. MR Zbl

[Royden 1963] H. L. Royden, *Real analysis*, Macmillan, New York, 1963. MR Zbl

[Sandier and Serfaty 2004] E. Sandier and S. Serfaty, "Gamma-convergence of gradient flows with applications to Ginzburg–Landau", *Comm. Pure Appl. Math.* **57**:12 (2004), 1627–1672. MR Zbl

[Santambrogio 2017] F. Santambrogio, "(Euclidean, metric, and Wasserstein) gradient flows: an overview", *Bull. Math. Sci.* **7**:1 (2017), 87–154. MR Zbl

[Serfaty 2011] S. Serfaty, "Gamma-convergence of gradient flows on Hilbert and metric spaces and applications", *Discrete Contin. Dyn. Syst.* **31**:4 (2011), 1427–1451. MR Zbl

[Silvestre 2017] L. Silvestre, "Upper bounds for parabolic equations and the Landau equation", *J. Differential Equations* **262**:3 (2017), 3034–3055. MR Zbl

[Strain and Wang 2020] R. M. Strain and Z. Wang, "Uniqueness of bounded solutions for the homogeneous relativistic Landau equation with Coulomb interactions", *Quart. Appl. Math.* **78**:1 (2020), 107–145. MR Zbl

[Villani 1998a] C. Villani, "On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations", *Arch. Ration. Mech. Anal.* **143**:3 (1998), 273–307. MR Zbl

[Villani 1998b] C. Villani, "On the spatially homogeneous Landau equation for Maxwellian molecules", *Math. Models Methods Appl. Sci.* **8**:6 (1998), 957–983. MR Zbl

[Villani 2009] C. Villani, *Optimal transport*: *old and new*, Grundlehren der Math. Wissenschaften **338**, Springer, 2009. MR Zbl

[Wu 2014] K.-C. Wu, "Global in time estimates for the spatially homogeneous Landau equation with soft potentials", *J. Funct. Anal.* **266**:5 (2014), 3134–3155. MR Zbl

JOSÉ A. CARRILLO: carrillo@maths.ox.ac.uk
Mathematical Institute, University of Oxford, Oxford, United Kingdom

MATIAS G. DELGADINO: matias.delgadino@math.utexas.edu
Department of Mathematics, University of Texas, Austin, TX, United States

LAURENT DESVILLETTES: desvillettes@math.univ-paris-diderot.fr
Université Paris Cité and Sorbonne Université, CNRS, IUF, IMJ-PRG, Paris, France

JEREMY S.-H. WU: jeremywu@math.ucla.edu
Department of Mathematics, University of California, Los Angeles, CA, United States

msp

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

Analysis & PDE

Volume 17   No. 4   2024