Understanding the Landau Equation as a Gradient Flow

Jeremy Sheung-Him Wu 伍尚謙

The Queen's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2022

Dedication

I sincerely appreciate the many people in my life who have played an important role, personally and academically, during my Ph.D. years.

I am wholeheartedly grateful to my supervisors José A. Carrillo and Matias G. Delgadino for their helpful supervision, constant support, and friendly guidance. I truly appreciate their dedication in making my Ph.D. a uniquely enjoyable experience. Muchas gracias.

I am fortunate to have been included in an extended research family while at Imperial and Oxford. The kindness and acceptance I received helped me feel comfortable adjusting to the intimidating prospect of a research degree. Special thanks are due to my Ph.D. siblings Franca, Markus, Rafa, Rishabh, and Sergio.

I am deeply indebted to my friends for many reasons, not least of which is because of their companionship, encouragement, and optimism. I especially cherish all the fond experiences and memories I have shared with Annina, Asad, Diana, Catalina, Eliana, Estelle, Hans, Håvard, Ian, Lorenz, Tommaso, and Vicki. My interactions with you brighten my day and I hope for many more bright days.

Returning to my roots, I would also like to thank my family and friends from Canada. Thank you to my parents for raising me to become the man I am now. Thank you, Annabelle, for being the best sister I could have asked for; you helped shape my moral character. Thank you, Adam, Alex, and Alton; you were there for me when I needed you and I can never repay our priceless friendship and found brotherhood.

This thesis is dedicated to the people who have entered my life and brought joy along with them. I have learned so much from you and I strive every day to be worthy of your positive impact in my life.

Declaration of Originality

I declare that the contents of this thesis are, to the best of my knowledge, original and my own work, except where otherwise indicated, cited, or commonly known, nor has any part of this thesis been submitted for a degree at another university.

The work in Chapter 2 and Appendix B is joint work with José A. Carrillo, Matias G. Delgadino, and Laurent Desvillettes. It is a preprint [31] submitted for publication.

The work in Chapter 3 and Appendices A and C is joint work with José A. Carrillo and Matias G. Delgadino. It has been published [32] in Nonlinear Analysis volume 219, page 112824 in June 2022.

The work in Chapter 4 is joint work with José A. Carillo, Jingwei Hu, and Li Wang. It has been published [33] in the Journal of Computational Physics: X volume 7, page 100066 in June 2020.

Contents

Abstract

This thesis provides and investigates the rigorous gradient flow viewpoint of the spatially homogeneous Landau equation for soft potentials. This extends the similar perspective recently developed for the Boltzmann equation with Maxwellian potentials [61]. Taking advantage of the H-theorem for the Landau equation, we construct a metric from a dynamic optimal transportation viewpoint [17, 60] for which the Landau equation can be viewed as the gradient flow of the Boltzmann entropy with respect to this metric. The gradient flow description is further reinforced by recovering the grazing collision limit from the Boltzmann equation to the Landau equation with simpler arguments and intuition in the spirit of Γ-convergence [113]. Finally, a robust and efficient particle approximation is numerically analysed for a regularised version of the Landau equation which preserves the gradient flow structure.

This thesis invites and advertises the use of gradient flow techniques to pursue some of the future research directions discussed here.

List of notations and symbols

$a \lesssim b$ or $a = \mathcal{O}(b)$	$a \leq Cb$ for $a, b > 0$ where $C > 0$
$a \lesssim_{\alpha,\beta,\dots} b$	$a \leq Cb$ for $a, b > 0$ where $C = C(\alpha, \beta, \dots) > 0$
$a \sim b$	$a \lesssim b$ and $b \lesssim a$ for $a, b > 0$
$\Pi[z]$	The projection onto $\{z\}^\perp$ for $z \in \mathbb{R}^d$ given by $I - \frac{z \otimes z}{\lvert z\rvert^2}$
$\langle v \rangle$	Japanese angle bracket meaning $\sqrt{1 + \lvert v\rvert^2}$ for $v \in \mathbb{R}^d$
$\mathcal{L}$	Lebesgue measure (on $\mathbb{R}^d$ or whatever space context dictates)
$\mathscr{P}(\Omega)$	Space of probability measures over the set Ω
$C_b(\Omega)$	Continuous and bounded functions on Ω
$\mu_n \overset{\sigma}{\rightharpoonup} \mu$ or $\mu_n \rightharpoonup \mu$	$\mu_n \in \mathscr{P}(\Omega)$ converges to $\mu \in \mathscr{P}(\Omega)$ in duality with $C_b(\Omega)$
$\mathscr{P}$	Space of probability measures over full space (usually $\mathbb{R}^d$)
$\mathscr{P}^a$ and $\mathscr{P}^a(\Omega)$	Members of $\mathscr{P}$ and $\mathscr{P}(\Omega)$ absolutely continuous with respect to $\mathcal{L}$
$m_p(\mu)$	p-th moment of μ given by $\int_{\mathbb{R}^d} \langle v \rangle^p \, d\mu(v)$ for $\mu \in \mathscr{P}$
$\mathscr{P}_p$	Space consisting of all $\mu \in \mathscr{P}$ such that $m_p(\mu) < +\infty$
$\mathcal{M}(\Omega)$	Space of signed Radon measures over the set Ω
$C_0(\Omega)$	Continuous functions on Ω vanishing towards the boundary
$m_n \overset{*}{\rightharpoonup} m$	$m_n \in \mathcal{M}(\Omega)$ converges to $m \in \mathcal{M}(\Omega)$ in duality with $C_0(\Omega)$
$\mathcal{M}^d(\Omega)$	Space of $\mathbb{R}^d$-valued signed Radon measures over the set Ω
L^1_+	$\{f \in L^1(\mathbb{R}^d) \mid f \geq 0$ and $\int_{\mathbb{R}^d} \lvert f(v)\rvert dv = 1\}$
L^1_κ	$\{f \in L^1(\mathbb{R}^d) \mid \int_{\mathbb{R}^d} \lvert f(v)\rvert \langle v \rangle^\kappa \, dv < +\infty\}$
$\mu = f\mathcal{L}$	$\mu \in \mathscr{P}^a$ and its density with respect to $\mathcal{L}$ is $f \in L^1_+$
$\partial^i f$ or $\partial_{v^i} f$	Partial derivative of f with respect to v^i component
$(D^2 f)^{ij} = \partial^{ij} f$	(i,j)-component of Hessian matrix of f given by $\partial^{ij} f = \partial_{v^i} \partial_{v^j} f$

We will always use μ to denote a probability measure. Moreover, we will consider curves $\mu : t \in [0, T] \mapsto \mu(t) \in \mathscr{P}$ which may be labelled by any of the following; μ, μ_t, $\mu(t)$. Similarly, f will usually denote a function (in some L^p space) and curves $f : t \mapsto f(t)$ may be frequently represented by f, f_t, $f(t)$. Subscripts, in particular subscript t, will **never** refer to partial differentiation.

We define the Boltzmann entropy of $\mu \in \mathscr{P}$ by

$$\mathcal{H}[\mu] = \mathcal{H}[f] := \begin{cases} \int f \log f d\mathcal{L}, & \mu = f\mathcal{L} \\ +\infty, & \text{otherwise} \end{cases} .$$

Its first variation $\frac{\delta \mathcal{H}}{\delta f} = \log f$ is defined by the formal limit

$$\lim_{\eta \downarrow 0} \frac{\mathcal{H}[f + \eta\phi] - \mathcal{H}[f]}{\eta} = \int \frac{\delta \mathcal{H}}{\delta f}\phi, \quad \forall \phi \in C_c^\infty \text{ such that } \int \phi = 0.$$

Regularisations and scalings play a significant role. To avoid confusion, $\varepsilon > 0$ is reserved **exclusively for the content in Chapters 2 and 4 and appendix B** whereas $\epsilon > 0$ is reserved **exclusively for the content in Chapter 3 and appendix C.**

List of Figures

Chapter 1

Introduction

These humble pages contain the fruits of my Ph.D., which is dedicated to the study of the spatially homogeneous **Landau** equation

$$\partial_t f(t,v) = Q_L(f(t,\cdot), f(t,\cdot))(v), \tag{L}$$

where the collision operator is given, for $\gamma \in [-3,0]$ and $\Pi[z]$ the projection onto $\{z\}^\perp$, by

$$Q_L(f,f) := \nabla_v \cdot \left(f(v) \int_{\mathbb{R}^3} f(v_*)|v-v_*|^{2+\gamma}\Pi[v-v_*](\nabla \log f(v) - \nabla \log f(v_*))dv_* \right).$$

The unknown to be solved for, $f(t,v)$, corresponds to the probability of finding a particle in a plasma [95, 37] at time $t > 0$ and with velocity $v \in \mathbb{R}^3$. The non-local structure of Q_L accounts for the interactions from particles with velocities $v_* \in \mathbb{R}^3$. Plasma is the fourth state of matter, consisting primarily of charged particles which can be artificially generated by applying a strong electromagnetic field to a gas at extremely high temperatures. Some examples include stars, lightning, aurorae, fluorescent lights, fusion reactors, ... [82, 87]. The abundance of plasma in our universe motivates the study of (L) to better understand its physical properties. Moreover, plasma is the focus of the ambitious ITER project to produce net energy through fusion [88]. While (L) is the primary equation studied here, this thesis also considers the spatially homogeneous **Boltzmann** equation below, which is closely related and significant in kinetic theory and mathematical physics. For a given collision kernel B, the evolution of a density, f, of particles colliding with each other in a gas [115, 36] is given by

$$\partial_t f = Q_B(f,f) := \int_{\mathbb{R}^3}\int_{\mathbb{S}^2} [f'f'_* - ff_*]B(|v-v_*|,\theta)d\sigma dv_*. \tag{B}$$

I have suppressed the dependence on time and abbreviated $f = f(v)$, $f_* = f(v_*)$, $f' = f(v')$, $f'_* = f(v'_*)$ (this convention will persist), where the post-collision velocities are determined by

$$v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2}\sigma, \quad v'_* = \frac{v + v_*}{2} - \frac{|v - v_*|}{2}\sigma, \quad \sigma \in \mathbb{S}^2.$$

Intuitively, the non-local collision operator Q_B considers binary collisions of particles in a gas. Given two particles with pre-collision velocities v, $v_* \in \mathbb{R}^3$, the formulas for v' and v'_* specify the respective post-collision velocities of the two particles (c.f. Figure 1.1). The second argument of B represents the angle of collisions $\theta \in [0, \pi/2]$ given implicitly by

$$k = \frac{v - v_*}{|v - v_*|}, \quad \cos\theta = k \cdot \sigma.$$

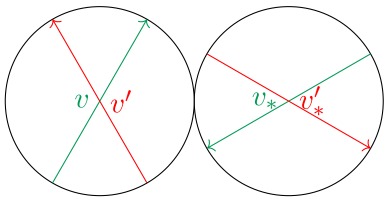


Figure 1.1: Hard sphere model of gas particles with green pre-collision velocities colliding and attaining red post-collision velocities.

Hilbert's famous sixth problem [84] suggests the importance of (B) to develop 'mathematically the limiting processes ... which lead from the atomistic view to the laws of motion of continua.' In formal terms, (B) acts as an intermediary between microscopic laws of physics (particle collisions) and macroscopic behaviour (the motion of fluids, e.g. phenomena described by the Navier-Stokes equations). The typical example [118, 2] of the Boltzmann kernel which is considered in this thesis is

$$B(|z|, \theta) = |z|^\gamma b(\theta), \quad \sin\theta\, b(\theta) =: \beta(\theta) \geq 0. \tag{1.1}$$

Besides the physical origins of (B), similar principles have inspired the use of kinetic models (choosing different collision kernels other than (1.1)) for opinion formulation [114], Elo-type models to compare player strength in games, such as chess [89], and general systems of interacting agents [105].

Crucial to the study of (L) and (B) is the evolution of the Boltzmann **entropy**

$$\mathcal{H}[f] = \int f(v) \log f(v) dv, \quad f \in L^1_+,$$

which I have defined here with the opposite sign conventionally used in physics. The quantity $-\mathcal{H}$ measures the volume of different microscopic configurations given fixed macroscopic observations [120].

This quantity also plays a similar role in Shannon's information theory as a measurement of uncertainty or redundancy [98, 42]. In plain terms, large values of $\mathcal{H}[f]$ correspond to order and total correlation of a system where f describes the probability of different microscopic states. Conversely, small values of $\mathcal{H}$ indicate disorder and uncertainty of a physical system. Supposing $f = f(t) \in L^1_+$ is a (smooth) solution to either (B) or (L), one has Boltzmann's famous **H-theorem** (formally proved in Section 1.1), which reads

$$\mathcal{H}[f(t_1)] \leq \mathcal{H}[f(t_2)], \quad \forall t_1 \geq t_2 > 0.$$

Physically, this result corresponds to the **second law of thermodynamics** and implies the **irreversibility** of time. Moreover, it suggests convergence of solutions f towards an equilibrium state which represents fully uncorrelated particles as time goes to infinity. Mathematically, the H-theorem is at the heart of [118] and forms the basis of our extension in Chapter 2. The Boltzmann and Landau equations are intimately related by the so-called **grazing collision limit** [51, 45]. This limiting process describes the derivation of (L) from (B) in which β concentrates on collisions with small angular deviation ($\theta \ll 1$). This corresponds to infinitesimally small differences between the pre- and post-collision velocities as Figure 1.2 below suggests.

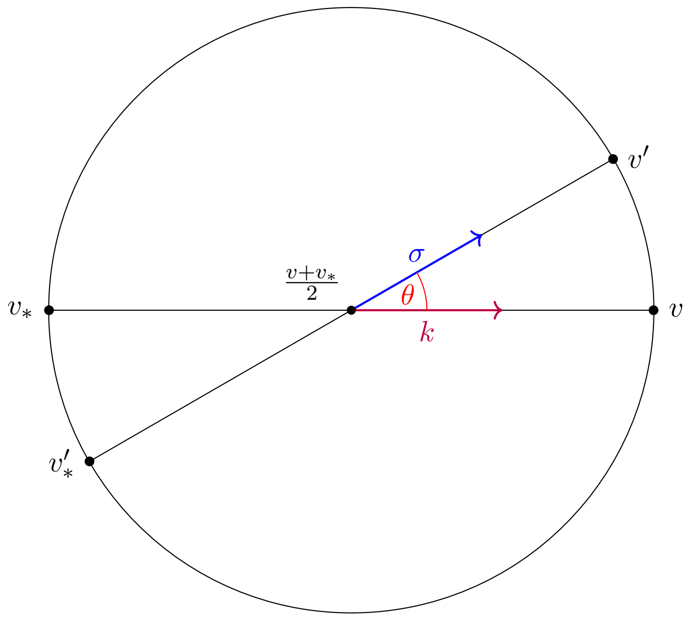


Figure 1.2: Geometry of elastic collisions in centre of momentum coordinates.

The grazing collision limit is studied in more detail later (c.f. Chapter 3), but I simply mention it for now to emphasise the connection between (B) and (L). In fact, Landau's original derivation [95] in 1936 of (L) served to replace (B) for $\gamma = -3$ corresponding to the Coulomb potential (owing to the highly singular kernel (1.1)), which is the most physically relevant parameter choice.

The singularity coming from $\gamma < 0$ and the degeneracy of Π are major sources of difficulty in the theory of (L) [78, 73, 69, 70, 65]. This thesis collects the various projects I have completed relating to (L) from the viewpoint of **gradient flows**, a popular and thriving area of mathematical research in recent times [6, 17, 90, 35, 111]. I invite readers to apply the sophisticated tools of gradient flows against the open problems of (L).

As the reader can already tell, I have written this thesis with more of a personal flair and self-expression than a typical academic paper. My narration freely alternates between the first person singular (mainly used here in the introduction and in Chapter 5) and plural (mainly used in Chapters 2 to 4 with the royal 'we') forms. This thesis is also organised in a somewhat chronological order. The format was motivated by my desire to write a coherent and pedagogical document. To be precise, the present introduction sets the necessary context for Chapters 2 to 4. Section 1.1 reviews some formal computations from the Boltzmann (B) and Landau (L) equations to develop the intuition motivating the rigorous arguments in the rest of this thesis. Section 1.2 reviews some elements from the extensive literature of gradient flow theory using a toy example in finite dimensions with a particular emphasis on the results applied and generalised here. Sections 1.3 to 1.5 recount the literature on the Boltzmann and Landau equations and serve as historical introductions for Chapters 2 to 4, respectively. Chapter 2 is the theoretical heart of this thesis and details the rigorous gradient flow description of (L). Chapter 3 revisits and simplifies the grazing collision limit which connects the Boltzmann equation to the Landau equation from the framework of gradient flows developed by Erbar [61] (for the Boltzmann equation) and Chapter 2 (for the Landau equation, see also [31]). Chapter 4 focuses on numerical experimentation to approximate the Landau equation by considering a regularisation of the Landau equation, which preserves its gradient flow structure. Finally, Chapter 5 summarises the results of this thesis together with future research questions I wish to pursue.

1.1 The blueprint

I now sketch some formal computations as motivation to set the stage for the rest of this thesis, in particular for Chapter 2. In the following, consider a (smooth) solution $f = f(t,v) \in L^1_+$ to (L). The H-theorem for (L) can be formally proven by using $\log f(t,v)$ as a test function

$$
\begin{aligned}
&\frac{d}{dt}\mathcal{H}[f(t,\cdot)] = \int_{\mathbb{R}^3} (\log f + 1)\partial_t f dv \\
&= -\int_{\mathbb{R}^3} f\nabla \log f \cdot \int_{\mathbb{R}^3} f_*|v-v_*|^{2+\gamma}\Pi[v-v_*](\nabla \log f - \nabla_* \log f_*)dv_* dv \\
&= -\frac{1}{2}\iint_{\mathbb{R}^6} ff_*|v-v_*|^{2+\gamma}(\nabla \log f - \nabla_* \log f_*)^T\Pi[v-v_*](\nabla \log f - \nabla_* \log f_*)dv_* dv.
\end{aligned}
$$

Integrating over time, this equality gives the so-called **entropy dissipation equality**

$$\mathcal{H}[f(0)] = \mathcal{H}[f(t)] + \int_0^t \underbrace{\left(\frac{1}{2}\iint_{\mathbb{R}^6} ff_*|v-v_*|^{2+\gamma}|\Pi[v-v_*](\nabla \log f - \nabla_* \log f_*)|^2 dv_* dv\right)}_{=:D_L(f)\geq 0} dt. \quad (EDE_L)$$

The positivity of the dissipation, D_L, comes from the fact that $\Pi[v-v_*]$ is a projection; the quadratic form is compressed by inserting an extra Π (it is an idempotent operator), since

$$z^T\Pi z = z^T\Pi^T\Pi z = (\Pi z)^T\Pi z = |\Pi z|^2 \geq 0, \quad \forall z \in \mathbb{R}^3.$$

Assuming $\mathcal{H}[f(0)] < +\infty$ to avoid pathological cases, (EDE_L) implies $\mathcal{H}[f(t)] \leq \mathcal{H}[f(0)]$, and the difference is given precisely by $\int_0^t D_L(f)dt < +\infty$. The lower bound of $\mathcal{H}$ comes from Jensen's inequality for the **relative entropy**

$$\mathcal{H}[f\,|\,M] := \int_{\mathbb{R}^3} \frac{f}{M}\log\left(\frac{f}{M}\right)Mdv, \quad M(v) = (2\pi)^{-\frac{3}{2}}\exp\left\{-\frac{|v|^2}{2}\right\},$$

recalling $x \mapsto x\log x$ is convex. My thesis illustrates the perspective of viewing functions $t \mapsto f(t)$ which satisfy (EDE_L) as gradient flow solutions to (L). More precisely, the goal here is to take (EDE_L) as the criteria for which the Landau equation is the steepest descent of $\mathcal{H}[f(t)]$ with respect to some metric. Identifying this metric is one of the major objectives of Chapter 2.

I have often looked to the heat equation as a helpful toy model to clarify the ideas discussed previously. Consider $u = u(t,x) \in L^1_+$ a solution to the heat equation in $\mathbb{R}^3$

$$\partial_t u = \Delta u. \tag{1.2}$$

There is a corresponding H-theorem for (1.2) to arrive at a simpler version of (EDE_L)

$$\mathcal{H}[u(0)] = \mathcal{H}[u(t)] + \int_0^t \underbrace{\int_{\mathbb{R}^3} u|\nabla \log u|^2 dx}_{=:D_h(u)}\, dt. \tag{1.3}$$

The entropy dissipation for the heat equation, D_h, is more commonly referred to as the Fisher information. It is a classical result [90] that (1.2) can be interpreted as the gradient flow of $\mathcal{H}$ with respect to the 2-Wasserstein optimal transportation distance W_2 on $\mathscr{P}_2$. In other words, solutions $u \in \mathscr{P}_2$ to (1.2) decrease $\mathcal{H}$ as fast as possible in the W_2 metric. To see how this example is related to (L), I drew inspiration from the seminal work of Benamou and Brenier [17]. They showed that the W_2 metric between $\mu_0, \mu_1 \in \mathscr{P}_2$ could be computed as

$$W_2^2(\mu_0, \mu_1) = \inf\left\{ \int_0^1 \int_{\mathbb{R}^3} |v(t,x)|^2 d\mu(t,x)dt \;\middle|\; (\mu, v) \in CE \right\}, \tag{1.4}$$

where the constraint $(\mu, v) \in CE$ means that the curves $t \mapsto \mu(t) \in \mathscr{P}_2$ and $v \in L^2_\mu((0,T) \times \mathbb{R}^3)$ solve the continuity equation (in the distributional sense) with endpoints given by μ_0, μ_1

$$\partial_t \mu + \nabla \cdot (\mu v) = 0, \quad \mu|_{t=0} = \mu_0, \quad \mu|_{t=1} = \mu_1.$$

Now suppose $u \in \mathscr{P}_2$ solves (1.2) and it has finite initial entropy $\mathcal{H}[u(0)] < +\infty$. Then the pair $(u, -\nabla \log u) \in CE$ with endpoints $\mu_0 = u(0), \mu_1 = u(1)$. Here, I am interchanging $du(x) = u(x)dx$ where u solves (1.2), an identification which will be frequently used. Moreover, the integral in (1.4) can be explicitly written with the pair giving

$$W_2^2(u(0), u(1)) \leq \int_0^1 \int_{\mathbb{R}^3} u|\nabla \log u|^2 = \int_0^1 D_h(u(t))dt.$$

In fact, up to some omitted technical details, we have the following infinitesimal version involving the **metric derivative with respect to** W_2

$$|\dot{u}|_{W_2}(t) := \lim_{h \to 0} \frac{W_2(u(t+h), u(t))}{|h|} = D_h(u(t)).$$

The approach my supervisors José A. Carrillo and Matias Delgadino, our collaborator Laurent Desvillettes, and I took to understanding the gradient flow structure of (L) was to determine, in a similar way to (1.4), the Landau metric d_L which satisfies

$$|\dot{f}|_{d_L}(t) := \lim_{h \to 0} \frac{d_L(f(t+h), f(t))}{|h|} = D_L(f(t)).$$

In Chapter 2, the construction of d_L is modelled after the right-hand side of (1.4) by infimising some lower semi-continuous energy integral subject to constraints which generalise the continuity equation. To make the analogy between D_L and D_h more clear, I considered the following differential operator $\tilde{\nabla}$ for functions $\phi = \phi(v)$ given by

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

D_L can then be rewritten with more visual similarity to D_h as

$$D_L(f) = \frac{1}{2} \iint_{\mathbb{R}^6} f f_* |\tilde{\nabla} \log f|^2 dv_* dv.$$

Roughly speaking, $\tilde{\nabla}$ replaces the usual gradient ∇ in the right-hand side (1.4). Let us formally recall the intuition for the heat equation. A solution of the heat equation, u, can also be interpreted as a solution of the continuity equation with velocity $-\nabla \log u$. The task now is to take a solution, f, of the Landau equation and deduce the form of the generalised continuity equation it solves with 'velocity' given by $-\tilde{\nabla} \log f$. The generalised continuity equation to be considered is something I call the 'grazing continuity equation' (GCE) which, for a given velocity $V = V(t, v, v_*) \in \mathbb{R}^3$, reads

$$\partial_t f + \frac{1}{2} \tilde{\nabla} \cdot (f f_* V) = 0. \tag{1.5}$$

The grazing continuity equation, as for the usual continuity equation in (1.4), should only be interpreted in the distributional sense. That is, for a test function $\phi \in C_c^\infty((0,T) \times \mathbb{R}^3)$, the statement $(f, V) \in GCE$ means that (1.5) holds when tested against ϕ

$$\int_0^T \int_{\mathbb{R}^3} \partial_t \phi(t, v) df_t(v) dt + \frac{1}{2} \int_0^T \iint_{\mathbb{R}^6} [\tilde{\nabla}\phi](v, v_*) \cdot V_t(v, v_*) df_t(v) df_t(v_*) dt = 0.$$

This takes care of the generalised continuity equation constraint. Finally, the integral to be infimised should be the $L^2_{ff_*}$ norm of V (mimicking the L^2_u norm of the velocity for the heat equation) leading to

$$d_L^2(f_0, f_1) := \inf \left\{ \frac{1}{2} \int_0^1 \iint_{\mathbb{R}^6} f f_* |V|^2 dv_* dv dt \,\middle|\, (f, V) \in GCE,\ f|_{t=0} = f_0,\ f|_{t=1} = f_1 \right\}.$$

This is the idea which led to the construction of the Landau metric d_L. This is explored in more detail in Chapter 2. We provide conditions under which weak solutions of (L) (H-solutions from Villani [118], to be precise) are equivalent to curves f which decrease $\mathcal{H}$ as fast as possible in the d_L metric (i.e. gradient flows). Notice that the choices of test functions $\phi \in \{1, v, |v|^2\}$ yield $\tilde{\nabla}\phi = 0$. Hence, curves f solving (1.5) a priori conserve mass, momentum, and energy. This motivates the choice of $\mathscr{P}_2$ as the space to equip with d_L as done in Chapter 2.

I will avoid any rigorous pointwise interpretation of the grazing continuity equation (1.5) (more precisely $\tilde{\nabla}\cdot$) in this thesis, but an excursion in this direction may be helpful to clarify the relationship between (1.5) and (L). Firstly, the notation $\tilde{\nabla}\cdot$ refers to the formal adjoint of $\tilde{\nabla}$ so that a generalised integration by parts formula holds for (smooth) scalar functions $\phi = \phi(v)$ and vector fields $A =$

$A(v, v_*) \in \mathbb{R}^3$

$$\iint_{\mathbb{R}^6} [\tilde{\nabla}\phi](v, v_*) \cdot A(v, v_*) dv_* dv = -\int_{\mathbb{R}^3} \phi(v)[\tilde{\nabla} \cdot A](v) dv.$$

Note that the left-hand side is an integral over $\mathbb{R}^6$, whereas the right-hand side is an integral over $\mathbb{R}^3$, and $[\tilde{\nabla} \cdot A]$ is a scalar function only depending on v. This was (and to some extent remains) a source of confusion to me, although it is, of course, owed to the fact that $\tilde{\nabla}$ sends functions of $v \in \mathbb{R}^3$ to functions of $(v, v_*) \in \mathbb{R}^6$. The non-locality of the Landau equation is embedded into the introduction of new variables from $\tilde{\nabla}$. Let us compute the precise form of $[\tilde{\nabla} \cdot A](v) \in \mathbb{R}$ given $A(v, v_*) \in \mathbb{R}^3$. Assuming ϕ and A are smooth and decay fast enough and noticing $\Pi[v - v_*]\nabla\left(|v - v_*|^{1+\frac{\gamma}{2}}\right) = 0$, the generalised integration by parts formula gives

$$\begin{aligned}\iint_{\mathbb{R}^6} \tilde{\nabla}\phi \cdot V &= \iint_{\mathbb{R}^6} |v - v_*|^{1+\frac{\gamma}{2}} \Pi[v - v_*](\nabla\phi - \nabla_*\phi_*) \cdot A \\ &= \int_{\mathbb{R}^3} \left\{ \int_{\mathbb{R}^3} \Pi[v - v_*]\nabla(|v - v_*|^{1+\frac{\gamma}{2}}\phi) \cdot A(v, v_*) dv \right\} dv_* \\ &\quad - \int_{\mathbb{R}^3} \left\{ \int_{\mathbb{R}^3} \Pi[v - v_*]\nabla_*(|v - v_*|^{1+\frac{\gamma}{2}}\phi_*) \cdot A(v, v_*) dv_* \right\} dv.\end{aligned}$$

Now, we use the usual integration by parts formula for the terms in the curly brackets. Notice that we have interchanged variables by Fubini's theorem.

$$\begin{aligned}\iint_{\mathbb{R}^6} \tilde{\nabla}\phi \cdot V &= -\int_{\mathbb{R}^3} \left\{ \int_{\mathbb{R}^3} |v - v_*|^{1+\frac{\gamma}{2}} \phi \nabla \cdot (\Pi[v - v_*]A(v, v_*))\, dv \right\} dv_* \\ &\quad + \int_{\mathbb{R}^3} \left\{ \int_{\mathbb{R}^3} |v - v_*|^{1+\frac{\gamma}{2}} \phi_* \nabla_* \cdot (\Pi[v - v_*]A(v, v_*))\, dv_* \right\} dv \\ &= -\int_{\mathbb{R}^3} \phi \left\{ \int_{\mathbb{R}^3} |v - v_*|^{1+\frac{\gamma}{2}} \nabla \cdot (\Pi[v - v_*]A(v, v_*)) dv_* \right\} dv \\ &\quad + \int_{\mathbb{R}^3} \phi \left\{ \int_{\mathbb{R}^3} |v - v_*|^{1+\frac{\gamma}{2}} \nabla \cdot (\Pi[v - v_*]A(v_*, v)) dv_* \right\} dv.\end{aligned}$$

In the last equality, we have simply swapped the labels $v \leftrightarrow v_*$ (notice the arguments of A in the last line are also swapped) and applied Fubini. Formally passing the divergence (in v) outside of the curly brackets, recalling again $\Pi[v - v_*]\nabla\left(|v - v_*|^{1+\frac{\gamma}{2}}\right) = 0$, we arrive at

$$[\tilde{\nabla} \cdot A](v) = \nabla_v \cdot \int_{\mathbb{R}^3} |v - v_*|^{1+\frac{\gamma}{2}} \Pi[v - v_*](A(v, v_*) - A(v_*, v)) dv_*.$$

With this pointwise expression for $\tilde{\nabla}\cdot$, we can substitute $V = -\tilde{\nabla} \log f = -|v - v_*|^{1+\frac{\gamma}{2}} \Pi[v - v_*](\nabla \log f - \nabla_* \log f_*)$ into the grazing continuity equation (1.5) to recover precisely the Landau equation (L).

Developing the $\frac{1}{2}\tilde{\nabla}\cdot(ff_*V)$ term in (1.5) with this substitution, we obtain

$$
\begin{aligned}
&-\frac{1}{2}\tilde{\nabla}\cdot(ff_*V)=\frac{1}{2}\tilde{\nabla}\cdot(ff_*\tilde{\nabla}\log f)\\
&=\frac{1}{2}\tilde{\nabla}\cdot(ff_*|v-v_*|^{1+\frac{\gamma}{2}}\Pi[v-v_*](\nabla\log f-\nabla_*\log f_*))\\
&=\frac{1}{2}\nabla\cdot\left(\int_{\mathbb{R}^3}|v-v_*|^{2+\gamma}\Pi[v-v_*](ff_*(\nabla\log f-\nabla_*\log f_*)-f_*f(\nabla_*\log f_*-\nabla\log f))dv_*\right)\\
&=\nabla\cdot\left(f\int_{\mathbb{R}^3}f_*|v-v_*|^{2+\gamma}\Pi[v-v_*](\nabla\log f-\nabla_*\log f_*)dv_*\right).
\end{aligned}
$$

This is exactly the right-hand side of the Landau equation (L). Just to re-emphasise the analogy with the heat equation, the previous computations show that the Landau equation (L) can be written as

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

which should be compared to the heat equation (1.2)

$$\partial_t u=\nabla\cdot(u\nabla\log u).$$

The gradient flow perspective is reinforced in Chapter 3 wherein the grazing collision limit connecting the Boltzmann equation (B) to the Landau equation (L) is recovered in this framework. One can formally derive the Landau equation in the case when most collisions occur with a small angle $\theta\ll 1$. More specifically, fix $\epsilon>0$ and extend β (c.f. (1.1)) from $[0,\pi/2]$ to the whole real line by zero. We consider the scaling (discussed in Section 3.2.1) that concentrates β around $\theta=0$ given by

$$\beta^\epsilon(\theta):=\frac{\pi^3}{\epsilon^3}\beta\left(\frac{\pi\theta}{\epsilon}\right),\quad \theta\in[0,\epsilon/2]. \tag{1.6}$$

Denoting the new collision kernel B^ϵ induced here through (1.1), this gives rise to a new collision operator Q_B^ϵ which replaces the right-hand side of (B). The grazing collision limit is precisely passing to $\epsilon\downarrow 0$. For a fixed smooth f, the formal computations of Degond and Lucquin-Desreux [45], and Desvillettes [51] show the convergence

$$Q_B^\epsilon(f,f)\overset{\epsilon\downarrow 0}{\to}C_\beta Q_L(f,f),\quad C_\beta=\frac{\pi}{8}\int_0^{\pi/2}\theta^2\beta(\theta)d\theta.$$

For simplicity, this thesis assumes $C_\beta=1$ which fixes a normalisation for β. Of course, while these preliminary computations established the formal 'convergence of the collision operators,' the natural question is rigorous convergence of solutions

$$\partial_t f^\epsilon = Q_B^\epsilon(f^\epsilon, f^\epsilon) \to \partial_t f = Q_L(f,f),$$

which is the main focus of Chapter 3.

In order to outline the formal idea of the grazing collision limit from the gradient flow perspective, I should discuss first the H-theorem for (B) (this equation was the original context of the H-theorem) for an arbitrary kernel B (including those B^ϵ generated by (1.6)). Again, suppose $f = f(t,v) \in L^1_+$ is a (smooth) solution to (B) and formally use $\log f$ as a test function

$$\frac{d}{dt}\mathcal{H}[f(t,\cdot)] = \int_{\mathbb{R}^3} (\log f + 1)\partial_t f dv = -\int_{v\in\mathbb{R}^3} (\log f + 1) \int_{v_*\in\mathbb{R}^3}\int_{\mathbb{S}^2} [f'f'_* - ff_*]B\, d\sigma dv_* dv$$
$$= -\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} \log f [f'f'_* - ff_*]B\, d\sigma dv_* dv.$$

The disappearance of the constant term is due to the fact that the collision map sending (v, v_*) to (v', v'_*) is involutive (this and the subsequent properties of the collision map are examined in detail in Chapter 3). As well, the collision kernel B is invariant under this change of variables. More precisely, the swapping of variables $(v, v_*) \leftrightarrow (v', v'_*)$ is volume preserving and results in the sign change

$$\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} [f'f'_* - ff_*]B\, d\sigma dv_* dv = -\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} [f'f'_* - ff_*]B\, d\sigma dv_* dv = 0.$$

In fact, after playing around with the various symmetries of this change of variables, we arrive at the entropy dissipation equality which proves the H-theorem for (B)

$$\mathcal{H}[f(0)] = \mathcal{H}[f(t)] + \int_0^t \underbrace{\left(\frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} [f'f'_* - ff_*]\log\frac{f'f'_*}{ff_*} B\, d\sigma dv_* dv\right)}_{=:D_B(f)\geq 0} dt. \tag{EDE_B}$$

The positivity of the Boltzmann dissipation, D_B, comes from the following fact (not so obvious to me at the time, so I convinced myself in Corollary A.1)

$$(x-y)\log\frac{x}{y} \geq 4|\sqrt{x} - \sqrt{y}|^2 \geq 0, \quad \forall x, y > 0,\, x \neq y.$$

In the same way that (EDE_L) was used as a model for the gradient flow interpretation of (L), so too is (EDE_B) the basis for the gradient flow interpretation of (B) [61]. In fact, the mathematical content of Chapter 2 was inspired by and extends the gradient flow perspective of the Boltzmann equation for $\gamma = 0$ by Erbar [61]. The gradient flow perspective of the grazing collision limit can be formally posed in the following way. **Suppose, for $\epsilon > 0$ the grazing collision parameter, $f^\epsilon = f^\epsilon(t,v) \in L^1_+$ is a sequence of functions which satisfy**

$$\mathcal{H}[f^\epsilon(0)] = \mathcal{H}[f^\epsilon(t)] + \int_0^t \underbrace{\left(\frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}[f^{\epsilon\prime}f^{\epsilon\prime}_* - f^\epsilon f^\epsilon_*]\log\frac{f^{\epsilon\prime}f^{\epsilon\prime}_*}{f^\epsilon f^\epsilon_*}B^\epsilon d\sigma dv_* dv\right)}_{=:D^\epsilon_B(f^\epsilon)} dt, \qquad (EDE_\epsilon)$$

where B^ϵ is a sequence of collision kernels induced through (1.1) **by the scaling of** (1.6) **for a fixed β. What conditions ensure f^ϵ has a convergent subsequence towards a limit, f, as $\epsilon \downarrow 0$ and that (EDE_ϵ) converges to (EDE_L)?** In other words, what mechanisms drive the convergence of the entropy dissipations $\int_0^t D^\epsilon_B(f^\epsilon) \to \int_0^t D_L(f)$? This question captures the formal gradient flow viewpoint we adopted to recover the grazing collision limit in Chapter 3.

Moving slightly away from pure theory, Chapter 4 discusses a numerical scheme to approximate solutions to (L). We consider a particle method to solve a regularised version of (L) with right-hand side, for $\varepsilon > 0$ (we use here ε as opposed to the grazing collision parameter ϵ previously), given by

$$Q^\varepsilon_L(f,f) = \nabla\cdot\left(f\int_{\mathbb{R}^3} f_*|v-v_*|^{2+\gamma}\Pi[v-v_*]\left(\nabla\psi_\varepsilon*\log\left[f*\psi_\varepsilon\right] - \nabla_*\psi_\varepsilon*\log\left[f_**\psi_\varepsilon\right]\right)dv_*\right). \qquad (L_\varepsilon)$$

Here the new features are coloured in green and ψ_ε denotes the Gaussian mollifier

$$\psi_\varepsilon(v) = \frac{1}{(2\pi\varepsilon)^{\frac{d}{2}}}\exp\left(-\frac{|v|^2}{2\varepsilon}\right). \qquad (\psi_\varepsilon)$$

Let me briefly explain the motivation behind the choice of ψ_ε and its appearance in (L_ε). Obviously, convolution with ψ_ε yields a smoothing property to the Landau equation, but the Gaussian profile was chosen because the stationary solutions of (L) consist entirely of these distributions (they are referred to as Maxwellians in the kinetic theory literature, a convention I now adopt). This is revisited in Lemma 4.7 based on arguments in [74]. As for the appearance of ψ_ε in somewhat arbitrary locations in (L_ε), we arrive at the interplay between the theoretical elements of Chapter 2 and the numerical motivations in Chapter 4. The Landau equation (L) should be viewed as the gradient flow of $\mathcal{H}$ with respect to d_L while (L_ε) should be viewed as the gradient flow of a regularised entropy, $\mathcal{H}_{2,\varepsilon}$, with respect to d_L. Here, the regularised entropy is given by

$$\mathcal{H}_{2,\varepsilon}[\mu] := \mathcal{H}[\mu*\psi_\varepsilon] = \int(\mu*\psi_\varepsilon)\log\left[\mu*\psi_\varepsilon\right], \quad \mu\in\mathscr{P}.$$

The regularisation by convolution against ψ_ε destroys the parabolic nature of (L). Thus, (L_ε) is viewed as a continuity equation with non-local and non-linear velocity given by

$$U^\varepsilon[f](v) := -\int_{\mathbb{R}^3} f_*|v-v_*|^{2+\gamma}\Pi[v-v_*](\nabla\psi_\varepsilon * \log[f*\psi_\varepsilon] - \nabla_*\psi_\varepsilon * \log[f_* * \psi_\varepsilon])dv_*.$$

The loss of diffusion precisely allows for a particle method. It makes sense to feed an empirical measure

$$\mu_0^N = \frac{1}{N}\sum_{i=1}^N \delta_{v_0^i}, \quad i = 1, \dots, N \in \mathbb{N},$$

as initial data to (L_ε) and expect that the evolution maintains the empirical measure structure by solving the system of N ODEs

$$\dot{v}^i(t) = U^\varepsilon[\mu^N(t)](v^i(t)), \quad \mu^N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{v^i(t)}, \quad i = 1, \dots, N.$$

In other words, the regularisation enables $\mu^N(t)$ to solve (L_ε) with initial condition μ_0^N. All the information is contained in the evolution of the trajectories v^i. This can be programmed into a computer, which is precisely the topic of Chapter 4.

Another benefit of the particular regularisation is that (L_ε) preserves nice structural properties similar to those of (L). I already alluded to this with the terminology of gradient flows, but (L_ε) satisfies a regularised version of the H-theorem; if $f = f(t,v) \in L^1_+$ is a (smooth) solution of (L_ε), then

$$\frac{d}{dt}\mathcal{H}_{2,\varepsilon}[f(t)] = -D_L^\varepsilon(f) :=$$
$$-\frac{1}{2}\iint_{\mathbb{R}^6} ff_*|v-v_*|^{2+\gamma}\left|\Pi[v-v_*]\left(\nabla\psi_\varepsilon * \log[f*\psi_\varepsilon] - \nabla_*\psi_\varepsilon * \log[f_* * \psi_\varepsilon]\right)\right|^2 dv_* dv \le 0.$$

Here, I have again marked in green the differences between the regularised dissipation D_L^ε and the unregularised dissipation D_L. The retention of these structural properties is beneficial for numerical analysis not only in the present context, but also for other PDE models [12, 13, 30].

1.2 Gradient flow theory

Recently, the gradient flow community has been very active and successful in PDEs, starting from the significant gradient flow landmarks by Jordan, Kinderleher, and Otto [90]; Benamou and Brenier [17]; Otto [103]; and the seminal reference book by Ambrosio, Gigli, and Savare [6]. Some of the advantages of gradient flow techniques include new insights into functional inequalities, stable numerical methods, and quantitative understanding of trends to equilibrium in tandem with uniqueness of solutions [35].

With reference to Chapter 2, the foundation set by Benamou and Brenier [17] with extensions [60, 34] helped consolidate the construction of the Landau metric d_L. These works not only provided the dynamic interpretation of optimal transportation distances, but also introduced new metrics allowing for more PDEs to be analysed in the framework of gradient flows. For example, the Benamou-Brenier form of the 2-Wasserstein distance (1.4) can be rewritten in flux variables by introducing $m = fv$ and for $f \in L^1_+$ so that

$$W_2^2(\mu_0, \mu_1) = \inf \left\{ \int_0^1 \int_{\mathbb{R}^3} \frac{|m(t,x)|^2}{f(t,x)} dxdt \,\middle|\, (f,m) \in CE^* \right\},$$

where the continuity equation constraint $(f,m) \in CE^*$ means (still in the distributional sense)

$$\partial_t f + \nabla \cdot m = 0, \quad f|_{t=0}\mathcal{L} = \mu_0, \quad f|_{t=1}\mathcal{L} = \mu_1.$$

One of the main ideas from [60] to create new distances is modifying the integrand of the infimised functional. Consider the function

$$\alpha : (x,y) \in \mathbb{R}^3 \times \mathbb{R}_{\geq 0} \mapsto \begin{cases} \frac{|x|^2}{y}, & y > 0 \\ 0, & x = 0,\ y = 0 \\ +\infty, & x \neq 0,\ y = 0 \end{cases} \in [0, +\infty].$$

Then the infimised functional for W_2 is nothing but

$$\int_0^1 \int_{\mathbb{R}^3} \alpha(m(t,x), f(t,x)) dxdt$$

and such functionals enjoy nice lower semi-continuity and convexity properties [25, 62] owing to the lower semi-continuity, joint convexity, and 1-homogeneity of α (c.f. Lemma 2.14). By choosing different α with the aforementioned properties, new dynamic distances can be generated. This idea is applied to construct d_L in Chapter 2.

I would now like to present some of the insights that can be gleaned from finite dimensional gradient flows, as many of these ideas are carried over to the infinite dimensional case of PDEs. I have adapted the discussions from [112, 35] for the context of this thesis. Consider $E : \mathbb{R}^d \to \mathbb{R}$ a smooth and convex function, then the gradient flow associated to E is the curve $x : [0, +\infty) \to \mathbb{R}^d$ satisfying the ODE

$$\frac{d}{dt} x(t) = -\nabla E(x(t)), \quad x(0) = x_0 \in \mathbb{R}^d. \tag{1.7}$$

Clearly the chain rule from calculus,

$$\frac{d}{dt}E(x(t)) = \nabla E(x(t)) \cdot \frac{d}{dt}x(t), \tag{1.8}$$

lends itself to the H-theorem corresponding to (1.7), because we have

$$\frac{d}{dt}E(x(t)) = -|\nabla E(x(t))|^2 \leq 0.$$

In a similar way to (EDE_L), (EDE_B), and (EDE_ϵ), $|\nabla E(x)|^2$ is the dissipation corresponding to (1.7). Moreover, the associated entropy dissipation equality reads

$$E(x(t)) - E(x(0)) = -\frac{1}{2}\int_0^t |\nabla E(x(s))|^2 ds - \frac{1}{2}\int_0^t \left|\frac{d}{ds}x(s)\right|^2 ds. \tag{1.9}$$

We will discuss the decision to include the derivative of x in (1.9), unlike what is initially written in (EDE_L), (EDE_B), and (EDE_ϵ). For now, it is clear that C^1 curves x satisfying (1.7) (i.e. classical solutions) also satisfy (1.9) (i.e. gradient flow solutions) by the previous discussion. In fact, the reverse is also true.

Lemma 1.1 (Gradient flow solutions are classical solutions)**.** *If $x \in C^1$ satisfies* (1.9)*, then x also satisfies* (1.7)*.*

Proof. Owing to the chain rule (1.8) and Young's inequality (with a minus), we have

$$\begin{aligned} E(x(t)) - E(x(0)) &= \int_0^t \frac{d}{ds}E(x(s))ds = \int_0^t \nabla E(x(s)) \cdot \frac{d}{ds}x(s)ds \\ &\geq -\frac{1}{2}\int_0^t |\nabla E(x(s))|^2 ds - \frac{1}{2}\int_0^t \left|\frac{d}{ds}x(s)\right|^2 ds. \end{aligned}$$

Since x satisfies (1.9), then the previous computation holds with equality. When equality is attained in Young's inequality, this implies co-linearity between $\frac{d}{ds}x(s)$ and $\nabla E(x(s))$ which, by the minus sign choice, yields (1.7). □

The idea of the proof of Lemma 1.1 is exactly the same as the proof of Theorem 2.8 and, by extension, Theorem 2.10. The key ingredient to proving the equivalence of (1.7) and (1.9) is the chain rule (1.8). In the context of Chapter 2, proving the infinite dimensional version of the chain rule was the major source of technical difficulty. For these non-trivial issues, I refer to Proposition 2.23 and Section 2.5.

Returning to the inclusion of $\left|\frac{d}{ds}x\right|^2$ in the entropy dissipation equality (1.9), this was clearly essential in the proof of Lemma 1.1. In terms of the Landau equation, and gradient flow PDEs in general, this object is replaced by the metric derivative. I neglected to include this term in the introductory discussions of (EDE_L) and (EDE_ϵ) for technical simplicity, but the reader can rest assured that the corresponding metric derivatives are present in Chapters 2 and 3.

One of the advantages of considering (1.9) over (1.7) is that (1.9) has nice stability properties with respect to compactness of curves x. In [109, 113] Sandier and Serfaty utilised (the infinite dimensional version of) (1.9) as a way to streamline the characterisation of the limit of evolutions that have a gradient flow structure. Effectively, the problem reduces to checking the lower semi-continuous convergence of the associated dissipations and metric derivatives. This idea is at the forefront of Chapter 3 where the limit from (EDE_ϵ) to (EDE_L) is made rigorous. Although this procedure is by now well-known, our contributions include a method to prove the detailed steps from Sandier and Serfaty. This approach has been heavily used in recent years in a wide array of scenarios. Making a non-exhaustive list, we mention the works in Cahn-Hilliard [39, 16, 50], diffusion to reaction limits [9], particle methods second order [30] and fourth order [100] non-linear diffusion, congested crowd motion [1], and dislocations [19].

As for existence of solutions to (1.7), consider the so-called JKO or de Giorgi minimisation scheme [90] for (1.7) given by fixing a time step $\tau > 0$ and iteratively finding

$$x_\tau^{n+1} = \operatorname{argmin}_{x\in\mathbb{R}^d} \left\{ E(x) + \frac{1}{2\tau}|x - x_\tau^n|^2 \right\}, \quad x_\tau^0 := x_0 \in \mathbb{R}^d, \quad n \in \mathbb{N}. \tag{JKO-finite}$$

Since E is convex, minimisers x_τ^{n+1} exist and are unique. Moreover, they can also be characterised as critical points of the expression in the following curly brackets

$$\nabla \left\{ E(x) + \frac{1}{2\tau}|x - x_\tau^n|^2 \right\}\bigg|_{x=x_\tau^{n+1}} = 0 \iff \frac{x_\tau^{n+1} - x_\tau^n}{\tau} = -\nabla E(x_\tau^{n+1}).$$

Notice the final expression is the implicit Euler discretisation of (1.7). Hence, under mild compactness hypotheses, x_τ^{n+1} converges to solutions of (1.7) as $\tau \downarrow 0$. In their paper [90], Jordan, Kinderlehrer, and Otto proved that solutions to a large class of Fokker-Planck equations can be constructed in a similar way to (JKO-finite). For example, in the specific case of the heat equation (1.2), they showed that it can be viewed as the 2-Wasserstein gradient flow of the Boltzmann entropy by looking at

$$u_\tau^{n+1} = \operatorname{argmin}_{u\in\mathscr{P}_2} \left\{ \mathcal{H}[u] + \frac{1}{2\tau}W_2^2(u, u_\tau^n) \right\}, \quad u_\tau^0 = u_0 \in \mathscr{P}_2, \quad n \in \mathbb{N}. \tag{JKO-heat}$$

The adjustment to accommodate Landau is clear; the 2-Wasserstein metric should be replaced with d_L, which is examined in Section 2.4.

1.3 The Landau equation

The Landau equation (L) is an important partial differential equation in kinetic theory. We have already discussed its use in describing colliding particles in plasma physics [95], as well its relation to the Boltzmann equation (B) through the grazing collision limit [45, 51, 118]. Similar to the Boltzmann equation (see [20] for a consistency result and related derivation issues), the rigorous derivation of the Landau equation from particle dynamics is still a huge challenge. 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 $\gamma = -3$ are known as the Maxwellian and Coulomb cases, respectively. The Coulomb potential is the most physically realistic case to consider and yet, it presents severe but interesting mathematical difficulties. I now present a brief history of known well-posedness results in the spatially homogeneous case, along with some intuitive computations for the difficulty when γ decreases. After these calculations, I will emphasise an important result by Desvillettes [52] which is central to Chapter 2.

Desvillettes and Villani [56, 57, 119] studied (L) in full detail for the hard and Maxwellian potential cases. In particular, they showed that solutions to (L) exist, instantaneously regularise, and are unique classical solutions when the initial data has finite mass and energy. These solutions exhibit rapid decay and are also bounded below by Maxwellians (Gaussian distributions).

When $\gamma < 0$, less can be said. In the case $\gamma \in [-2, 0)$, Fournier and Guérin [65] proved global existence and uniqueness of weak solutions conditional to L^p propagation and high order moment propagation (the exponents for L^p and order of moments depend on γ). Their approach was to apply techniques from probability and view (L) as a stochastic process (i.e. write the corresponding SDE for random variables distributed by solutions to (L)). The conditions outlined in [65] were guaranteed in [122] for L^p propagation (see also [3]) and [28] for moment propagation (see the references therein).

For $\gamma \in [-4, -2)$, even more difficulties arise concerning uniqueness and regularity of solutions, for example. To begin, while the previous results for L^p and moment propagation for $\gamma \in [-2, 0)$ were proven globally, when $\gamma \in (-3, -2)$, such estimates are only known locally-in-time or, equivalently, in a small data framework [65, 3]. I will now focus on the Coulomb case, although the results stated below may apply when $\gamma \in (-3, -2)$. One of the first significant results concerning well-posedness for Coulomb is due to Guo [78], who proved global existence and uniqueness of classical solutions in a near-Maxwellian setting. Fournier proved local-in-time uniqueness of weak solutions as long as they remained bounded in L^∞_v [63]. Concerning potential blow up in different norms, we mention Gualdani and Guillen [73], who considered radially symmetric (see also [72, 75]) and monotonically decreasing solutions of (L). They proved that classical solutions exist for all times unless the $L^{\frac{3}{2}+}$ norm blows up near the maximal time. The recent result of [55] shows that the $\dot{H}^1$ semi-norm may blow up in finite time but will eventually (with explicitly computable time based on initial data) remain bounded. Another partial answer to the question of blow up is that the set of singular times has Hausdorff dimension less than $\frac{1}{2}$ [69].

Regarding weaker notions of solutions, i.e. not requiring finite L^p norms, I mention here the global existence of renormalized solutions [5] and H-solutions by Villani [118]. To gain some intuition for the challenges associated to smaller γ I formally present now a few of the mathematical estimates. This will also demonstrate the philosophy of H-solutions, which we extend with the gradient flow perspective of this thesis. We begin with a 'classical' formulation of weak solutions to (L) by splitting the collision operator into the natural gain and loss terms

$$\begin{aligned} Q_L(f,f) &= \nabla\cdot\left(\int_{\mathbb{R}^3}|v-v_*|^{2+\gamma}\Pi[v-v_*](f_*\nabla f - f\nabla_* f_*)dv_*\right) \\ &= \nabla\cdot\left(\underbrace{\left\{\int_{\mathbb{R}^3} f_*|v-v_*|^{2+\gamma}\Pi[v-v_*]\right\}}_{=:A[f]}\nabla f - \underbrace{\left\{\int_{\mathbb{R}^3}|v-v_*|^{2+\gamma}\Pi[v-v_*]\nabla_* f_*\right\}f}_{=:B[f]}\right). \end{aligned} \tag{1.10}$$

Since the solution-dependent coefficients are convolutions, we see

$$A[f] = \{(|\cdot|^{2+\gamma}\Pi[\cdot]) * f\}(v), \quad B[f] = \mathrm{div}A[f] = \{(|\cdot|^{2+\gamma}\mathrm{div}\Pi[\cdot]) * f\}(v).$$

Here, we have again taken advantage of the fact that $\nabla(|v-v_*|^{2+\gamma}) \in \ker\Pi[v-v_*]$. By formally integrating by parts, we could define the action of $Q_L(f,f)$ without differentiating f on test functions ϕ by

$$\begin{aligned} &-\int\nabla\phi^T A[f]\nabla f + \int\nabla\phi\cdot(\nabla\cdot A[f])f = \int D^2\phi : A[f]f + 2\int\nabla\phi\cdot(\nabla\cdot A[f])f \\ &= \iint_{v,v_*} f_* f|v-v_*|^{2+\gamma}\partial^{ij}\phi(v)\Pi^{ij}[v-v_*] + 2\iint_{v,v_*} f_* f\partial^i\phi(v)\partial^j\Pi^{ij}[v-v_*] \\ &= \frac{1}{2}\iint ff_*|v-v_*|^{2+\gamma}(\partial^{ij}\phi(v)+\partial^{ij}\phi(v_*))\Pi^{ij}[v-v_*] \\ &\qquad\qquad - 2\iint ff_*|v-v_*|^{\gamma}(v-v_*)^i\cdot(\partial^i\phi(v)-\partial^i\phi(v_*)) \\ &=: \langle Q_L(f,f),\phi\rangle_{\text{classic}}. \end{aligned} \tag{1.11}$$

Here, I have expanded the definition of $A[f]$, symmetrised with the change of variables $v \leftrightarrow v_*$, and used the fact that, in d dimensions, the projection satisfies

$$\partial^j \Pi^{ij}[x] = -(d-1)|x|^{-2}x^i.$$

Firstly, the separation of gain and loss terms in (1.10) gives the interpretation of (L) as a parabolic equation with diffusion matrix $A[f]$ and drift $B[f]$ (hence (L) is also referred to as the Fokker-Planck equation). The diffusion matrix satisfies the following coercivity estimate [56, 3]

$$\xi^T A[f]\xi \geq C \langle v\rangle^\gamma |\xi|^2, \quad \forall \xi \in \mathbb{R}^d, \quad \gamma \geq -3, \tag{1.12}$$

where $C > 0$ is a constant depending only on bounds for $\int \langle v\rangle^2 f$ and $\mathcal{H}[f]$. These quantities are formally controlled in the evolution of (L) anyway; the moments up to second order are conserved by observing $\langle Q_L(f,f), \phi\rangle_{\text{classic}} = 0$ when $\phi \in \{1,\, v,\, |v|^2\}$ and the entropy is bounded via (EDE_L). The degenerate coercivity from (1.12) already suggests additional difficulty in applying parabolic theory for $\gamma < 0$.

Returning to (1.11), both integrals there can be estimated by

$$|\langle Q_L(f,f), \phi\rangle_{\text{classic}}| \lesssim \|\phi\|_{W_v^{2,\infty}} \iint f f_* |v - v_*|^{2+\gamma}.$$

Here, the criticality of $\gamma = -2$ is very clear when defining the action of $Q_L(f,f)$ on ϕ through (1.11). When $\gamma \geq -2$, this definition makes sense by a priori moment estimates, which are propagated anyway [28, 122]. However, when $\gamma < -2$, at least some finite L^p norm is required for $p > 1$, but we have previously discussed that this is only propagated locally-in-time or for small data.

Despite this apparent obstruction for the very soft potentials, Villani's H-solutions [118] sidestep these L^p requirements. This notion takes advantage of (EDE_L) assuming finite initial entropy $\mathcal{H}[f_0] < +\infty$ to use $D_L(f) \in L^1_t$ in the following formulation of weak solutions. Instead of splitting the gain and loss terms like in (1.11), we directly integrate by parts $Q_L(f,f)$ against test functions ϕ and symmetrise with $v \leftrightarrow v_*$ to define

$$\begin{aligned}
&- \iint f f_* |v - v_*|^{2+\gamma} \nabla\phi^T \Pi[v - v_*](\nabla \log f - \nabla_* \log f_*) dv_* dv \\
&= -\frac{1}{2} \iint f f_* |v - v_*|^{2+\gamma} (\nabla\phi - \nabla_*\phi_*)^T \Pi[v - v_*](\nabla \log f - \nabla_* \log f_*) dv_* dv \\
&=: \langle Q_L(f,f), \phi\rangle_H.
\end{aligned}$$

The Cauchy-Schwarz inequality and the Mean Value Theorem applied to ϕ for $\gamma \in [-4,-2)$ yield

$$\begin{aligned}
\int_0^T |\langle Q_L(f,f),\phi\rangle_H|\,dt &\lesssim \left(\int_0^T \iint ff_*|v-v_*|^{2+\gamma}|\nabla\phi - \nabla_*\phi_*|^2\right)^{\frac{1}{2}}\left(\int_0^T D_L(f)\right)^{\frac{1}{2}} \\
&\lesssim \left(\int_0^T D_L(f)\right)^{\frac{1}{2}} \times \begin{cases} \|\phi\|_{L_t^\infty W_v^{1,\infty}}\left(\int_0^T \iint ff_*|v-v_*|^{2+\gamma}\right)^{\frac{1}{2}}, & \gamma \in [-2,0) \\ \|\phi\|_{L_t^\infty W_v^{2,\infty}}\left(\int_0^T \iint ff_*|v-v_*|^{4+\gamma}\right)^{\frac{1}{2}}, & \gamma \in [-4,-2) \end{cases} \\
&\lesssim \|\phi\|_{L_t^\infty W_v^{2,\infty}}\left(\int_0^T D_L(f)\right)^{\frac{1}{2}}\left(\int_0^T \int (1+|v|^2)f\right)^{\frac{1}{2}}.
\end{aligned}$$

Therefore, no a priori L^p assumptions on f are required to make sense of $\langle Q_L(f,f),\phi\rangle_H$. In particular, we require bounds only on the physical quantities (mass, energy, entropy, and entropy-dissipation) which are formally conserved or bounded anyway. This illustrates the advantage of H-solutions by exploiting finite (L^1 in time) entropy-dissipation.

From the outset, solutions defined by $\langle Q_L,\phi\rangle_H$ are weaker than $\langle Q_L,\phi\rangle_{\text{classic}}$. This gap was resolved by Desvillettes [52, 53, 54] (c.f. precisely Theorem 2.32). This result captures the idea that finite entropy-dissipation implies functional regularity; a weighted Fisher information and 'cross Fisher information' are controlled by the Landau entropy-dissipation. In his paper, Desvillettes uses Theorem 2.32 with a Sobolev embedding (hence gaining L^p integrability) to prove that H-solutions are classical weak solutions; $\langle Q_L,\phi\rangle_{\text{classic}}$ is well-defined for H-solutions. Theorem 2.32 is also one of the fundamental results used for our analysis of the gradient flow structure of the Landau equation presented in Chapter 2.

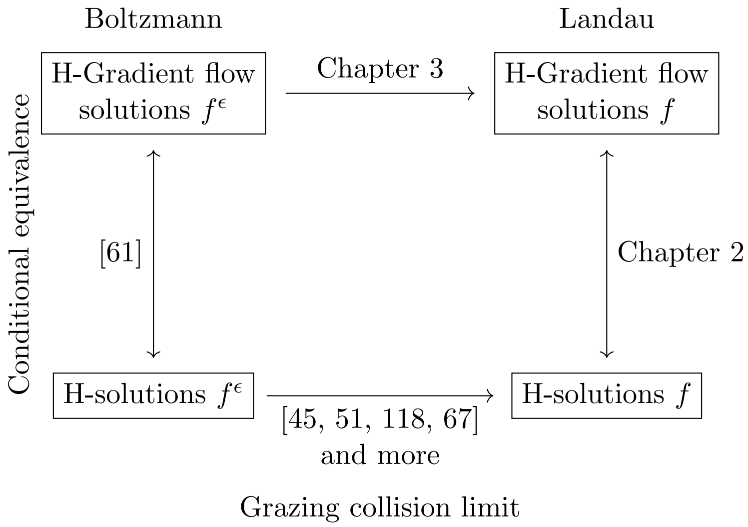
We close this survey of the theoretical analysis of (L) by mentioning some other recent and interesting results for the spatially inhomogeneous Landau equation

$$\partial_t f + v\cdot\nabla_x f = Q_L(f,f), \quad f = f(t,x,v), \quad x \in \Omega \subset \mathbb{R}^d, \quad v \in \mathbb{R}^d.$$

The ellipticity estimate (1.12) has an analogue in this case, however, the point to emphasise is that the collision operator induces diffusion only in the v variable. The x variable gains regularity through mixing with the v variable by the $v\cdot\nabla_x$ transport operator. This vague argument has been investigated by [70, 83, 86, 77, 76] not only for Landau, but also for general kinetic equations. Finally, the seminal theory of renormalised solutions [58, 59] applies to the space inhomogeneous Landau equation as well, yielding global existence of such solutions [5]. The results mentioned in this paragraph have also been applied to the Boltzmann equation, which is surveyed momentarily.

1.4 The Boltzmann equation

The Boltzmann equation (B) is the central equation in kinetic theory modelling particle collisions in a gas, as well as many other interacting particle systems [37]. Its connection to the Landau equation has already been introduced via the grazing collision limit. The purpose of Chapter 3 is to recover and simplify the grazing collision limit from the perspective of gradient flows building on the works of Erbar [61] and Chapter 2 (c.f. the diagram below). Complementing the overview in Section 1.3, I now discuss the historical development of well-posedness results for (B) in relation to the grazing collision limit.



The earliest well-posedness result for usual weak solutions to (B) is due to Arkeryd [7, 8] who required cut-off assumptions on the collision kernel (replacing $|v - v_*|^{2+\gamma}$ with something bounded near $v = v_*$). In particular, this excludes the physically relevant soft potential cases $\gamma < 0$. Nevertheless, this well-posedness theory was sufficient for Arsen'ev and Buryak [10] in 1990, who rigorously proved convergence in the grazing collision limit from Boltzmann to Landau. Removing the cut-off assumption, Villani was able to prove the grazing collision limit for soft and hard potentials using his notion of H-solutions [118]. Here, we again see the advantage of using (EDE_B). Shortly after, in a collaboration with Alexandre [4, 5] they upgraded from weak to strong convergence in the grazing collision limit by applying the regularity estimate they achieved with Desvillettes and Wennberg [2]. The argument for the gain in compactness relied on velocity average techniques [71] applied to renormalized solutions [59, 58]. The main result in [2] was the precursor (for Boltzmann) to Theorem 2.32 (for Landau); finite entropy-dissipation implies functional regularity.

More recently, Godinho [67] and He [79] gave quantitative rates of convergence for short times in the grazing collision limit of weak solutions. Godinho's proof relies on the construction of weak solutions from the well-posedness theory of Fournier and Mouhot [66] and Fournier and Guérin [64].

The presentation in Chapter 3 seeks to simplify the grazing collision limit globally-in-time with minimal assumptions (finite initial mass, energy, and entropy) in comparison to Godinho. In particular, the well-posedness theory he uses requires at least 7th order moments (possibly more depending on γ) and holds for short times. Our simplification follows the program of Γ-convergence set by Sandier and Serfaty [113, 109].

1.5 Numerical analysis of (L)

The goal of Chapter 4 is to introduce a deterministic particle method to numerically approximate solutions to (L). Here, I present a historical overview firstly for deterministic particle methods for general PDEs and secondly for the numerical approximations applied to (L).

There have been several strategies to accommodate particle methods for diffusive-type equations in the literature by introducing suitable regularisations of the flux for the continuity equation [108]. The case of the heat equation $\frac{\partial \rho}{\partial t} = \Delta\rho$ was considered in [49, 107] by interpreting the Laplacian as induced by a velocity field u, $\Delta\rho = -\nabla \cdot (u\rho)$, $u = -\nabla\rho/\rho$, and regularising the numerator and denominator separately by convolution with a mollifier. Well-posedness of the resulting system of ordinary differential equations and a priori estimates relevant to the method were studied in [92], and subsequently extended to nonlinear diffusion [102, 97, 99]. Variations of these methods allowing the weights to change in time were also analysed in [47, 48]. The main disadvantage of these existing deterministic particle methods is that, with the exception of [97] for the porous medium equation $\frac{\partial \rho}{\partial t} = \Delta\rho^2$, they do not preserve the gradient flow structure [97]. For further background on deterministic particle methods, we refer to the review [38], and for particle methods applied to transport equations, we refer to [41, 40, 43]. As mentioned earlier in Section 1.1, we have followed the strategy in [30] of regularising the free energy functional instead in order to retain the gradient flow structure at the particle method level.

To approximate Q_L, a popular approach is to use the Fourier-Galerkin spectral method [104]. This technique takes advantage of the convolutional property of the collision integral so that the resulting method can be implemented efficiently using fast Fourier transform (FFT). To be specific, the total complexity of one time evaluation of the collision operator requires $\mathcal{O}(N_v^d \log N_v)$ complexity, where N_v is the number of Fourier modes in each velocity dimension. As we shall see, the proposed particle method would require $\mathcal{O}(N^2)$ complexity, where N is the total number of particles. Hence, in terms of efficiency, it may not be as fast as the spectral method. However, it is able to preserve all the physical properties of the equation: positivity; conservation of mass, momentum, and energy; and entropy dissipation. This is in contrast to the spectral method, wherein the truncated Fourier approximation destroys the structure of the solution (only mass is conserved, no positivity, no conservation of momentum and energy, no entropy decay). Furthermore, $\mathcal{O}(N^2)$ is the direct cost of the particle method (a naïve implementation). With the help of the fast summation technique such as the treecode, this cost can be reduced to $\mathcal{O}(N \log N)$. We will explore this acceleration in the current paper while an in-depth study will be deferred to future work.

It is important to mention that the particle-in-cell (PIC) method [18, 85, 116] is currently the dominant method to solve Vlasov-type equations (kinetic equations with spatial inhomogeneity) which is essentially a particle method. Hence, our proposed method is a natural candidate to be coupled with the PIC methodology to yield an efficient Lagrangian solver for the inhomogeneous Landau equation. The numerical exploration of these ideas for inhomogeneous problems is certainly a research topic of great interest, constituting a major future direction. Other proposed techniques which preserve the main properties of the Landau operator include Eulerian methods based on mesh discretisations in velocity, see [46, 93, 23] and the references therein. However, they are more difficult to incorporate within the PIC approach for spatially inhomogeneous problems.

1.6 Contributions

Concluding this thesis' introduction, I would now like to discuss the results here in relation to the literature mentioned previously.

The main result of Chapter 2 is Theorem 2.10, which gives conditions for which Villani's H-solutions are equivalent to the gradient flows of the Landau equation for $\gamma \in (-3, 0]$. This sets the foundation to apply the sophisticated machinery of gradient flows [6] for the analysis of (L). In particular, I addressed the open question concerning uniqueness of weak solutions in Section 1.3 for the very soft potentials $\gamma \in (-4, -2)$. Owing to pioneering results from [35], the theory of gradient flows provides a systematic framework to investigate uniqueness of solutions. This is beyond the scope of this thesis, however, it is certainly an interesting area of future investigation. Pre-empting the discussion on numerics for Chapter 4, the construction of the Landau metric in Theorem 2.7 introduces the JKO scheme [90] as a new numerical method to solve (L). Although the JKO scheme is by now classical for gradient flow theory, its application here to the spatially homogeneous Landau equation is novel. This brings variational approaches to bear when approximating (L). Moreover, by construction, the JKO scheme dissipates entropy, and such structure-preserving methods are important in numerics [12, 13].

Chapter 3 simplifies the argument for the grazing collision limit from Boltzmann to Landau. Here, we make use of the relevant mechanisms (bounded mass, energy, and entropy dissipation) to prove this limit on a global time interval which improves upon Godinho's result [67]. Moreover, the grazing collision limit verifies the compatibility between Erbar's Boltzmann metric ([61]) and the Landau metric

(Chapter 2) in the passage of this limit. In [61], Erbar proves that the Boltzmann metric is an upper bound for the 1-Wasserstein metric in the case $\gamma = 0$. We are able to extend this estimate to the soft potentials (c.f. Lemma 3.12).

Although the collision operator Q_L is diffusive (c.f. (1.12)), Chapter 4 provides numerical evidence that the particle method introduced there is (almost) second order accurate. It extends the ideas [30, 44] to kinetic equations with an emphasis on preserving the entropy dissipation structure. In particular, the characterisation of Maxwellian steady states remains true with our regularisation.

This thesis rigorously adopts the gradient flow framework of the Landau equation. Of course, each chapter lends itself to the pursuit of different research directions. Some of these avenues are discussed in Chapter 5.

Chapter 2

Gradient Flow structure of the Landau Equation

The content of this chapter is based on joint work with José A. Carrillo, Matias G. Delgadino, and Laurent Desvillettes. It is a preprint [31] submitted for publication.

This chapter develops the gradient flow viewpoint of the Landau equation which is used throughout this thesis. Our approach is heavily inspired by Erbar's gradient flow description of the Boltzmann equation [61] for Maxwellian molecules ($\gamma = 0$). We are able to extend this formulation in the case of Landau for soft potentials within the range $-3 < \gamma \leq 0$. We make precise here the ideas presented in Section 1.1. In particular, we construct and explore properties of a tailor-made metric d_L in Section 2.2 after reviewing some concepts from gradient flow theory and stating our results in Section 2.1. We prove that (L) (resp. (L_ε) for a different mollifer) is the gradient flow of $\mathcal{H}$ (resp. $(\mathcal{H}_\varepsilon)$ defined later) with respect to d_L in Section 2.5 (resp. Section 2.3). Construction of gradient flow solutions using the JKO scheme [90] with the regularised entropy $(\mathcal{H}_\varepsilon)$ is studied in Section 2.4.

2.1 Preliminaries and the main results

2.1.1 Notations and definitions

We use the Japanese angle bracket notation

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

For $\varepsilon > 0$, we denote our regularisation kernel to be an exponential distribution

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

We also define other mollifiers with different tail behaviour

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

for $s > 0$; we point out some of the limitations and restrictions on $s > 0$ in the later estimates. We shall refer to $G^{2,\varepsilon}$ as the Maxwellian regularisation. 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;

$$\mathscr{P}_{p,E}(\mathbb{R}^d) := \left\{ \mu \in \mathscr{P}_p(\mathbb{R}^d) \,\middle|\, m_p(\mu) \le E \right\}.$$

$\mathcal{M}$ and $\mathcal{M}^d$ abbreviate $\mathcal{M}(\mathbb{R}^{2d})$ and $\mathcal{M}^d(\mathbb{R}^{2d})$, the space of signed scalar and $\mathbb{R}^d$-valued Radon measures with the weak-* topology against the space $C_c(\mathbb{R}^{2d})$ and $(C_c(\mathbb{R}^{2d}))^d$, respectively. For $T > 0$, $\mathcal{M}_T^d$ is the space of $\mathbb{R}^d$-valued signed Radon measures on $\mathbb{R}^d \times \mathbb{R}^d \times [0, T]$.

For $\mu \in \mathscr{P}$, we define a family of regularised entropies $\mathcal{H}_\varepsilon[\mu]$ by

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

note the distinction from $\mathcal{H}_{2,\varepsilon}$ in Chapter 1. Formally, one can calculate the first variation of this functional in $\mathscr{P}_2$ as

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

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

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

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 $\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu} = 1 + G^\varepsilon * \log[\mu * G^\varepsilon]$. 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 $\frac{\delta \mathcal{F}}{\delta f}$, we refer to the $\mathcal{F}$ Landau equation as

$$\partial_t f = \nabla \cdot \left(f \int_{\mathbb{R}^d} f_* |v - v_*|^{2+\gamma} \Pi[v - v_*] \left(\nabla \frac{\delta \mathcal{F}}{\delta f} - \nabla_* \frac{\delta \mathcal{F}_*}{\delta f_*} \right) dv_* \right). \tag{2.1}$$

Similar to the formal computations in Section 1.1, the $\mathcal{F}$ Landau equation (2.1) can be more concisely written using ($\tilde{\nabla}$) 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 (2.1) with $\frac{\delta \mathcal{F}}{\delta f}$, one obtains an analogy of Boltzmann's H-theorem with the functional $\mathcal{F}$;

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

We will refer to $D_{\mathcal{F}}$ as the $\mathcal{F}$ dissipation with specific mention of $D_L = D_{\mathcal{H}}$ for the usual Landau entropy-dissipation. These notations induce our notion of weak solutions to the $\mathcal{F}$ Landau equation (2.1) closely following Villani's H-solutions [118].

Definition 2.1 (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 (2.1) if the following hold.

1. $f_t \in L^1_+$ for all $t \in [0,T]$ with uniformly bounded second moment; $\sup_{t\in[0,T]} m_2(f_t) < +\infty$.
2. The functional $\mathcal{F}$ evaluated along the curve is bounded by its initial value

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

3. The $\mathcal{F}$ dissipation is time-integrable;

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

4. For every test function $\phi \in C_c^\infty((0,T)\times\mathbb{R}^d)$, equation (2.1) is satisfied in weak form

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

For $\varepsilon > 0$, we will refer to the weak $\mathcal{H}_\varepsilon$ 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 [118], is essential in our analysis.

2.1.2 Quick review of gradient flow theory

We recall the basic definitions of gradient flow theory that can be found in more generality in [6, Chapter 1]. Throughout this section, (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 2.2 (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 $s \le t \in (a,b)$

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

Among all possible functions m in Definition 2.2, one can make the following minimal selection.

Definition 2.3 (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 [6, Theorem 1.1.2].

Definition 2.4 (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

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

Using Young's inequality and moving everything to one side, the inequality in Definition 2.4 implies

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

If the reverse inequality also holds, one obtains the stronger Energy Dissipation *Equality.* This leads to the notion of gradient flows.

Definition 2.5 (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 non-increasing and the following inequality holds

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

F has the following natural candidates for upper gradient.

Definition 2.6 (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

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

2.1.3 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 2.7 (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 $\tau \in \mathscr{P}_2$ the subset $\mathscr{P}_\tau := \left\{\mu \in \mathscr{P}_{2,m_2(\tau)}(\mathbb{R}^d) \,|\, d_L(\mu, \tau) < \infty\right\}$ is a complete geodesic space.*

The content of this theorem is that our new proposed distance 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. A necessary condition is that the probability measures μ_0 and μ_1 should have the same second moment. In the construction of d_L detailed in Section 2.2, 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 $0 \le t,\, s \le T$.

Theorem 2.8 (varepsilon equivalence)**.** *Fix any $\varepsilon, E > 0, \gamma \in [-4, 0]$. 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 $\mathcal{H}_\varepsilon$ with respect to its upper gradient $\sqrt{D_{\mathcal{H}_\varepsilon}}$ 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 2.4) which is the following

Theorem 2.9 (Existence of curves of maximal slope)**.** *For any $\varepsilon, E > 0, \gamma \in [-4, 0]$, 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 $\mathcal{H}_\varepsilon$ with respect to its upper gradient $\sqrt{D_{\mathcal{H}_\varepsilon}}$.*

Remark 2.1. *The choice of an exponential convolution kernel G^ε is perhaps unnatural compared to the Maxwellian regularisation $G^{2,\varepsilon}$ for the regularised entropy $\mathcal{H}_\varepsilon$. We discuss in more detail the estimates that fail using $G^{2,\varepsilon}$ in Remark 2.7 as it pertains to Theorem 2.8. With respect to Theorem 2.9, the* *general construction of some curve can be done even with the Maxwellian regularisation. However, due to the same lack of estimates, this curve might not be a curve of maximal slope with respect to* $\sqrt{D_{\mathcal{H}_\varepsilon}}$*. This is discussed in Remark 2.8. On the other hand, the choice of a Maxwellian regularisation is considered numerically in Chapter 4.*

Motivated by the numerical simulations in Chapter 4, Theorems 2.8 and 2.9 provide the theoretical support to (L_ε). In order to treat the full Landau equation (L) i.e. pass $\varepsilon \downarrow 0$, more assumptions are required.

Theorem 2.10 (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*

(A2.1) *(Moments and L^p) Assume that there exists some $0 < \eta \leq \gamma + 3$ such that*

$$\langle v \rangle^{2-\gamma} f_t(v) \in L^\infty_t \left(0, T; L^1_v \cap L_v^{\frac{3-\eta}{3+\gamma-\eta}}(\mathbb{R}^3)\right).$$

(A2.2) *(Finite entropy) We assume that the initial entropy is finite*

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

(A2.3) *(Finite entropy-dissipation) We assume that the entropy-dissipation of f is integrable in time*

$$D_L(f_t) = D_{\mathcal{H}}(f_t) = \frac{1}{2} \iint_{\mathbb{R}^6} f f_* \left| \tilde{\nabla} \frac{\delta \mathcal{H}}{\delta f} \right|^2 dv dv_* =$$
$$\frac{1}{2} \iint_{\mathbb{R}^6} f f_* |v - v_*|^{\gamma+2} |\Pi[v - v_*](\nabla \log f - \nabla_* \log f_*)|^2 dv dv_* \in L^1_t(0, T).$$

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

Remark 2.2. *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 **(A2.1)**–**(A2.3)** exist (they are strong and unique under extra conditions). We refer to [122], and Appendix B of [54] when $\gamma > -2$, for details.*

*When $\gamma \in (-3, -2)$, Assumption **(A2.1)** is not known to hold for global weak solutions with large initial data. Solutions satisfying **(A2.1)**–**(A2.3)** are nevertheless known to exist for initial data close to equilibrium (cf. [78], in a more general spatially inhomogeneous context), or in the Coulomb case $\gamma = -3$ (in that case $\frac{3-\eta}{3+\gamma-\eta}$ being replaced by ∞) for large initial data, but for short times [55, 11].*

*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 by the third author and Villani [56, 57]. We believe that our results also carry to the hard potentials. In particular, the exponents in assumption (**A2.1**) should be modified to*

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

We emphasise that these conditions are guaranteed since the required moments and L^p integrability are propagated from appropriate initial data when $\gamma > 0$ [56, 57]. This condition appears in [53, Corollary 2.7]. It is the hard potential version of Theorem 2.32 which is crucial to the proof of Theorem 2.10. 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 2.33, for example.

*The Coulomb case $\gamma = -3$ is excluded because assumption (**A3.1**) is insufficient in the case $\gamma = -3$. This is discussed more in Remark 2.10.*

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 (L), or equivalently, contructing solutions of the original Landau equation by passing to the limit $\varepsilon \downarrow 0$ in Theorem 2.9. Some of the difficulties in achieving this result are the propagation of moments for the regularised Landau equation uniformly in ε and the compactness of sequences with bounded-in-ε regularised entropy dissipation $D_{\mathcal{H}_\varepsilon}$. The rest of this chapter is devoted to showing the main four theorems in the next four sections.

2.2 The Landau metric d_L

Our approach to defining the distance d_L mentioned in Theorem 2.7 closely follows the dynamic formulation of transport distances originally due to Benamou and Brenier [17] and further extended by Dolbeault, Nazaret, and Savaré [60]. We also refer the reader to Erbar [61] for a similar approach. The proof of Theorem 2.7 is contained in Section 2.2.3 after establishing a few preliminary notions. The formulation of d_L consists of two components:

1. The continuity equation constraint which inspires our generalisation in Section 2.2.1.

2. The ‘action’ or energy functional $\mathcal{A}_L$ which inspires our generalisation in Section 2.2.2.

2.2.1 Grazing continuity equation

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

$$\partial_t \mu_t + \frac{1}{2}\tilde{\nabla} \cdot M_t = 0, \quad \text{in } (0,T) \times \mathbb{R}^d, \tag{2.2}$$

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

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

For some regularity properties, we will also need to assume the following moment condition

$$\int_0^T \iint_{\mathbb{R}^{2d}} (1 + |v| + |v_*|) d|M_t|(v,v_*) dt < \infty. \tag{2.3}$$

The curves $(\mu_t)_{t\in[0,T]}, (M_t)_{t\in[0,T]}$ are Borel families of measures belonging to $\mathscr{P}$ and $\mathcal{M}^d$ respectively. We will refer to $\mu = \mu_t$ from the pair as a *curve* and $M = M_t$ as a *grazing rate*. We first establish some a-priori properties of solutions to the grazing continuity equation.

Lemma 2.11 (Continuous representative). *For families $(\mu_t), (M_t)$ in $\mathscr{P}$ and $\mathcal{M}^d$, respectively, satisfying the grazing continuity equation and the finite moment condition* (2.3)*, there exists a unique weakly continuous representative curve $(\tilde{\mu}_t)_{t\in[0,T]}$ such that $\tilde{\mu}_t = \mu_t$ 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 following 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 [6, 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 this 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 the grazing continuity equation, the distributional time derivative is

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

Using a Mean-Value estimate for $\gamma \in [-4,-2)$, we have the following estimates depending on $\gamma \in [-4,0]$,

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

The precise estimate used above reads

$$|\tilde{\nabla}\zeta| \lesssim \begin{cases} |v-v_*|^{1+\frac{\gamma}{2}} \sup|\nabla\zeta|, & \gamma\in[-2,0] \\ |v-v_*|^{2+\frac{\gamma}{2}} \sup|D^2\zeta|, & \gamma\in[-4,-2) \end{cases} \lesssim_\gamma (1+|v|+|v_*|) \begin{cases} \sup|\nabla\zeta|, & \gamma\in[-2,0] \\ \sup|D^2\zeta|, & \gamma\in[-4,-2) \end{cases}.$$

Hence, for every $\zeta \in C_c^2(\mathbb{R}^d)$, $t \mapsto \mu_t(\zeta)$ belongs to $W^{1,1}(0,T)$ (and $\dot{\mu}_t(\zeta)$ is also its weak derivative). The rest of the proof proceeds as in [6, Lemma 8.1.2] using the C^2-norm of ζ for the soft potentials $\gamma \in [-4,-2)$ as opposed to just the C^1-norm of ζ. To elaborate, there are two final properties to check so that the continuous representative belongs in $\mathscr{P}$.

1. We need to show that μ_t extends to a continuous curve. Using [22, Theorem 8.2] to relate $\mu_t(\zeta) - \mu_s(\zeta) = \int_s^t \mu_\lambda(\zeta) d\lambda$, the previous estimate shows that there is a unique extension to a continuous curve $\{\tilde{\mu}_t\}_{t\in[0,T]}$ in $(C_c^2(\mathbb{R}^d))'$ since

$$|\mu_t(\zeta) - \mu_s(\zeta)| \le \|\zeta\|_{C^2} \iint_s^t \int_{\mathbb{R}^{2d}} (1+|v|+|v_*|) d|M_\lambda|(v,v_*) d\lambda.$$

2. We are done once we can establish $\{\mu_t\}_{t\in[0,T]}$ is tight. This can be accomplished by choosing particular test functions ζ approximating large balls and applying the previous estimate.

The last point to prove is the formula

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

This follows from a standard approximation argument by picking particular functions $\eta_\delta \in C_c^1(0,T)$ and $\phi \in C_c^\infty((0,T)\times\mathbb{R}^{2d})$ such that η_δ approximates the indicator function $\chi_{[t_1,t_2]}$ as $\delta \downarrow 0$ in the grazing continuity equation

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

□

Lemma 2.12 (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 $\mathscr{P}$, $\mathcal{M}^d$ respectively satisfying* (2.2) *and the moment condition* (2.3)*. Assume further that*

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

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

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

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

Proof. We show the proof of the conservation of energy for $\gamma \in [-4,-2]$ since the conservation of momentum is similar. To compress some notations, we denote $w = |v - v_*|^{1+\frac{\gamma}{2}}$. We consider a fixed $\varphi \in C_c^\infty(B_2)$ which satisfies

$$0 \le \varphi \le 1, \quad \varphi|_{B_1} = 1, \text{ and } |\nabla\varphi| \lesssim 1.$$

We define $\varphi_R(v) := \varphi(v/R)$ for $R > 0$ and use this as a test function in the grazing continuity equation to obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d} |v|^2 \varphi_R(v)\, d\mu_t(v) - \int_{\mathbb{R}^d} |v|^2 \varphi_R(v)\, d\mu_0(v) \\
&\quad = \int_0^t \iint_{\mathbb{R}^{2d}} w\Pi \left(v\varphi_R(v) + |v|^2 \frac{\nabla\varphi(v/R)}{R} - v_*\varphi_R(v_*) - |v_*|^2 \frac{\nabla\varphi(v_*/R)}{R} \right) dM_s(v,v_*)ds
\end{aligned} \tag{2.4}$$

We estimate one half of the terms in (2.4) using the cancellation from the projection Π to obtain

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

The last inequality is obtained by the Mean-Value Theorem when $|v - v_*| \le 1$ since the properties of ϕ_R give

$$|v\phi_R(v) - v_*\phi_R(v_*)| \le |v - v_*| \sup_{w\in B_R} \left(|\phi_R(w)| + \frac{|w|}{R} |\nabla\phi(w/R)| \right) \lesssim |v - v_*|.$$

Hence, using the triangle inequality for the case $|v - v_*| \ge 1$, we have the following estimate

$$w\left|v\varphi_R(v) - v_*\varphi_R(v_*)\right| \lesssim \begin{cases} |v - v_*|^{2+\frac{\gamma}{2}} & |v - v_*| \le 1 \\ |v| + |v_*| & |v - v_*| \ge 1. \end{cases}$$

Similarly, using that $\nabla\varphi_R$ is supported in $B_{2R}\setminus B_R$ and that $\left|\partial^i\left\{|v|^2\frac{\partial^j\varphi(v/R)}{R}\right\}\right|\lesssim 1$, we obtain that

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

where we have controlled the difference with the Mean-Value Theorem. From the previous bounds, we can use (2.3) to take $R\to\infty$ in (2.4) 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 proof for conservation of momentum involve testing the grazing continuity equation against $v_i\phi_R$ where v_i is the i-th component of v. The case $\gamma\in[-4,-2]$ follows the same argument as just presented. For $\gamma\in[-2,0]$, the estimates can be more blunt since the weight is no longer singular. □

Remark 2.3. *Note that as γ increases into the range $(-2,0]$, the weight function $w=|v-v_*|^{1+\frac{\gamma}{2}}$ starts adding growth so the Mean-Value estimates in Lemma 2.12 no longer apply unless more moments of M are assumed than* (2.3)*.*

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

Definition 2.13 (Grazing continuity equation)**.** For some terminal time $T>0$, we define 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 (2.3)

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

3. The grazing continuity equation (2.2) is satisfied in the distributional sense. That is, for every $\phi\in C_c^\infty((0,T)\times\mathbb{R}^d)$,

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

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 $GCE(\lambda,\nu)$ as those pairs $(\mu, M) \in GCE_T$ such that $\mu_0 = \lambda,\ \mu_T = \nu$. For $E > 0$, we will speak of curves $(\mu, M) \in GCE_T^{2,E} = GCE_T^E$ such that $\sup_{t\in[0,T]} m_2(\mu_t) \leq E$.

2.2.2 Action of a curve

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

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

Lemma 2.14. *α is jointly lower semi-continuous, convex, and positively 1-homogeneous.*

Proof. These properties are easily, if tediously, verified. In the case $s \neq 0$, the convexity of α follows since the Hessian of $\frac{|u|^2}{2s}$ has non-negative determinant. □

For fixed $\mu \in \mathscr{P}$, $M \in \mathcal{M}^d$, we define $\mu^1 \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d)$ by

$$\mu^1(dv, dv_*) := \mu \otimes \mu(dv, dv_*) = \mu(dv)\mu(dv_*).$$

Consider $\tau \in \mathcal{M}$ given by $\tau = \mu^1 + |M|$ and the decompositions $\mu^1 = f^1\tau$ and $M = N\tau$. We define the action functional as

$$\mathcal{A}_L(\mu, M) := \iint_{\mathbb{R}^{2d}} \alpha(N, f^1) d\tau. \tag{2.5}$$

This is well-defined by the 1-homogeneity of α; if $\sigma \in \mathcal{M}$ is another measure which dominates μ^1 and $|M|$, then also $\tau \ll \sigma$ so $\frac{d\tau}{d\sigma}$ exists and we can write

$$\iint_{\mathbb{R}^{2d}} \alpha\left(\frac{dM}{d\sigma}, \frac{d\mu^1}{d\sigma}\right) d\sigma = \iint_{\mathbb{R}^{2d}} \alpha\left(\frac{dM}{d\tau}, \frac{d\mu^1}{d\tau}\right) \frac{d\tau}{d\sigma} d\sigma.$$

The following lemma establishes a more concrete expression for the action functional.

Lemma 2.15. *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 $\mathcal{A}_L(\mu, M) < \infty$. Then, M is absolutely continuous with respect to $ff_* dv dv_*$ given by density $U : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ such that $M = ff_* U dv dv_* = m dv dv_*$ and*

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

Proof. The proof is identical to [61, Lemma 3.6] up to appropriate modifications but we present it here for clarity. Consider the same dominating measure $\tau \in \mathcal{M}$ from before the statement of the lemma with the same notations for the densities;

$$\tau = \mu \otimes \mu + |M|, \quad \mu \otimes \mu = f^1 \tau, \quad M = N\tau.$$

For a fixed measurable subset $A \subset \mathbb{R}^{2d}$, suppose

$$0 = \iint_A f f_* dv_* dv = \mu \otimes \mu(A) = \iint_A f^1 d\tau(v, v_*).$$

We wish to deduce from this that $M(A) = 0$. Since τ is a positive measure, we must have $f^1 = 0$ for almost every $(v, v_*) \in A$. By the assumption of finite action, we deduce that $\alpha(N, f^1) = 0$ for almost every $(v, v_*) \in A$. This implies $N = 0$ for almost every $(v, v_*) \in A$ which leads to $M(A) = 0$ and we are done. □

Lemma 2.16 (Lower semi-continuity of action functional)**.** *The action functional $\mathcal{A}_L$ as defined in* (2.5) *is jointly lower semi-continuous in both arguments. Specifically, if $\mu_n \stackrel{\sigma}{\rightharpoonup} \mu$ weakly in $\mathscr{P}$ and $M_n \stackrel{*}{\rightharpoonup} M$ weakly-* in $\mathcal{M}^d$, we have*

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

Proof. This result is an application of the general lower semi-continuity results which can be found in [25, Theorem 3.4.3] and [62, Theorem 5.25] since α satisfies the required convexity, lower semi-continuity, and homogeneity assumptions by Lemma 2.14. □

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

Lemma 2.17. *Let $F : \mathbb{R}^{2d} \to [0, \infty]$ be measurable and fix any $\mu \in \mathscr{P}$, $M \in \mathcal{M}^d$. We have the following estimate*

$$\iint_{\mathbb{R}^{2d}} F(v, v_*) d|M|(v, v_*) \leq \sqrt{2}\mathcal{A}_L(\mu, M)^{\frac{1}{2}} \left(\iint_{\mathbb{R}^{2d}} F(v, v_*)^2 d\mu(v) d\mu(v_*) \right)^{\frac{1}{2}}. \tag{2.6}$$

Proof. This proof follows [61, Lemma 3.8]. We assume $\mathcal{A}_L(\mu, M) < +\infty$ or else (2.6) 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. By considering $\tau = \mu \otimes \mu + |M|$, we use Cauchy-Schwarz to estimate

$$\begin{aligned}\iint_{\mathbb{R}^{2d}} F d|M|(v, v_*) &\leq \iint_{\mathbb{R}^{2d}} F \left|\frac{dM}{d\tau}\right| d\tau(v, v_*) = \iint_{\mathbb{R}^{2d}} F \left(\left|\frac{dM}{d\tau}\right| \Big/ \sqrt{2\frac{d\mu \otimes \mu}{d\tau}}\right) \sqrt{2\frac{d\mu \otimes \mu}{d\tau}} d\tau \\ &\leq \left(\iint_{\mathbb{R}^{2d}} \alpha\left(\frac{dM}{d\tau}, \frac{d\mu \otimes \mu}{d\tau}\right) d\tau\right)^{\frac{1}{2}} \left(\iint_{\mathbb{R}^{2d}} 2F^2 d\mu \otimes \mu\right)^{\frac{1}{2}} \\ &= \sqrt{2}\mathcal{A}_L(\mu, M)^{\frac{1}{2}} \left(\iint_{\mathbb{R}^{2d}} F(v, v_*)^2 d\mu(v) d\mu(v_*)\right)^{\frac{1}{2}}.\end{aligned}$$

□

Remark 2.4. *Suppose we have* $\mu_t \in \mathscr{P}$ *such that* $\int_0^T m_2(\mu_t) dt < +\infty$, *then the previous estimate* (2.6) *for* $M \in \mathcal{M}_T^d$ *and* $F = 1 + |v| + |v_*|$ *yields*

$$\int_0^T \iint_{\mathbb{R}^{2d}} (1 + |v| + |v_*|) d|M_t|(v, v_*) dt \lesssim \int_0^T \mathcal{A}_L(\mu_t, M_t)^{\frac{1}{2}} \left(1 + 2\int_{\mathbb{R}^d} |v|^2 \, d\mu_t\right)^{\frac{1}{2}} dt. \tag{2.7}$$

Therefore, the time integrability of both the action and the second moment of μ_t *imply* M *satisfies the moment condition* (2.3). *In the sequel, we will be considering curves satisfying precisely this property to guarantee* (2.7).

Proposition 2.18. *Let* $(\mu_t^n, M_t^n)_n$ *be a sequence in* GCE_T *such that* $(\mu_0^n)_n$ *is tight and we have the following uniform bounds*

$$\sup_{n \in \mathbb{N}} \int_0^T \int_{\mathbb{R}^d} |v|^2 \, d\mu_t^n(v) dt < \infty \qquad \text{and} \qquad \sup_{n \in \mathbb{N}} \int_0^T \mathcal{A}_L(\mu_t^n, M_t^n) \, dt < \infty. \tag{2.8}$$

Then, there exists $(\mu_t, M_t) \in GCE_T$ *such that, possibly after extracting a subsequence, we have the following convergences*

$$\begin{aligned} \mu_t^n &\overset{\sigma}{\rightharpoonup} \mu_t \quad \text{weakly in } \mathscr{P}, \quad \forall t \in [0, T] \\ M_t^n dt &\overset{*}{\rightharpoonup} M_t dt \quad \text{weakly-* in } \mathcal{M}_T^d \end{aligned}.$$

Furthermore, along this subsequence we have the following lower semi-continuity

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

Sketch of the proof. This result follows from a similar proof to [60, Lemma 4.5] and [61, Proposition 3.11] which we sketch. The second moment bound for μ^n in (2.8) produces a limit μ. Recalling the application of Lemma 2.17 in Remark 2.4, the bounded action in (2.8) and the estimate (2.7) produce a limit $\tilde{M} \in \mathcal{M}((0, T) \times \mathbb{R}^{2d})$ along a subsequence of $M_t^n dt$. The time integrated version of Lemma 2.17 implies that $\tilde{M}$ can be disintegrated with respect to Lebesgue measure; $\tilde{M} = M_t dt$ for a curve $t \in$

$(0, T) \mapsto M_t \in \mathcal{M}^d$. The lower semi-continuity follows from Fatou's lemma and Lemma 2.16. $\square$

2.2.3 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 GCE_T^E for $T > 0$ some terminal time and arbitrary $E > 0$.

Definition 2.19. For $\lambda, \nu \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ we define the (square of the) Landau distance by

$$d_L^2(\lambda, \nu) := \inf \left\{ T \int_0^T \mathcal{A}_L(\mu_t, M_t) dt \,\middle|\, (\mu, M) \in GCE_T^E(\lambda, \nu) \right\}. \tag{2.9}$$

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 characterisation of d_L which can be seen in other PDE contexts such as [61, 60].

Lemma 2.20. *Given $\lambda, \nu \in \mathscr{P}_{2,E}(\mathbb{R}^d)$, we have*

$$d_L(\lambda, \nu) = \inf \left\{ \int_0^T \sqrt{\mathcal{A}_L(\mu_t, M_t)} dt \,\middle|\, (\mu, M) \in GCE_T^E(\lambda, \nu) \right\}. \tag{2.10}$$

Proof. Here we reproduce the same proof as in [60, Theorem 5.4]. We denote, only for this proof, $\overline{d}_L(\lambda, \nu) = RHS_{(2.10)}$. By Cauchy-Schwarz, we have that for every $(\mu, M) \in GCE_T^E$

$$\int_0^T \sqrt{\mathcal{A}_L(\mu_t, M_t)} dt \le T^{\frac{1}{2}} \left(\int_0^T \mathcal{A}_L(\mu_t, M_t) dt \right)^{\frac{1}{2}}.$$

From here, we obtain

$$\overline{d}_L(\lambda, \nu) \le d_L(\lambda, \nu),$$

and so we turn to establishing the reverse inequality. As in [6, Lemma 1.1.4], for every $\delta > 0$ and $t \in [0, T]$, define

$$s_\delta(t) := \int_0^t \{\delta + \mathcal{A}_L(\mu_r, M_r)\}^{\frac{1}{2}} dr.$$

One sees that $s_\delta(\cdot)$ is strictly increasing and its derivative satisfies

$$s'_\delta(t) = \{\delta + \mathcal{A}_L(\mu_t, M_t)\}^{\frac{1}{2}} \ge \delta^{\frac{1}{2}}.$$

Define $S_\delta := s_\delta(T)$ so that $s_\delta : [0, T] \mapsto [0, S_\delta]$ is a strictly increasing surjection. Therefore, its inverse, denoted by t_δ is well-defined and is Lipschitz continuous with a bound on its derivative given by

$$t'_\delta(s) \le \delta^{-\frac{1}{2}}, \quad \forall s \in (0, S_\delta).$$

Moreover, an application of the inverse function theorem gives the form

$$t'_\delta(s_\delta(t)) = (\delta + \mathcal{A}_L(\mu_t, M_t))^{-\frac{1}{2}}.$$

By classical arguments on time rescaling (such as in [6, Lemma 8.1.3]), by defining $\mu^\delta_s = \mu_{t_\delta(s)}$ and $M^\delta_s = t'_\delta(s) M_{t_\delta(s)}$ one has that $(\mu^\delta, M^\delta) \in GCE^{2,E}_{S_\delta}$ with $\mu^\delta_0 = \lambda$, $\mu^\delta_{S_\delta} = \nu$ so by definition

$$d^2_L(\lambda, \nu) \leq S_\delta \int_0^{S_\delta} \mathcal{A}(\mu^\delta_s, M^\delta_s) ds = S_\delta \int_0^T \frac{\mathcal{A}_L(\mu_t, M_t)}{\delta + \mathcal{A}_L(\mu_t, M_t)} (\delta + \mathcal{A}_L(\mu_t, M_t))^{\frac{1}{2}} \, dt.$$

The equality follows from a change of variables in the integral from s to t. The term in the denominator appears because the action functional $\mathcal{A}_L$ is 2-homogeneous in its second argument. Continuing, we add δ to the numerator of the fraction and obtain the further estimate

$$d^2_L(\lambda, \nu) \leq \cdots \leq S_\delta \int_0^T (\delta + \mathcal{A}_L(\mu_t, M_t))^{\frac{1}{2}} \, dt = S^2_\delta.$$

Passing to the limit superior as $\delta \downarrow 0$ with the reverse Fatou lemma, we have

$$\begin{aligned} d_L(\lambda, \nu) &\leq \limsup_{\delta \downarrow 0} S_\delta = \limsup_{\delta \downarrow 0} \left\{ \int_0^T (\delta + \mathcal{A}_L(\mu_t, M_t))^{\frac{1}{2}} \, dt \right\} \\ &\leq \int_0^T \sqrt{\mathcal{A}_L(\mu_t, M_t)} dt = \overline{d}_L(\lambda, \nu). \end{aligned}$$

□

Proposition 2.21 (Minimising curve)**.** *Suppose that μ_0, $\mu_1 \in \mathscr{P}_{2,E}(\mathbb{R}^d)$ satisfy $d_L(\mu_0, \mu_1) < \infty$. Then there exists a curve $(\mu, M) \in GCE^{2,E}_1(\mu_0, \mu_1)$ attaining the infimum of* (2.9) *(equivalently, also* (2.10)*) and $\mathcal{A}_L(\mu_t, M_t) = d^2_L(\mu_0, \mu_1)$ for almost every $t \in [0, 1]$.*

Proof. The existence of a curve $(\mu, M) \in GCE^{2,E}_1(\mu_0, \mu_1)$ follows from the direct method of calculus of variations where the lower semi-continuity comes from Proposition 2.18. As for the formula $\mathcal{A}_L(\mu_t, M_t) = d^2_L(\mu_0, \mu_1)$ for almost every $t \in [0, 1]$, notice that Cauchy-Schwarz as well as the equivalent representations (2.9) and (2.10) give

$$\int_0^1 \sqrt{\mathcal{A}_L(\mu_t, M_t)} dt \leq \left(\int_0^1 \mathcal{A}_L(\mu_t, M_t) dt \right)^{\frac{1}{2}} = d_L(\mu_0, \mu_1) = \int_0^1 \sqrt{\mathcal{A}_L(\mu_t, M_t)} dt.$$

Clearly, these are all equalities which, by Cauchy-Schwarz, implies that $\sqrt{\mathcal{A}_L(\mu_t, M_t)}$ is a constant. □

We now have all the ingredients to prove Theorem 2.7.

Proof of Theorem 2.7. 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.

d_L distinguishes points

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 2.21 we can find $(\mu, M) \in GCE_1^{2,E}(\mu_0,\mu_1)$ which is a minimising curve and moreover $0 = d_L(\mu_0,\mu_1) = \mathcal{A}_L(\mu_t, M_t)$ 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 $(\mu, 0) \in GCE_1^{2,E}(\mu_0,\mu_1)$.

Symmetry

Symmetry follows because time can be reversed for every curve. For instance, if $(\mu, M) \in GCE_T^E(\mu_0,\mu_1)$, then one can check that the pair

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

belong to $GCE_T^E(\mu_1,\mu_0)$ with the same action.

Triangle inequality

We sketch the argument using a glueing lemma as in [60, Lemma 4.4]. Without loss of generality, 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 $d_L(\mu^1,\mu^2) < \infty$. By Proposition 2.21, we can find minimising curves connecting these probability measures

$$(\mu^{0\to 1}, M^{0\to 1}) \in GCE_1^{2,E}(\mu^0,\mu^1), \quad (\mu^{1\to 2}, M^{1\to 2}) \in GCE_1^{2,E}(\mu^1,\mu^2).$$

We take their concatenation from time 0 to 1 is given by

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

One can check that $(\mu, M) \in GCE_1^{2,E}(\mu^0,\mu^2)$, 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 have

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

d_L-convergence/boundedness implies weak convergence/compactness

Fix a sequence $(\mu^n)_{n\in\mathbb{N}} \in \mathscr{P}_{2,E}$ and $\mu^\infty \in \mathscr{P}_{2,E}$ such that $d_L(\mu^\infty, \mu^n) \to 0$ as $n \to \infty$. By Proposition 2.21, take minimising curves $(\nu^n, M^n) \in GCE_1^{2,E}(\mu^\infty, \mu^n)$ such that

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

By compactness in Proposition 2.18, there are limits $(\nu, M) \in GCE_1^{2,E}$ such that $\nu^n \rightharpoonup \nu$ and $M^n \overset{*}{\rightharpoonup} M$ up to a subsequence. Moreover, the lower semi-continuity in Proposition 2.18 gives

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

hence $M = 0$ so that ν is constant in time. Since $\nu(0) = \mu^\infty$, this implies $\mu^\infty = \nu(1) = \lim_{n\to\infty} \mu^n$ 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 [61], 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 2.21 guarantees the existence of a minimising curve $(\mu, M) \in GCE_1^{2,E}(\mu_0, \mu_1)$. One easily sees that this also induces a minimising curve for intermediate times. More precisely, for every $0 \leq r \leq s \leq 1$, we have that $(t \mapsto \mu_{t+r}, t \mapsto M_{t+r}) \in GCE_{s-r}^{2,E}(\mu_r, \mu_s)$ also minimises $d_L(\mu_r, \mu_s)$.

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 2.18, 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 semi-continuity 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

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

which implies $\mu^\infty \in \mathscr{P}_\tau$. □

Following the philosophy of Otto [103], we present the Riemannian geometrical interpretation of $(\mathscr{P}_{2,E}, d_L)$.

Proposition 2.22 (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* $\mathcal{M}^d$ *such that* $(\mu, M) \in GCE_T^E$ *with the property that*

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

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

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

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

$$M_t = U\mu_t \otimes \mu_t \qquad and \qquad U \in T_\mu := \overline{\{\tilde{\nabla}\phi \,|\, \phi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu_t\otimes\mu_t)}$$

such that $(\mu, \tilde{M}) \in GCE_T^E(\mu_0, \mu_T)$ *where we have equality*

$$|\dot{\mu}|^2(t) = \mathcal{A}_L(\mu_t, \tilde{M}_t) < +\infty, \quad a.e.\ t \in (0,T).$$

In other words, for a fixed μ*, the measure* $\tilde{M}$ *is minimal for the action functional so if* $\eta \in \mathcal{M}^d$ *is* $\tilde{\nabla}\cdot$*-free, meaning*

$$\iint_{\mathbb{R}^{2d}} \tilde{\nabla}\xi \cdot d\eta(v, v_*) = 0, \quad \forall \xi \in C_c^\infty(\mathbb{R}^d),$$

then there holds $\mathcal{A}_L(\mu_t, \tilde{M}_t) \leq \mathcal{A}_L(\mu_t, \tilde{M}_t + \eta)$*.*

Proof. We proceed exactly as in Dolbeault et al. [60, Theorem 5.17]. We start with the 'if' ($\Longleftarrow$) direction. Lemma 2.20 gives for $0 \leq r \leq s \leq T$

$$d_L(\mu_r, \mu_s) \leq \int_r^s \sqrt{\mathcal{A}_L(\mu_t, M_t)}dt < \infty.$$

The proof of Lemma 2.20 also says that $\sqrt{\mathcal{A}_L(\mu_t, M_t)} \in L^2(0,T)$. Hence μ is absolutely continuous with respect to d_L and by Lebesgue differentiation theorem, we also obtain

$$|\dot{\mu}|^2(t) \leq \mathcal{A}_L(\mu_t, M_t), \quad \text{a.e. } t \in [0,T].$$

For the 'only if' ($\Longrightarrow$) direction, assume μ is absolutely continuous. Lemma 1.1.4 of [6] says that there is an arc-length reparameterization of $t \mapsto \mu_t$ which is Lipschitz (this was similarly used in the proof of Lemma 2.20). So without loss of generality, assume that our original map $t \mapsto \mu_t$ is Lipschitz so that the metric derivative is bounded $|\dot{\mu}| \leq 1$. For fixed $N \in \mathbb{N}$ we take a mesh of $[0,T]$ with step size $\tau = 2^{-N}T$. Let $k = 1, \ldots, 2^N$. Recalling Definitions 2.2 and 2.3, $|\dot{\mu}|(t)$ can be thought of as the minimal m in the right-hand side of Definition 2.2 so that the following inequality holds

$$d_L(\mu_{(k-1)\tau}, \mu_{k\tau}) \leq \int_{(k-1)\tau}^{k\tau} |\dot{\mu}|(t)dt. \tag{2.11}$$

Squaring this inequality and using Cauchy-Schwarz gives

$$d_L^2(\mu_{(k-1)\tau}, \mu_{k\tau}) \leq \left(\int_{(k-1)\tau}^{k\tau} |\dot{\mu}|(t)dt\right)^2 \leq \tau \int_{(k-1)\tau}^{k\tau} |\dot{\mu}|^2(t)dt.$$

Finally, division by τ gives

$$\tau^{-1} d_L^2(\mu_{(k-1)\tau}, \mu_{k\tau}) \leq \int_{(k-1)\tau}^{k\tau} |\dot{\mu}|^2(t)dt. \tag{2.12}$$

Since $|\dot{\mu}| \in L^2(0,T)$, inequality (2.11) gives $d_L(\mu_{(k-1)\tau}, \mu_{k\tau}) < \infty$. Hence, by Proposition 2.21, we can find minimising curves $(\mu^{k,N}, M^{k,N}) \in GCE_T^E\left(\mu_{(k-1)\tau}, \mu_{k\tau}\right)$ such that (going back to Definition 2.19)

$$\tau \int_0^\tau \mathcal{A}_L(\mu_t^{k,N}, M_t^{k,N})dt = d_L^2(\mu_{(k-1)\tau}, \mu_{k\tau}).$$

Now we define $(\mu^N, M^N) \in GCE_T^E(\mu_0, \mu_T)$ as the concatenation of all the curves $(\mu^{k,N}, M^{k,N})$. This is done in a similar way as in the proof of the triangle inequality for Theorem 2.7. More explicitly, we define the curves

$$\mu_t^N := \begin{cases} \mu_t^{1,N}, & 0 \leq t \leq \tau \\ \mu_{t-\tau}^{2,N}, & \tau \leq t \leq 2\tau \\ \vdots & \vdots \\ \mu_{t-(2^N-1)\tau}^{2^N,N}, & (2^N-1)\tau \leq t \leq T \end{cases}, \quad M_t^N := \begin{cases} M_t^{1,N}, & 0 \leq t \leq \tau \\ M_{t-\tau}^{2,N}, & \tau \leq t \leq 2\tau \\ \vdots & \vdots \\ M_{t-(2^N-1)\tau}^{2^N,N}, & (2^N-1)\tau \leq t \leq T \end{cases}.$$

We compute the time integral of the square root of the action on this concatenated curve with the help of (2.12) and the Lipschitz assumption

$$\begin{aligned}\int_0^T \mathcal{A}_L(\mu_t^N, M_t^N)dt &= \sum_{k=1}^{2^N} \int_0^\tau \mathcal{A}_L(\mu_t^{k,N}, M_t^{k,N})dt = \sum_{k=1}^{2^N} \tau^{-1} d_L^2(\mu_{(k-1)\tau}, \mu_{k\tau}) \\ &\leq \sum_{k=1}^{2^N} \int_{(k-1)\tau}^{k\tau} |\dot{\mu}|^2(t)dt = \int_0^T |\dot{\mu}|^2(t)dt < \infty.\end{aligned}$$

This computation is independent of $N \in \mathbb{N}$ so taking the supremum, we have

$$\sup_{N \in \mathbb{N}} \int_0^T \mathcal{A}_L(\mu_t^N, M_t^N)dt < \infty.$$

As well, we have $(\mu_0^N)_{N\in\mathbb{N}} = (\mu_0^{1,N})_{N\in\mathbb{N}} = \{\mu_0\}$ which is tight as a singleton in $\mathscr{P}_{2,E}(\mathbb{R}^d)$. There is also a uniform second moment bound that is obtained directly from the minimising curves $(\mu^{k,N}, M^{k,N}) \in GCE_T^E\left(\mu_{(k-1)\tau}, \mu_{k\tau}\right)$. These conditions fulfill Proposition 2.18 so, up to a subsequence, there exists $(\tilde{\mu}, M) \in GCE_T^E(\mu_0, \mu_T)$ such that $\mu^N \rightharpoonup \tilde{\mu}$ and $M^N \overset{*}{\rightharpoonup} M$. One has that $\tilde{\mu}_t = \mu_t$ for every $t \in [0,T]$ because this equality holds for every $t = \frac{i}{2^N}T$ for every $N \in \mathbb{N}$ and $i = 1, \dots 2^N$. The union over all i, N of this dyadic partition of $[0,T]$ is dense, so by extension we have $\tilde{\mu}_t = \mu_t$. By lower semi-continuity, the previous calculations show

$$T^{-1}d_L^2(\mu_0, \mu_T) \leq \int_0^T \mathcal{A}_L(\mu_t, M_t)dt \leq \liminf_{N\to\infty} \int_0^T \mathcal{A}_L(\mu_t^N, M_t^N)dt \leq \int_0^T |\dot{\mu}|^2(t)dt.$$

Since $|\dot{\mu}|$ is minimal, we have $\mathcal{A}_L(\mu_t, M_t) = |\dot{\mu}|^2(t)$ for almost every $t \in [0,T]$. Uniqueness of the M satisfying this equality follows because the problem

$$\min_M \int_0^T \mathcal{A}_L(\mu_t, M_t)dt, \quad \text{s.t. } (\mu, M) \in GCE_T^E(\mu_0, \mu_T)$$

is a linear programming problem with a strictly convex (in M) objective function with a convex constraint on M (the grazing continuity equation).

For the statement regarding the tangent space, notice that whenever $\eta \in T_\mu^\perp$, if $(\mu, M) \in GCE_T^E(\mu_0, \mu_T)$, then also $(\mu, M+\eta) \in GCE_T^E(\mu_0, \mu_T)$ assuming that $\mathcal{A}_L(\mu, M+\eta) < +\infty$. To be precise, $\eta \in T_\mu^\perp$ means whenever $\xi \in C_c^\infty(\mathbb{R}^d)$, there holds

$$\iint_{\mathbb{R}^{2d}} \tilde{\nabla}\xi \cdot d\eta = 0.$$

The first order optimality condition of the minimality of $\tilde{M}$ is precisely

$$\mathcal{A}_L(\mu, \tilde{M}) \leq \mathcal{A}_L(\mu, \tilde{M} + \eta), \quad \forall \eta \in T_\mu^\perp.$$

□

2.3 Energy dissipation equality

The goal in this section is to prove Theorem 2.8 which states that the notions of gradient flow solutions coincide with ε-solutions to the Landau equation. To fix ideas, we recall the regularised entropy functionals $(\mathcal{H}_\varepsilon)$ acting on probability measures

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

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

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

The crucial ingredient to prove Theorem 2.8 is the following

Proposition 2.23 (Chain Rule ε)**.** *Fix $\gamma \in [-4, 0]$ and suppose $(\mu, M) \in GCE_T^E$ and*

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

Then, $\sup_{t\in[0,T]} \mathcal{H}_\varepsilon[\mu_t] < \infty$ and the 'chain rule' holds

$$\mathcal{H}_\varepsilon[\mu_r] - \mathcal{H}_\varepsilon[\mu_s] = \frac{1}{2}\int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\left[\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu}\right] \cdot dM_t dt, \quad \forall 0 \le s \le r \le T. \tag{2.13}$$

Remark 2.5. *Recall the expression for the dissipation*

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

Using a time integrated version of Lemma 2.17, we have the estimate

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

Therefore, under the hypothesis of Proposition 2.23, we have that

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

which implies that $D_\varepsilon(\mu_t)^{\frac{1}{2}}$ is a strong upper gradient of $\mathcal{H}_\varepsilon$, see Definition 2.4.

Taking Proposition 2.23 (and Lemma 2.26 in the sequel which depends only on $\varepsilon > 0$ and uniform second moments) for granted, we can prove Theorem 2.8.

Proof of Theorem 2.8. Throughout, $\mu = f\mathcal{L}$ is a curve of probability measures with uniformly bounded second moment.

Weak ε-solution $\implies$ Curve of maximal slope

Consider f an ε-solution to the Landau equation. Define $m = -ff_*\tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta f}$ so that the pair of measures $(\mu = f\mathcal{L}, M = m\mathcal{L}\otimes\mathcal{L})$ therefore belong to GCE_T^E. Indeed, the distributional grazing continuity equation from Definition 2.13 is precisely the weak ε Landau equation. Based on the definition of M

and the finite $\mathcal{H}_\varepsilon$ dissipation, we have the bound

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

which implies the weak continuity of μ. By Proposition 2.22, we have

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

Using Proposition 2.23, we have for any $0 \le s \le r \le T$

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

According to Definition 2.5, 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 $\mathcal{H}_\varepsilon$ with respect to the upper gradient $\sqrt{D_\varepsilon}$. Since μ is absolutely continuous with respect to d_L, Proposition 2.22 guarantees existence of a unique curve $M : t \in [0, T] \mapsto M_t \in \mathcal{M}_L$ such that $\int_0^T \sqrt{\mathcal{A}(\mu_t, M_t)}dt < \infty$ and $|\dot{\mu}|^2(t) = \mathcal{A}(\mu_t, M_t)$ a.e. $t \in [0, T]$. Furthermore, the pair $(\mu, M) \in GCE_T^E$. According to Lemma 2.15, let $M = m\mathcal{L} \otimes \mathcal{L}$ for some measurable function m. We apply the chain rule (2.13) with Cauchy-Schwarz and Young's inequalities with minus signs in the follow computations.

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

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 Cauchy-Schwarz and Young's inequalities, this implies that $\frac{m}{\sqrt{ff_*}} = -\sqrt{ff_*}\tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta f}$. As in the previous direction, the weak ε Landau equation coincides with the grazing continuity equation when m is equal to $-ff_*\tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta f}$. □

Remark 2.6. *Assuming Proposition 2.23 is true with $\mathcal{H}_\varepsilon$ replaced by $\mathcal{H}$, the proof of Theorem 2.10 follows almost identically. Achieving this by taking the very technical limit $\varepsilon \downarrow 0$ of* (2.17) *is the focus of Section 2.5.*

The rest of this section is devoted to proving Proposition 2.23. We need some preliminary estimates. The following result is a variation of [27, Lemma 2.6].

Lemma 2.24 (Carlen-Carvalho [27])**.** *Fix $\mu \in \mathscr{P}_{2,E}$ for some $E > 0$ and $\varepsilon > 0$. Then, there exists a constant $C = C(\varepsilon, E) > 0$ such that*

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

Proof. Starting with an upper bound, we easily see

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

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 $\varepsilon > 0$ small enough

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

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

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

At this point, we appeal to Chebyshev's inequality using the second moment of μ 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 $1 - \frac{E}{R^2} \geq \frac{1}{2}$ to uniformly lower bound the integral $\int_{|v'| \leq R} d\mu(v')$ away from 0 and then conclude the result after applying logarithms. □

Lemma 2.25 (log-derivative estimates)**.** *For fixed $\varepsilon > 0$ we have the formula*

$$\nabla G^\varepsilon(v) = -\frac{1}{\varepsilon}\left\langle \frac{v}{\varepsilon} \right\rangle^{-1} G^\varepsilon(v)\frac{v}{\varepsilon}. \tag{2.14}$$

For $\mu \in \mathscr{P}$, denoting $\partial^i = \frac{\partial}{\partial v^i}$ and $\partial^{ij} = \frac{\partial^2}{\partial v^i \partial v^j}$, we obtain

$$|\nabla \log(\mu * G^\varepsilon)(v)| \leq \frac{1}{\varepsilon}, \quad \left|\partial^{ij} \log(\mu * G^\varepsilon)(v)\right| \leq \frac{4}{\varepsilon^2}. \tag{2.15}$$

Proof. Equation (2.14) is a direct computation after noticing

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

The first order log-derivative estimate of (2.15) is calculated using formula (2.14) to obtain

$$\begin{aligned}|\nabla(\mu * G^\varepsilon)(v)| = |\mu * \nabla G^\varepsilon(v)| \le \frac{1}{\varepsilon}\int_{\mathbb{R}^d}\left\langle \frac{v-v'}{\varepsilon}\right\rangle^{-1}\left|\frac{v-v'}{\varepsilon}\right| G^\varepsilon(v-v')d\mu(v')\\ \le \frac{1}{\varepsilon}\int_{\mathbb{R}^d} G^\varepsilon(v-v')d\mu(v') = \frac{1}{\varepsilon}(\mu * G^\varepsilon)(v).\end{aligned}$$

For the second order derivatives, we first look at $\partial^{ij}\mu * G^\varepsilon$ which can be computed with the help of (2.14)

$$\begin{aligned}&|\partial^{ij}\mu * G^\varepsilon(v)| = \left|\partial^i\left(-\frac{1}{\varepsilon}\int_{\mathbb{R}^d}\left\langle\frac{v-v'}{\varepsilon}\right\rangle^{-1}\frac{v^j-v'^j}{\varepsilon}G^\varepsilon(v-v')d\mu(v')\right)\right| =\\ &\left|\frac{1}{\varepsilon^2}\int_{\mathbb{R}^d}\left(\left\langle\frac{v-v'}{\varepsilon}\right\rangle^{-3}\frac{v^i-v'^i}{\varepsilon}\frac{v^j-v'^j}{\varepsilon} - \delta^{ij}\left\langle\frac{v-v'}{\varepsilon}\right\rangle^{-1}\right.\right.\\ &\qquad\qquad\left.\left. + \left\langle\frac{v-v'}{\varepsilon}\right\rangle^{-2}\frac{v^i-v'^i}{\varepsilon}\frac{v^j-v'^j}{\varepsilon}\right)G^\varepsilon(v-v')d\mu(v')\right|\\ &\le \frac{3}{\varepsilon^2}\mu * G^\varepsilon(v).\end{aligned}$$

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

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

□

Lemma 2.26. *Fix $\varepsilon > 0$ 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 $\gamma \in [-2, 0]$:*

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

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

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

In particular, it holds

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

Proof. We first expand the expression for $\tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta\mu}$ to be used throughout this proof.

$$\begin{aligned}\tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta\mu} &= \tilde{\nabla} G^\varepsilon * \log(\mu * G^\varepsilon)(v,v_*) \\ &= |v-v_*|^{1+\frac{\gamma}{2}}\Pi[v-v_*](\nabla_v G^\varepsilon * \log(\mu * G^\varepsilon)(v) - \nabla_{v_*} G^\varepsilon * \log(\mu * G^\varepsilon)(v_*)) \\ &= |v-v_*|^{1+\frac{\gamma}{2}}\Pi[v-v_*]\int_{\mathbb{R}^d} G^\varepsilon(v')\left(\frac{\nabla\mu * G^\varepsilon}{\mu * G^\varepsilon}(v-v') - \frac{\nabla\mu * G^\varepsilon}{\mu * G^\varepsilon}(v_*-v')\right)dv'.\end{aligned} \tag{2.16}$$

1. <u>*Moderately soft case* $\gamma \in [-2,0]$*:*</u> We use (a concave version of) the triangle inequality (valid since $1+\frac{\gamma}{2} \geq 0$) and the first estimate of (2.15) to bound the last line of (2.16)

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

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

<u>$|v-v_*| \geq 1$*:*</u>

In the far field, we have $|v-v_*|^{1+\frac{\gamma}{2}} \leq 1$ hence we can brutally estimate (2.16) using again the first estimate of (2.15) to obtain the estimate

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

<u>$|v-v_*| \leq 1$*:*</u>

We can treat the singularity from the weight with a mean-value estimate and the second estimate of (2.15)

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

Inserting this into (2.16), we have

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

□

Remark 2.7. *Originally, I considered the general family of convolution kernels $G^{s,\varepsilon}$ described in Section 2.1.1. In particular, the choice $s = 2$ is very natural since these profiles are Maxwellians which are known to be stationary solutions to the Landau equation. However, it is not clear how one should prove the analogous estimates of Lemma 2.25 without assuming $\mu \in \mathscr{P}$ has compact support.*

The previous results will be used in the proof of Proposition 2.23.

Proof of Proposition 2.23. To prove (2.13), our strategy is to regularise the pair (μ, M) in time with parameter $\rho > 0$ and differentiate the regularisation. Then we obtain uniform bounds in ρ needed to take the limit $\rho \downarrow 0$.

<u>*Finite regularised entropy*</u>

We have the following chain of inequalities

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

The first inequality comes from Lemma 2.24 because $\log(\mu_t * G^\varepsilon)$ has linear growth (uniform in time) while in the second inequality, $\mu_t * G^\varepsilon$ has as many moments as μ_t with computable constants.

<u>*Time regularisation with $\rho > 0$*</u>

Without loss of generality, take μ the weakly time continuous representative (Lemma 2.11) and M the optimal grazing rate (Proposition 2.22) achieving the finite distance d_L. We first regularise the pair (μ, M) in time for a fixed parameter $\rho > 0$ as follows. Take a mollifier $\eta \in C_c^\infty(\mathbb{R})$ with the following properties

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

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

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

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

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

We derive equation (2.13) using this regularised grazing continuity equation. Consider

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

which we differentiate with respect to t by appealing to the Dominated Convergence Theorem. Firstly, due to the time regularisation, we have

$$\partial_t \left\{ (\mu_t^\rho * G^\varepsilon) \log(\mu_t^\rho * G^\varepsilon) \right\} = \left[(\partial_t \mu_t^\rho) * G^\varepsilon \right] (\log(\mu_t^\rho * G^\varepsilon) + 1).$$

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

$$\begin{aligned} &\left| \frac{1}{h} [(\mu_{t+h}^\rho * G^\varepsilon) \log\big(\mu_{t+h}^\rho * G^\varepsilon\big) - (\mu_t^\rho * G^\varepsilon) \log(\mu_t^\rho * G^\varepsilon)] \right| \\ &\leq \frac{1}{h} \left| (\mu_{t+h}^\rho * G^\varepsilon) - (\mu_t^\rho * G^\varepsilon) \right| \sup_{s \in [t, t+h]} \left| \log(\mu_s^\rho * G^\varepsilon) + 1 \right|. \end{aligned}$$

where we have used the Mean Value theorem with the chain rule. Applying Lemma 2.24, we obtain

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

We apply the Mean Value Theorem on the difference quotient again to get

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

Since μ has finite second order moments, this last expression belongs to L_v^1. By the Dominated Convergence Theorem,

$$\frac{d}{dt} \mathcal{H}_\varepsilon[\mu_t^\rho] = \int_{\mathbb{R}^d} \left[(\partial_t \mu_t^\rho) * G^\varepsilon \right] (\log(\mu_t^\rho * G^\varepsilon) + 1) dv = \int_{\mathbb{R}^d} (\partial_t \mu_t^\rho) \cdot [G^\varepsilon * \log(\mu_t^\rho * G^\varepsilon)] dv$$

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^ρ. Integrating in t, we obtain

$$\begin{aligned} \mathcal{H}_\varepsilon[\mu_r^\rho] - \mathcal{H}_\varepsilon[\mu_s^\rho] &= \int_s^r \int_{\mathbb{R}^d} (\partial_t \mu_t^\rho) \cdot [G^\varepsilon * \log(\mu_t^\rho * G^\varepsilon)] dv dt \\ &= \frac{1}{2} \int_s^r \iint_{\mathbb{R}^{2d}} [\tilde{\nabla} G^\varepsilon * \log(\mu_t^\rho * G^\varepsilon)] \cdot dM_t^\rho dt \\ &= \frac{1}{2} \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla} \frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t^\rho} \cdot dM_t^\rho dt. \end{aligned} \tag{2.17}$$

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

Estimates on the right-hand side of (2.17)*:*

According to Lemma 2.26, we have the estimate

$$\left| \tilde{\nabla} \frac{\delta \mathcal{H}_\varepsilon}{\delta \mu^\rho} \right| \lesssim_{\varepsilon, E} |v|^p + |v_*|^p, \quad 0 \leq p \leq 1.$$

By the first moment assumption of M_t (2.3), we have

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

This estimate also extends to M_t^ρ by definition of the mollification

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

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

<u>*Convergence for $\rho \downarrow 0$:*</u>

Firstly, we establish the following identity which will be useful later. For fixed functions f^1, f^2 we have

$$\begin{aligned} &\tilde{\nabla}[G^\epsilon * f^1] - \tilde{\nabla}[G^\epsilon * f^2] \\ &= |v - v_*|^{1+\frac{\gamma}{2}} \Pi[v - v_*](\nabla[G^\epsilon * f^1] - \nabla[G^\epsilon * f^2] - (\nabla_*[G^\epsilon * f^1]_* - \nabla_*[G^\epsilon * f^2]_*)) \\ &= |v - v_*|^{1+\frac{\gamma}{2}} \Pi[v - v_*] \int_{\mathbb{R}^d} (\nabla G^\epsilon(v - v') - \nabla G^\epsilon(v_* - v'))(f^1(v') - f^2(v')) dv'. \end{aligned} \tag{2.18}$$

Using the weak time continuity of μ, we can consider

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

The $\cdot$ stands for the convoluted variable. Since t belongs to a compact set, the function $t \mapsto \langle \mu_t, G^\varepsilon(v' - \cdot)) \rangle$ is *uniformly* continuous from the weak continuity of μ. In particular, using the continuity in v' and the lower bound from Lemma 2.24 we conclude that for any $R > 0$

$$|\log(\mu_t^\rho * G^\varepsilon) - \log(\mu_t * G^\varepsilon)| \to 0 \qquad \text{uniformly on } B_R. \tag{2.19}$$

Therefore by Lemma 2.24, abbreviating the weight to $w = |v - v_*|^{1+\frac{\gamma}{2}}$ and using (2.18) with $f^1 = \log(\mu_t^\rho * G^\varepsilon)$ and $f^2 = \log(\mu_t * G^\varepsilon)$, we have

$$\begin{aligned} &\left| \tilde{\nabla} \frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t^\rho} - \tilde{\nabla} \frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t} \right| = |\tilde{\nabla} G^\varepsilon * \log(\mu_t^\rho * G^\varepsilon)(v, v_*) - \tilde{\nabla} G^\varepsilon * \log(\mu_t * G^\varepsilon)(v, v_*)| \\ &\leq \int_{\mathbb{R}^d} w |\nabla G^\varepsilon(v - v') - \nabla G^\varepsilon(v_* - v')| \left| \log\big(\mu_t^\rho * G^\varepsilon(v')\big) - \log\big(\mu_t * G^\varepsilon(v')\big) \right| dv' \\ &\leq C_\varepsilon w \int_{B_{R_0}^c} |\nabla G^\varepsilon(v - v') - \nabla G^\varepsilon(v_* - v')| \langle v' \rangle \, dv' \\ &\quad + \left(\sup_{B_{R_0}} |\log(\mu_t^\rho * G^\varepsilon) - \log(\mu_t * G^\varepsilon)| \right) w \int_{B_{R_0}} |\nabla G^\varepsilon(v - v') - \nabla G^\varepsilon(v_* - v')| \, dv'. \end{aligned}$$

For a fixed (v, v_*), we obtain the convergence to zero by taking $\rho \to 0$ 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 $R > 0$

$$\left|\tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t^\rho}(v, v_*) - \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t}(v, v_*)\right| \to 0 \qquad \text{uniformly on } [0, T] \times B_R \times B_R. \tag{2.20}$$

We turn to the limit estimate for the right hand side of (2.17). For any $R > 0$, we have

$$\begin{aligned}
&\left|\int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t^\rho} \cdot dM_t^\rho dt - \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t} \cdot dM_t dt\right| \\
&\le \left|\int_s^r \iint_{\mathbb{R}^{2d}} \left(\tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t^\rho} - \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t}\right) \cdot dM_t^\rho dt\right| + \left|\int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t} \cdot dM_t^\rho dt - \int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t} \cdot dM_t dt\right| \\
&\le \int_s^r \iint_{B_R \times B_R} \left|\tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t^\rho} - \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t}\right| d|M_t^\rho| dt + \int_s^r \iint_{(B_R \times B_R)^C} \left|\tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t^\rho} - \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t}\right| d|M_t^\rho| dt + o(1).
\end{aligned}$$

The last term is $o(1)$ as $\rho \downarrow 0$ by the estimates in the previous step using the moment condition (2.3) of M and M^ρ. By sending $\rho \downarrow 0$ (the first term vanishes due to (2.20)) and then sending $R \to \infty$ (the second term vanishes again from the previous step), we obtain the convergence

$$\lim_{\rho \downarrow 0} \frac{1}{2}\int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t^\rho} \cdot dM_t^\rho dt = \frac{1}{2}\int_s^r \iint_{\mathbb{R}^{2d}} \tilde{\nabla}\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu_t} \cdot dM_t^\rho dt. \tag{2.21}$$

Convergence of the left-hand side of (2.17)

By (2.19), Lemma 2.24, and the uniform bound on the second moment, we have that

$$\begin{aligned}
|\mathcal{H}_\varepsilon[\mu_t^\rho] - \mathcal{H}_\varepsilon[\mu_t]| &\le \int_{\mathbb{R}^d} |(\mu_t^\rho * G^\varepsilon)\log(\mu_t^\rho * G^\varepsilon)(v) - (\mu_t * G^\varepsilon)\log(\mu_t * G^\varepsilon)(v)| dv \\
&\to 0, \quad \text{as } \rho \downarrow 0.
\end{aligned}$$

Therefore, by the previous equation and (2.21) we can take $\rho \downarrow 0$ in (2.17) to obtain

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

which is the desired result. □

2.4 JKO scheme for ε-Landau equation

This section is devoted to the proof of Theorem 2.9 after some preliminary results. Our construction of curves of maximal slope in Theorem 2.9 uses the minimising movement/variational approximation scheme of Jordan, Kinderlehrer, and Otto [90]. Fix a small time step $\tau > 0$ and initial datum $\mu_0 \in$

$\mathscr{P}_{2,E}(\mathbb{R}^d)$ and consider the recursive minimisation procedure for $n \in \mathbb{N}$

$$\nu_0^\tau := \mu_0, \qquad \nu_n^\tau \in \operatorname{argmin}_{\lambda \in \mathscr{P}_{2,E}} \left[\mathcal{H}_\varepsilon[\lambda] + \frac{1}{2\tau} d_L^2(\nu_{n-1}^\tau, \lambda) \right]. \tag{2.22}$$

Then, we concatenate these minimisers into a piecewise constant-in-time curve by setting

$$\mu_0^\tau := \mu_0, \qquad \mu_t^\tau := \nu_n^\tau, \text{ for } t \in ((n-1)\tau, n\tau]. \tag{2.23}$$

The scheme given by (2.22) and (2.23) satisfies the abstract formulation in [6] giving

Proposition 2.27 (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* (2.22)*. Furthermore, up to a subsequence of μ_t^τ described in* (2.23) *as $\tau \to 0$, there exists a locally absolutely continuous curve $(\mu_t)_{t\geq 0}$ such that*

$$\mu_t^\tau \rightharpoonup \mu_t, \qquad \forall t \in [0, \infty).$$

Proof. Our metric setting is $(\mathscr{P}_{\mu_0}, d_L)$ (see Theorem 2.7) 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 [61] which consists in verifying [6, Assumptions 2.1 a,b,c]. These assumptions are listed and verified now.

1. **$\mathcal{H}_\varepsilon$ 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 $\mu_n * G^\varepsilon \rightharpoonup \mu * G^\varepsilon$ 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 $\mathcal{H}_\varepsilon[\mu] = \mathcal{H}[\mu * G^\varepsilon]$, we achieve the first property.

2. **$\mathcal{H}_\varepsilon$ is lower bounded:** By Lemma 2.24 for fixed $\varepsilon > 0$, $\log(\mu * G^\varepsilon)$ 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

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

3. **d_L-bounded sets are relatively sequentially σ-compact:** This is one of the consequences from Theorem 2.7.

The existence of minimisers, ν_n^τ, to (2.22) and limits along subsequences as $\tau \downarrow 0$, μ_t, to (2.23) is guaranteed from [6, Corollary 2.2.2] and [6, Proposition 2.2.3], respectively. $\square$

At the abstract level, the limit curve constructed in Proposition 2.27 has no relation to $\sqrt{D_\varepsilon}$. The following lemmata bridge this gap.

Lemma 2.28. *For any $\mu_0 \in \mathscr{P}_2$, we have*

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

Proof. For fixed $\varepsilon, R_1, R_2 > 0$ and $\gamma \in \mathbb{R}$, take $T > 0$ from Theorem B.1 and the unique weak solution $\mu \in C([0,T]; \mathscr{P}_2(\mathbb{R}^d))$ to

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

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

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

The notation J_0^ε from Appendix B means

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

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

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

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

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

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

$$\sqrt{D_\varepsilon^{R_1,R_2}(\mu_0)} \leq |\partial \mathcal{H}_\varepsilon|(\mu_0), \quad \forall R_1,\, R_2 > 0.$$

Owing to the many regularisations applied, the ε-entropy-dissipation $\mu \mapsto D_\varepsilon^{R_1,\,R_2}(\mu)$ 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

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

As functions of R_1, R_2 individually, $D_\varepsilon^{R_1,R_2}(\mu_0)$ is non-decreasing. Furthermore, the integrand of $D_\varepsilon^{R_1,R_2}(\mu_0)$ converges to the integrand of $D_\varepsilon(\mu_0)$ 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. □

Lemma 2.29. *$|\partial^- \mathcal{H}_\varepsilon|$ is a strong upper gradient for $\mathcal{H}_\varepsilon$ 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 2.7, we have $d_L(\lambda, \nu) < \infty$. Now by Proposition 2.21, there exists a pair of curves $(\mu, M) \in GCE_1^E$ connecting λ, ν and $\mathcal{A}(\mu_t, M_t) = d_L^2(\lambda, \nu)$ for almost every $t \in [0, 1]$. Using Remark 2.5 and Lemma 2.28, we have

$$|\mathcal{H}_\varepsilon[\lambda] - \mathcal{H}_\varepsilon[\nu]| \leq \int_0^1 \sqrt{D_\varepsilon(\mu_t)}|\dot{\mu}|(t)dt \leq \int_0^1 |\partial^- \mathcal{H}_\varepsilon|(\mu_t)|\dot{\mu}|(t)dt.$$

□

We can now prove Theorem 2.9 so that we can relate curves of maximal slope to weak solutions of the ε-Landau equation.

Proof of Theorem 2.9. Take a limit curve μ_t constructed in Proposition 2.27. By the previous Lemma 2.29, the assumptions of [6, Theorem 2.3.3] are fulfilled so μ is a curve of maximal slope with respect to $|\partial^- \mathcal{H}_\varepsilon|$ and satisfies the associated energy dissipation inequality

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

The inequality of Lemma 2.28 gives

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

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

Remark 2.8. *The results of Proposition 2.27 and Lemma 2.28 can be generalised to other regularisation kernels $G^{s,\varepsilon}$, in particular, the Maxwellian regularisation. However, this is not the case for Lemma 2.29 since the proof relies on Proposition 2.23, see Remark 2.7.*

2.5 Recovering the full Landau equation as $\varepsilon \downarrow 0$

Theorems 2.8 and 2.9 provide the basic existence theory for the $\varepsilon > 0$ approximation of the Landau equation. In this section, we prove the $\varepsilon \downarrow 0$ analogue of Theorem 2.8 which is Theorem 2.10. Recall we fix $d = 3$ in this setting. 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 (**A2.2**) 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 (**A2.2**) will refer to this bound. Take the following claim for granted.

Claim 2.30. *Assume (**A2.1**), (**A2.2**), (**A2.3**) and let M be any grazing rate such that $(\mu, M) \in GCE_T^E$ and*

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

Then we have the chain rule

$$\mathcal{H}[\mu_r] - \mathcal{H}[\mu_s] = \frac{1}{2}\int_s^r \iint_{\mathbb{R}^6} \tilde{\nabla}\left[\frac{\delta\mathcal{H}}{\delta\mu}\right] \cdot dM_t dt. \tag{2.24}$$

Accepting Claim 2.30 on faith, we can prove Theorem 2.10.

Sketch of the proof of Theorem 2.10. Recalling the proof of Theorem 2.8, we see that the crucial ingredient is the chain rule (2.13) in Proposition 2.23. By following the steps of the proof of Theorem 2.8 and using (2.24) instead of (2.13), one completes the proof of Theorem 2.10. □

We dedicate this section to proving Claim 2.30 i.e. (2.24). Equation (2.24) is clearly the $\varepsilon \downarrow 0$ limit of (2.13). The left-hand side of (2.24) can be obtained from the left-hand side of (2.13) using the finite entropy assumption (**A2.2**) and exploiting the convolution structure of $\mathcal{H}_\varepsilon$.

The difficulty remains in deducing that the right-hand side of (2.13) converges to the right-hand side of (2.24) as $\varepsilon \downarrow 0$ given by

$$\int_0^T \iint_{\mathbb{R}^6} \tilde{\nabla} \frac{\delta \mathcal{H}_\varepsilon}{\delta \mu} \cdot dM_t dt \to \int_0^T \iint_{\mathbb{R}^6} \tilde{\nabla} \frac{\delta \mathcal{H}}{\delta \mu} \cdot dM_t dt, \quad \varepsilon \downarrow 0, \tag{2.25}$$

under the additional assumptions (**A2.1**)–(**A2.3**) on $\mu = f\mathcal{L}$. For this, we quote here an extended version of the Dominated Convergence Theorem which we will need from [106, Chapter 4, Theorem 17].

Theorem 2.31 (Extended Dominated Convergence Theorem (EDCT))**.** *Let $(H_\varepsilon)_{\varepsilon>0}$ and $(I_\varepsilon)_{\varepsilon>0}$ be sequences of measurable functions on a measurable space X satisfying $I_\varepsilon \geq 0$ and suppose there exists measurable functions H, I satisfying*

1. *$|H_\varepsilon| \leq I_\varepsilon$ for every $\varepsilon > 0$ and pointwise a.e.*
2. *H_ε and I_ε converge pointwise a.e. to H and I, respectively.*
3. $$\lim_{\varepsilon \downarrow 0} \int_X I_\varepsilon = \int_X I < \infty.$$

Then, we have the convergence

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

Setting $M = m\mathcal{L} \otimes \mathcal{L}$ (valid by Lemma 2.15) and using Young's inequality on the right-hand side of (2.13), we have the majorants

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

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 EDCT 2.31 with $X = (0,T) \times \mathbb{R}^6$ to prove (2.25) once we show

$$\int_0^T \iint_{\mathbb{R}^6} f f_* \left| \tilde{\nabla} \left[\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu} \right] \right|^2 dv_* dv dt \to \int_0^T \iint_{\mathbb{R}^6} f f_* \left| \tilde{\nabla} \left[\frac{\delta \mathcal{H}}{\delta \mu} \right] \right|^2 dv_* dv dt, \quad \varepsilon \downarrow 0. \tag{2.26}$$

The pointwise a.e. convergence hypothesis of EDCT 2.31 is straightforward based on the regularisation of $\mathcal{H}_\varepsilon$ through G^ε. Focusing on (2.26), we will use a standard Dominated Convergence Theorem (DCT) for the integration in the t variable, by proving

$$\begin{aligned}
&\iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu}\right]\right|^2 dv_* dv \to \iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}}{\delta \mu}\right]\right|^2 dv_* dv, \quad \text{a.e. } t, \\
&\iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}_\varepsilon}{\delta \mu}\right]\right|^2 dv_* dv \leq C \iint_{\mathbb{R}^6} \frac{1}{2} f f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}}{\delta \mu}\right]\right|^2 dv_* dv, \quad \text{a.e. } t \quad \forall \varepsilon > 0,
\end{aligned} \tag{2.27}$$

where $C > 0$ is a constant independent of $\varepsilon > 0$. The estimate of (2.27) guarantees the L_t^1 majorisation due to the finite entropy-dissipation assumption (**A2.3**). Our estimates in this section accomplish both the convergence and the estimate of (2.27) by nested application of EDCT 2.31. The significance of all three assumptions (**A2.1**)–(**A2.3**) will be apparent in proving the convergence in (2.27).

Remark 2.9. *In this section, the only properties of G^ε we use are that it is a non-negative radial approximate identity with sufficiently many moments. As in the construction of minimising movement curves in Section 2.4, the results of this section can be achieved with other radial approximate identities.*

2.5.1 Outline of technical strategy to prove (2.27)

The need to apply EDCT 2.31 instead of the more classical Lebesgue DCT is because we are unable to find direct L_v^1 majorants for the function $v \to f \int_{\mathbb{R}^3} f_* \left|\tilde{\nabla}\left[\frac{\delta \mathcal{H}_\varepsilon}{\delta f}\right]\right|^2 dv_*$. The key observations that allow us to prove (2.27) are that we do not need our majorants to be integrable and moreover, we can rely on the self-adjointness of convolution against radial exponentials (which I hereafter refer to as SACRE) to reveal a convergent majorant in ε. The unofficial name 'SACRE' is quoted for simplicity although, in principle, it is a manifestation and exchange of weak-strong convergence as detailed below.

Step 1: Find majorants and appeal to EDCT 2.31

We seek to find pointwise a.e. majorants in the v variable

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

where $I_\varepsilon^1(v)$ satisfies the hypothesis for the majorant in EDCT 2.31. We then show that I_ε^1 converges pointwise to some I^1. This is possible 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 Item 3 of EDCT 2.31

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

Step 2: Use SACRE with G^ε

To show the integral convergence for I^1_ε, we find functions A^1 and B^1 such that

$$I^1_\varepsilon(v) \leq A^1(v)(G^\varepsilon * B^1)(v)$$

and apply EDCT 2.31. As in the previous step, the pointwise convergence is easily proved. Hence, we need to verify the convergence of the integrals

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

The key observation is applying SACRE to obtain

$$\int_{\mathbb{R}^3} A^1(G^\varepsilon * B^1) = \int_{\mathbb{R}^3} \underbrace{(G^\varepsilon * A^1)B^1}_{=:I^2_\varepsilon}.$$

Therefore, we have reduced the problem to showing integral convergence Item 3 of EDCT for I^2_ε (as the pointwise convergence is again easily proved).

<u>*Step 3: Reiterate step 2*</u>

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

$$I^2_\varepsilon(v) \leq A^2(v)(G^\varepsilon * B^2)(v).$$

Again the pointwise convergence for the majorant follows easily, hence we only need to check the integral convergence Item 3 of EDCT 2.31 given by

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

Using SACRE, we study instead the integral convergence of

$$I^3_\varepsilon(v) = (G^\varepsilon * A^2)B^2.$$

Eventually, after a finite number of times of finding majorants and applying SACRE, we will obtain a majorant I^i_ε for which the estimates and the convergence as $\varepsilon \to 0$ follows from the standard Lebesgue DCT, using the bound of the weighted Fisher information in terms of the entropy-dissipation (see Theorem 2.32) and assumption (**A2.3**). Although I^i_ε and I^j_ε are not necessarily directly comparable for $i \neq j$, the only property we are concerned with proving is $\int I^i_\varepsilon \to \int I$.

The key point of this strategy is that while $A^1(G^\varepsilon * B^1)$ may not have an integrable majorant, it is enough to study $(G^\varepsilon * A^1)B^1$ for which integrable majorants are more easily available.

2.5.2 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 one of my co-authors in this project, Laurent Desvillettes [52, 53].

Theorem 2.32 (Desvillettes [52, 53])**.** *Suppose* $\gamma \in (-4, 0]$ *and let* $f \geq 0$ *be a probability density belonging to* $L^1_{2-\gamma} \cap L\log L(\mathbb{R}^3)$*. We have*

$$\int_{\mathbb{R}^3} f(v)\langle v\rangle^\gamma |\nabla \log f|^2\, dv + \int_{\mathbb{R}^3} f(v)\, \langle v\rangle^\gamma\, |v \times \nabla \log f|^2\, dv \leq C(1 + D_L(f)),$$

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 [54, 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 (**A2.2**) enters in the sequel since the constant $C > 0$ in Theorem 2.32 depends on bounds for $\mathcal{H}[f]$. To decompose the entropy-dissipation in a manageable way that makes the cross Fisher term more apparent, notice that Lemma A.1 allows

$$|v|^2 |\Pi[v]\nabla \log f|^2 = |v \times \nabla \log f|^2.$$

The following lemma shows how we use assumption (**A2.1**) to control the singularity of the weight.

Lemma 2.33. *Given* $\gamma \in (-3, 0]$*, assume that* f *satisfies (**A2.1**) for some* $0 < \eta \leq \gamma + 3$*, then we have for a.e.* t

$$\int_{\mathbb{R}^3} f_*(t)|v - v_*|^\gamma dv_* \leq C_1(t)\, \langle v\rangle^\gamma\,, \quad \int_{\mathbb{R}^3} f_*(t)|v_*|^2|v - v_*|^\gamma dv_* \leq C_2(t)\, \langle v\rangle^\gamma\,, \tag{2.28}$$

where

$$\begin{aligned} \|C_1\|_{L^\infty(0,T)} &\lesssim_{\gamma,\eta} \|\langle\cdot\rangle^{-\gamma} f(t)\|_{L^\infty\left(0,T;L^1\cap L^{\frac{3-\eta}{3+\gamma-\eta}}(\mathbb{R}^3)\right)} \\ \|C_2\|_{L^\infty(0,T)} &\lesssim_{\gamma,\eta} \|\langle\cdot\rangle^{2-\gamma} f(t)\|_{L^\infty\left(0,T;L^1\cap L^{\frac{3-\eta}{3+\gamma-\eta}}(\mathbb{R}^3)\right)}. \end{aligned}$$

Remark 2.10. *The inequalities of* (2.28) *are essential in the proof of* (2.27)*. The assumption (**A3.1**) enters as a sufficient condition to prove both inequalities of* (2.28)*. In the case* $\gamma = -3$*, the integrals in* (2.28) *are not well-defined, since* $v \mapsto |v|^{-3}$ *is too singular in* $\mathbb{R}^3$*. It may be possible to consider the Riesz transforms of* $f_*(t)$ *and* $f_*(t)|v_*|^2$ *as an alternative to the integrals in* (2.28)*. The overall* *structure in the proof of* (2.27) *would, however, be completely different in this case. As we shall see in Section 2.5.3, we really manipulate integrals of the form* (2.28)*. For $\gamma = -3$, it is not entirely clear how to work with the Riesz transforms (up to a multiplicative constant)*

$$\text{p.v.} \int_{\mathbb{R}^3} f_*(t) \frac{(v - v_*)}{|v - v_*|^4} dv_*, \quad \text{p.v.} \int_{\mathbb{R}^3} f_*(t) |v_*|^2 \frac{(v - v_*)}{|v - v_*|^4} dv_*,$$

instead of the integrals in (2.28)*.*

Proof. We will only prove the first inequality of (2.28) since the second inequality uses the same procedure. We split the estimation for local $|v| \leq 1$ and far-field $|v| \geq 1$.

<u>$|v| \leq 1$:</u>

We split the integral over v_* into two regions

$$\begin{aligned} \int_{\mathbb{R}^3} f_* |v - v_*|^\gamma dv_* &= \int_{|v - v_*| \geq 1} f_* |v - v_*|^\gamma dv_* + \int_{|v - v_*| \leq 1} f_* |v - v_*|^\gamma dv_* \\ &\leq 1 + \int_{|v - v_*| \leq 1} f_* |v - v_*|^\gamma dv_*, \end{aligned}$$

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

$$\int_{|v - v_*| \leq 1} f_* |v - v_*|^\gamma dv_* \leq \|f * (\chi_{B_1} |\cdot|^\gamma)\|_{L^\infty} \leq \|f\|_{L^{\frac{3-\eta}{3+\gamma-\eta}}} \|\chi_{B_1} |\cdot|^\gamma\|_{L^{\frac{-3+\eta}{\gamma}}} \leq \left(\frac{\omega_2}{\eta}\right)^{\frac{-3+\eta}{\gamma}} \|f\|_{L^{\frac{3-\eta}{3+\gamma-\eta}}}.$$

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

<u>$|v| \geq 1$:</u>

Once again, we split the integral into two parts

$$\begin{aligned} \int_{\mathbb{R}^3} f_* |v - v_*|^\gamma dv_* &= \int_{|v_*| \leq \frac{1}{2}|v|} f_* |v - v_*|^\gamma dv_* + \int_{|v_*| \geq \frac{1}{2}|v|} f_* |v - v_*|^\gamma dv_* \\ &\leq 2^{-\gamma} |v|^\gamma \int_{|v_*| \leq \frac{1}{2}|v|} f_* dv_* + 2^{-\gamma} |v|^\gamma \int_{|v_*| \geq \frac{1}{2}|v|} f_* |v_*|^{-\gamma} |v - v_*|^\gamma dv_*. \end{aligned}$$

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

$$|v - v_*| \geq |v| - |v_*| \geq \frac{1}{2}|v|, \quad 1 \leq 2^{-\gamma} |v|^\gamma |v_*|^{-\gamma}.$$

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

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

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

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

Lemma 2.34. *For any $p \in \mathbb{R}$, we have*

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

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

Proof. We use Peetre's inequality in Lemma A.2 to introduce $v-w$ into the angle brackets

$$\begin{aligned}
&\int_{\mathbb{R}^d} \langle w\rangle^p G^\varepsilon(v-w)dw \le 2^{|p|/2}\langle v\rangle^p \int_{\mathbb{R}^d} \langle v-w\rangle^{|p|} G^\varepsilon(v-w)dw \\
&= 2^{|p|/2}\langle v\rangle^p \int_{\mathbb{R}^d} (1+|w|^2)^{\frac{|p|}{2}} \varepsilon^{-d} G(w/\varepsilon)dw = 2^{|p|/2}\langle v\rangle^p \int_{\mathbb{R}^d} (1+\varepsilon^2|w|^2)^{\frac{|p|}{2}} G(w)dw \\
&\le C_{|p|}\langle v\rangle^p \left[1+\varepsilon^{|p|}\int_{\mathbb{R}^d} |w|^{|p|}G(w)dw\right] \le C_{|p|}\left[1+\varepsilon^{|p|}m_{|p|}(G)\right]\langle v\rangle^p
\end{aligned}$$

□

We stress that Peetre's inequality in Lemma A.2 is necessary for the estimate of Lemma 2.34 with *non-positive* 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 2.35 (Twisted integration by parts)**.** *Let $f, g \in C_c^\infty(\mathbb{R}^3)$. 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 given in components by

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

2.5.3 Proof of (2.27) using EDCT 2.31

We start by decomposing and estimating the integrand of D_ε. With the help of Lemma A.1, we expand the square term of the integrand to see

$$\begin{aligned}\left|\tilde{\nabla}\left[\frac{\delta\mathcal{H}_\varepsilon}{\delta\mu}\right]\right|^2 &= |v-v_*|^{2+\gamma}|\Pi[v-v_*](b^\varepsilon * a^\varepsilon - b^\varepsilon * a^\varepsilon_*)|^2 \\ &\leq |v-v_*|^\gamma(4|v\times(b^\varepsilon * a^\varepsilon)|^2 + 4|v_*\times(b^\varepsilon * a^\varepsilon_*)|^2 \\ &\qquad + 4|v\times(b^\varepsilon * a^\varepsilon_*)|^2 + 4|v_*\times(b^\varepsilon * a^\varepsilon)|^2) \\ &\leq 4|v-v_*|^\gamma \underbrace{|v\times(b^\varepsilon * a^\varepsilon)|^2}_{①} + 4|v-v_*|^\gamma \underbrace{|v_*\times(b^\varepsilon * a^\varepsilon_*)|^2}_{②} \\ &\qquad + 4|v|^2|v-v_*|^\gamma \underbrace{|b^\varepsilon * a^\varepsilon_*|^2}_{③} + 4|v_*|^2|v-v_*|^\gamma \underbrace{|b^\varepsilon * a^\varepsilon|^2}_{④},\end{aligned}$$

where we use the shorthand notation

$$b^\varepsilon = G^\varepsilon \qquad \text{and} \qquad a^\varepsilon = \nabla\log(G^\varepsilon * f). \tag{2.29}$$

Since 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 $\varepsilon \downarrow 0$. As well, each ⓘ for $i = 1,2,3,4$ converge pointwise a.e. to

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

By EDCT 2.31, to show the integral convergence in (2.27), it suffices to show, for example,

$$\iint_{\mathbb{R}^6} ff_*|v-v_*|^\gamma ①\, dvdv_* \to \iint_{\mathbb{R}^6} ff_*|v-v_*|^\gamma \frac{|v\times\nabla f|^2}{f^2} dvdv_*,$$

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

2.5.3.1 Term ④

We seek to show in the limit $\varepsilon \downarrow 0$,

$$\begin{aligned}&\iint_{\mathbb{R}^6} ff_*|v_*|^2|v-v_*|^\gamma|b^\varepsilon * a^\varepsilon|^2 dv_* dv = \int_{\mathbb{R}^3}\left(\int_{\mathbb{R}^3} f_*|v_*|^2|v-v_*|^\gamma dv_*\right) f|b^\varepsilon * a^\varepsilon|^2 dv \\ &\to \int_{\mathbb{R}^3}\left(\int_{\mathbb{R}^3} f_*|v_*|^2|v-v_*|^\gamma dv_*\right)\frac{|\nabla f|^2}{f} dv.\end{aligned} \tag{2.30}$$

By the reordering of integrations written above, we now think of the double integral over v, v_* of $ff_*|v_*|^2|v-v_*|^\gamma|b^\varepsilon * a^\varepsilon|^2$ as a single integral of the function $\left(\int_{\mathbb{R}^3} f_*|v_*|^2|v-v_*|^\gamma dv_*\right) f|b^\varepsilon * a^\varepsilon|^2$ over v. To be precise, we wish to apply Theorem 2.31 with $X = \mathbb{R}^3$ and $H_\varepsilon = (\int_{\mathbb{R}^3} f_*|v_*|^2|v-v_*|^\gamma dv_*) f|b^\varepsilon * a^\varepsilon|^2$. We can use Cauchy-Schwarz on the convolution integral to absorb the power term as follows

$$\begin{aligned}
|b^\varepsilon * a^\varepsilon|^2 = \left|\int_{\mathbb{R}^3} b^\varepsilon(v-w)a^\varepsilon(w)dw\right|^2 &\leq \left(\int_{\mathbb{R}^3} \langle w\rangle^{-\gamma} b^\varepsilon(v-w)dw\right)\left(\int_{\mathbb{R}^3} b^\varepsilon(v-w)\langle w\rangle^\gamma |a^\varepsilon(w)|^2 dw\right) \\
&\leq C\langle v\rangle^{-\gamma} b^\varepsilon * [\langle\cdot\rangle^\gamma |a^\varepsilon(\cdot)|^2],
\end{aligned}$$

where the last inequality comes from Lemma 2.34. Continuing with Lemma 2.33, we have

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

By EDCT 2.31, we reduce the problem to showing in the limit $\varepsilon \downarrow 0$

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

This is where we use SACRE, Step 2 of our general strategy 2.5.1. Application of SACRE and further simplification using the specific forms of a^ε and b^ε (see (2.29)) yields

$$\int_{\mathbb{R}^3} fb^\varepsilon * [\langle\cdot\rangle^\gamma|a^\varepsilon|^2]dv = \int_{\mathbb{R}^3}[b^\varepsilon * f]\langle v\rangle^\gamma |a^\varepsilon|^2 dv = \int_{\mathbb{R}^3}\langle v\rangle^\gamma \frac{|b^\varepsilon * \nabla f|^2}{b^\varepsilon * f}dv. \tag{2.31}$$

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

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

Next, we notice that the function $\beta:\ (F, f) \mapsto \frac{|F|^2}{f}$ is jointly convex in $F \in \mathbb{R}^3$ and $f > 0$ (see Lemma 2.14), so we can use Jensen's inequality with $b^\varepsilon = G^\varepsilon$ as the reference probability measure to obtain a further pointwise majorant for the integrand of (2.31)

$$\begin{aligned}
&\frac{|b^\varepsilon * \nabla f|^2}{b^\varepsilon * f}(v) = \beta(b^\varepsilon * \nabla f,\, b^\varepsilon * f)(v) = \beta\left(\int_{\mathbb{R}^d}\nabla f(v-y)b^\varepsilon(y)dy,\, \int_{\mathbb{R}^d} f(v-y)b^\varepsilon(y)dy\right) \\
&\leq \int_{\mathbb{R}^d}\beta(\nabla f(v-y), f(v-y))b^\varepsilon(y)dy = \int_{\mathbb{R}^d}\frac{|\nabla f(v-y)|^2}{f(v-y)}b^\varepsilon(y)dy = b^\varepsilon * \left[\frac{|\nabla f|^2}{f}\right](v).
\end{aligned}$$

Using EDCT 2.31 again, we reduce the problem to showing in the limit $\varepsilon \downarrow 0$

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

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

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

Now, we are in a position to apply the classical Dominated Convergence Theorem. We notice that we have the pointwise convergence

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

Furthermore, using Lemma 2.34, we can estimate $b^\varepsilon * \langle\cdot\rangle^\gamma$ uniformly in ε to find the domination

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

Using Theorem 2.32, the finite entropy-dissipation assumption (**A2.3**), and uniformly bounded entropy (**A2.2**) (remember the constant in Theorem 2.32 depends also on bounds for the entropy) we know that the right-hand side belongs to L^1_v 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

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

We have closed the argument for the convergence of (2.30) after retracing the previous estimates with EDCT 2.31.

2.5.3.2 Term ①

We seek to show in the limit $\varepsilon \downarrow 0$,

$$\begin{aligned} \iint_{\mathbb{R}^6} f f_* |v - v_*|^\gamma |v \times (b^\varepsilon * a^\varepsilon)|^2 dv_* dv &= \int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} f_* |v - v_*|^\gamma dv_* \right) f |v \times (b^\varepsilon * a^\varepsilon)|^2 dv \\ \to \int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} f_* |v - v_*|^\gamma dv_* \right) \frac{|v \times \nabla f|^2}{f} dv \end{aligned} \tag{2.32}$$

using the same strategy of nested applications of EDCT 2.31 like in the previous Section 2.5.3.1. We will encounter difficulty when trying to use Jensen's inequality due to the cross Fisher information term. As in the previous Section 2.5.3.1, we have written this double integral over v, v_* as a single integral over v. By EDCT 2.31 and Lemma 2.33, it suffices to show the integral convergence of

$$\int_{\mathbb{R}^3} \langle v\rangle^\gamma f |v \times (b^\varepsilon * a^\varepsilon)|^2 dv \to \int_{\mathbb{R}^3} \langle v\rangle^\gamma \frac{|v \times \nabla f|^2}{f} \tag{2.33}$$

to obtain the integral convergence of (2.32). Pointwise, we can make the following manipulations

$$\begin{aligned} v \times (b^\varepsilon * a^\varepsilon) &= v \times \left(\int_{\mathbb{R}^3} G^\varepsilon(v-w) \nabla \log(f * G^\varepsilon(w)) dw \right) \\ &= v \times \left(\int_{\mathbb{R}^3} \nabla G^\varepsilon(v-w) \log(f * G^\varepsilon(w)) dw \right) = \int_{\mathbb{R}^3} w \times \nabla G^\varepsilon(v-w) \log(f * G^\varepsilon(w)) dw \\ &= \int_{\mathbb{R}^3} G^\varepsilon(v-w) w \times \nabla \log(f * G^\varepsilon(w)) dw, \end{aligned} \tag{2.34}$$

where we have used the radial symmetry of G^ε to get the cancellation $(v-w) \times \nabla G^\varepsilon(v-w) = 0$ and the twisted integration by parts Lemma 2.35 (note that we do not pick up any signs in the integration by parts since the variable w appears with a minus sign in the arguments of G^ε).

We apply Cauchy-Schwarz, multiply and divide by $\langle w \rangle^\gamma$, and use Lemma 2.34 to obtain

$$\begin{aligned} |v \times (b^\varepsilon * a^\varepsilon)|^2 &\leq \left(\int_{\mathbb{R}^3} G^\varepsilon(v-w) \langle w \rangle^{-\gamma} dw \right) \left(\int_{\mathbb{R}^3} G^\varepsilon(v-w) \langle w \rangle^{\gamma} \left| w \times \frac{\nabla f * G^\varepsilon(w)}{f * G^\varepsilon(w)} \right|^2 dw \right) \\ &\lesssim_\gamma \langle v \rangle^{-\gamma} \left(\int_{\mathbb{R}^3} G^\varepsilon(v-w) \langle w \rangle^{\gamma} \left| w \times \frac{\nabla f * G^\varepsilon(w)}{f * G^\varepsilon(w)} \right|^2 dw \right). \end{aligned}$$

Remembering that this majorant holds pointwise on the integrand of (2.33), we multiply by $\langle v \rangle^\gamma f(v)$ and obtain

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

Now, we recognise a convolution inside the brackets. Hence, using SACRE we can re-write

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

Using EDCT 2.31, 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 2.36.

$$\frac{|v \times \nabla f * G^\varepsilon(v)|^2}{f * G^\varepsilon(v)} \leq \int_{\mathbb{R}^3} G^\varepsilon(v-w) \frac{|w \times \nabla f(w)|^2}{f(w)} dw. \tag{2.35}$$

Proof of Claim 2.36. We start by repeating a similar argument to (2.34). Using that G^ε is radially symmetric and the twisted integration by parts Lemma 2.35 we obtain

$$\begin{aligned} v \times \nabla f * G^{\varepsilon}(v) &= v \times \left(\int_{\mathbb{R}^3} \nabla G^{\varepsilon}(v-w) f(w) dw \right) \\ &= \int_{\mathbb{R}^3} w \times \nabla G^{\varepsilon}(v-w) f(w) dw \\ &= \int_{\mathbb{R}^3} G^{\varepsilon}(v-w) \underbrace{(w \times \nabla_w f(w))}_{=:F(w)} dw. \end{aligned}$$

Therefore, since $(F, f) \mapsto \frac{|F|^2}{f}$ is jointly convex in $F \in \mathbb{R}^3$ and $f > 0$ (again, see Lemma 2.14), we apply Jensen's inequality to the left-hand side of (2.35) as in Section 2.5.3.1 to see

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

which proves the claim. □

Continuing, by EDCT 2.31, we seek to establish the convergence of the integral

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

Finally, the integrand of the right-hand side has a majorant due to Lemma 2.34

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

Once again using Theorem 2.32 and (**A2.3**) and (**A2.2**), we obtain that the right hand side belongs to $L^1_v(\mathbb{R}^3)$ a.e. $t \in (0, T)$. Using the usual Dominated Convergence theorem, we have established the convergence of the integral. Tracing back the estimates, this takes care of the convergence of the term ① and establishes the convergence in (2.33).

We note that the estimates in the previous subsections not only establish the a.e. pointwise convergence of (2.27), but also the majorisation

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

where

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

by Lemma 2.33. Hence, using assumption (**A2.3**) and (2.27) we can apply Lebesgue DCT to pass to the limit in the time integral and finally prove the desired chain rule in Claim 2.30.

Chapter 3

Gradient Flow perspective of the grazing collision limit

The content of this chapter is based on joint work with José A. Carrillo and Matias G. Delgadino. It has been published [32] in Nonlinear Analysis volume 219, page 112824 in June 2022.

This chapter recovers the grazing collision limit from the Boltzmann equation to the Landau equation using the gradient flow machinery developed previously in Chapter 2. In Section 1.1, I outlined convergence from (EDE_ϵ) to (EDE_L). However, from the finite dimensional example considered in Section 1.2, it was important to include the metric derivative. This is incorporated into the energy dissipation *inequality* (see Definition 3.1 below and compare with Definition 2.5). The statement of the main result can then be made shortly after in Section 3.1. Section 3.2 covers the main assumptions; in particular, the scaling associated to the grazing collision limit is discussed in Section 3.2.1. Some visualisations of the geometry of collisions are given in Section 3.2.2 and a coordinate system is introduced which makes the grazing collision limit more tractable. The original computations are recalled at the formal level in Section 3.2.3 The technical proofs are deferred to Sections 3.4 to 3.6 where a novel convex analytic form of the dissipations and metric derivatives is exploited.

3.1 Main Result

In the following definition, we refer to the Boltzmann and Landau metric derivatives. We have already seen what the Landau metric and metric derivative look like from Section 2.2. The Boltzmann metric, d_ϵ, and metric derivative, $|\dot{f}|_\epsilon$, are defined similarly and will be recalled later in Section 3.3. For now, these objects should be taken for granted subordinate to the collision kernel B^ϵ for fixed $\epsilon > 0$ (c.f. assumption (**A3.2**) and the discussion in Section 3.2.1). I will use here $\epsilon > 0$ to distinguish with $\varepsilon > 0$ used previously in Chapter 2. As a preview, B^ϵ is generated from a fixed kernel B in the form (1.1) by the scaling

$$\sin\theta\, B^\epsilon(r,\theta) = \frac{\pi}{\epsilon^3}\sin\left(\frac{\pi\theta}{\epsilon}\right) B\left(r, \frac{\pi\theta}{\epsilon}\right).$$

Definition 3.1 (H-gradient flows for Boltzmann and Landau)**.** For $\epsilon > 0$ and $T > 0$, we say that an absolutely continuous curve $f^\epsilon : t \in [0,T] \mapsto L^1_+(\mathbb{R}^3)$ with respect to the Boltzmann metric d_ϵ is an *H-gradient flow* solution to the Boltzmann equation associated to the kernel B^ϵ if the Energy Dissipation Inequality (abbreviated EDI) holds for every $t \in [0,T]$

$$\mathcal{H}[f^\epsilon(t)] + \frac{1}{2}\int_0^t D^\epsilon_B(f^\epsilon(s))\, ds + \frac{1}{2}\int_0^t |\dot{f}^\epsilon|^2_\epsilon(s)\, ds \le \mathcal{H}[f^\epsilon(0)] < \infty, \qquad (EDI_\epsilon)$$

(recall D^ϵ_B from (EDE_ϵ)) and it dissipates the second moment

$$\int_{\mathbb{R}^3} |v|^2 f^\epsilon_t(v)dv \le \int_{\mathbb{R}^3} |v|^2 f^\epsilon_s(v)dv < \infty, \quad \forall\, 0 \le s \le t \le T. \qquad (3.1)$$

Likewise, for $T > 0$ an absolutely continuous curve $f : t \in [0,T] \mapsto L^1_+(\mathbb{R}^3)$ with respect to the Landau metric is an *H-gradient flow* solution to the Landau equation if the Energy Dissipation Inequality holds for every $t \in [0,T]$

$$\mathcal{H}[f(t)] + \frac{1}{2}\int_0^t D_L(f(s))\, ds + \frac{1}{2}\int_0^t |\dot{f}|^2_L(s)\, ds \le \mathcal{H}[f(0)] < \infty,$$

and it dissipates the second moment as in (3.1) replacing f^ϵ_t by f_t.

Remark 3.1. *The notion of H-gradient flow is strictly weaker than the notion of curves of maximal slope introduced in [6]. More specifically, we do not require here that the dissipations D^ϵ_B and D_L to be strong upper gradients, let alone that the chain rule holds. Recall this was the content of Section 2.5 in Chapter 2 in the case of Landau. The focus here is to dispense with the very technical assumption (**A2.1**) from Theorem 2.10 and focus only on the physically relevant mechanisms in the grazing collision limit.*

Remark 3.2. *The more classical notion of renormalized solutions (for Boltzmann or Landau) are weak solutions that also dissipate entropy*

$$\mathcal{H}[f(t)] + \int_0^t D(f(t)) \le \mathcal{H}[f_0], \quad D = D_B \text{ or } D_L.$$

It can be checked that renormalized solutions are also H-gradient flow solutions, see Remark 3.11. The existence of renormalized solutions (and hence H-gradient flow solutions) can be shown subject to boundedness assumptions on the initial data, see [4, Corollary 2.1 and Appendix].

For H-gradient flows, the initial entropy controls the entropy at later times as well as the integrability of the dissipation and the metric derivative. We therefore consider H-gradient flows of Boltzmann, f^ϵ, subject to the following assumptions.

(**A3.1**) For every $\epsilon > 0$, we assume that the initial probability densities $f^\epsilon(0) = f_0^\epsilon$ converge in the weak topology to some probability density f_0. Furthermore, we assume a uniform second moment bound and convergence in entropy

$$\sup_{\epsilon>0} \int_{\mathbb{R}^3} (1+|v|^2) f_0^\epsilon(v) dv < +\infty, \quad \mathcal{H}[f_0^\epsilon] \overset{\epsilon\downarrow 0}{\to} \mathcal{H}[f_0] < +\infty.$$

(**A3.2**) There exists $\gamma \in [-4, 0)$, such that the ϵ-collision kernel satisfies

$$B^\epsilon(r,\theta)\sin\theta = r^\gamma \beta^\epsilon(\theta),$$

where

$$\beta^\epsilon(\theta) = \frac{\pi^3}{\epsilon^3}\beta\left(\frac{\pi\theta}{\epsilon}\right), \quad \theta \in (0, \epsilon/2).$$

The function β satisfies

$$\sup_{\theta\in[\delta,\pi/2]} \beta(\theta) < +\infty \quad \forall \delta > 0, \quad \operatorname{supp}\beta \in [0, \pi/2],$$

and that there exists $\nu \in (0, 2)$ and $c_1 > 0$ such that

$$c_1 \theta^{-1-\nu} \le \beta(\theta), \quad \forall \theta \in [0, \pi/2].$$

The most important quantitative assumption on the kernel is *finite angular momentum transfer* [118]

$$\int_0^{\pi/2} \theta^2 \beta(\theta)\, d\theta = \frac{8}{\pi}. \tag{3.2}$$

Remark 3.3. *The choice of $\frac{8}{\pi}$ in* (3.2) *is a normalisation constant that fixes $C_\beta = 1$ as in Section 1.1.*

Remark 3.4. *Our results also readily generalise to more general interaction kernels B^ϵ which do not decouple or satisfy the specific scaling of item (**A3.2**). As in [2, 4], we could consider kernels that satisfy the following bound on the total cross section*

$$T^\epsilon(|v-v_*|) := \int_0^{\pi/2} \theta^2 \sin\theta B^\epsilon(|v-v_*|,\theta) d\theta \le C(|v-v_*|^{-4}+1)\omega(|v-v_*|^2). \tag{3.3}$$

Here, $C > 0$ is a constant and ω is a bounded positive function such that $\omega(r) \to 0$ as $r \to \infty$ and

$r \to 0$.

As for the grazing collision limit $\epsilon \downarrow 0$, we require that there exists a function T such that

$$|T^\epsilon(r) - T(r)| \le o(1) \left(r^{-4} + 1\right) \omega(r), \quad \text{as } \epsilon \downarrow 0. \tag{3.4}$$

Up to a multiplicative constant, the limiting Landau collision operator reads

$$Q_L(f,f) = \nabla \cdot \left(f \int_{\mathbb{R}^3} f_* T(|v-v_*|)|v-v_*|^2 \Pi[v-v_*](\nabla \log f - \nabla_* \log f_*) dv_* \right).$$

We will discuss this generalisation in more detail in Remark 3.15.

Remark 3.5. *The Coulombic collision kernel which couples $\gamma = -3$ with $\nu = 2$ (see the discussion in Section 3.2.1) fails the finite angular momentum transfer* (3.2)*. In this case, a minor logarithmic cut-off adjustment is needed; for example [118] we consider instead*

$$\beta^\epsilon(\theta) := \frac{1}{\log \epsilon^{-1}} 1_{\theta \ge \epsilon} \beta(\theta).$$

Under this new scaling, we require

$$\int_0^{\pi/2} \theta^2 \beta^\epsilon(\theta) d\theta \overset{\epsilon \downarrow 0}{\to} \frac{8}{\pi}.$$

Our methods can be adapted to cover this case as well. The strong compactness estimate as it is written in Appendix C technically fails, but since the cut-off disappears in the grazing collision limit, the necessary strong compactness is still valid [4, 5].

Theorem 3.2 (Grazing collision limit of H-gradient flow solutions from Boltzmann to Landau)**.** *Suppose $(f^\epsilon)_{\epsilon>0}$ is a family of H-gradient flow solutions to the Boltzmann equation associated to B^ϵ satisfying assumptions (**A3.1**) and (**A3.2**). Then there exists f an H-gradient flow solution to the Landau equation, such that up to a subsequence $f^\epsilon(t) \rightharpoonup f(t)$ for every $t \in [0,T]$.*

Moreover, for $\gamma \in [-2,0]$ and fixed $\phi \in \dot{W}^{1,\infty}(\mathbb{R}^3)$ (respectively, $\gamma \in [-4,-2)$ and fixed $\phi \in \dot{W}^{2,\infty}(\mathbb{R}^3)$), the function $t \mapsto \int f(t)\phi$ is Hölder continuous with exponent $\frac{1}{2}$.

Proof. Our starting point is (EDI_ϵ) from the definition of H-gradient flow solutions to the Boltzmann equation. By definition and the finite initial quantities in (**A3.1**), we have the uniform bounds

$$\sup_{\epsilon>0} \max \left(\sup_{t\in[0,T]} \int_{\mathbb{R}^3} (1+|v|^2) f_t^\epsilon(v) dv, \ \sup_{t\in[0,T]} \mathcal{H}[f^\epsilon(t)], \ \int_0^T D_B^\epsilon(f_t^\epsilon) dt, \ \int_0^T |\dot{f}^\epsilon|_\epsilon^2(t) dt \right) < +\infty. \tag{3.5}$$

These uniform bounds and assumption (**A3.2**) are used in the following steps.

1. Extract a convergent subsequence of f^ϵ and some limit f with the claimed time regularity (Section 3.4).

2. Establish the estimate (Section 3.5)

$$\liminf_{\epsilon\downarrow 0} D_B^\epsilon(f^\epsilon(t)) \geq D_L(f(t)) \qquad \text{a.e. } t \in (0,T).$$

3. Establish the estimate (Section 3.6)

$$\liminf_{\epsilon\downarrow 0} |\dot{f}^\epsilon(t)|_\epsilon^2 \geq |\dot{f}(t)|_L^2 \qquad \text{a.e. } t \in (0,T).$$

Next, we pass to the limit $\epsilon \downarrow 0$ in (EDI_ϵ) using (**A3.1**), Fatou's Lemma, Step 2, Step 3, and the lower semi-continuity of $\mathcal{H}$ to obtain

$$\begin{aligned}\mathcal{H}[f_0] = \liminf_{\epsilon\downarrow 0} \mathcal{H}[f_0^\epsilon] &\geq \liminf_{\epsilon\downarrow 0} \mathcal{H}[f^\epsilon(t)] + \frac{1}{2}\liminf_{\epsilon\downarrow 0}\int_0^t D_B^\epsilon(f^\epsilon(s))ds + \frac{1}{2}\liminf_{\epsilon\downarrow 0}\int_0^t |\dot{f}^\epsilon|_\epsilon^2(s)ds \\ &\geq \mathcal{H}[f(t)] + \frac{1}{2}\int_0^t D_L(f(s))ds + \frac{1}{2}\int_0^t |\dot{f}|_L^2(s)ds,\end{aligned}$$

which implies that f is an H-gradient flow solution to Landau. □

Remark 3.6. *According to the same second moment and entropy bounds from assumptions (**A3.1**) and (**A3.2**), we recover the results in Villani [118], that is, the convergence from f^ϵ to f weakly in $L^p((0,T); L^1(\mathbb{R}^3))$ for every $1 \leq p < +\infty$.*

Remark 3.7 (Affine Representation)**.** *The main idea to showing Step 2 (Section 3.5) and Step 3 (Section 3.6) of the previous proof is to rewrite these expressions in what we will hereafter refer to as **affine representation**. In the context of optimal transport gradient flows, this method was first utilised by Otto [103, Equation (187)] for the Fisher information. More explicitly, we have the characterisation of the Fisher information as*

$$\int_{\mathbb{R}^d} |\nabla\sqrt{f}|^2 = \sup_{\psi\in C_c^\infty(\mathbb{R}^d)} \left\{2\int_{\mathbb{R}^d} \nabla\sqrt{f}\cdot\nabla\psi - \int_{\mathbb{R}^d}|\nabla\psi|^2\right\}.$$

The left-hand side (quadratic in $\nabla\sqrt{f}$) is equal to a supremum over particular affine expressions of $\nabla\sqrt{f}$ on the right-hand side. We establish a similar equality for both the dissipations and the metric derivatives. Taking the Landau dissipation for example and denoting the differential operator $\tilde{\nabla} =$

$|v-v_*|^{1+\frac{\gamma}{2}}\Pi[v-v_*](\nabla-\nabla_*)$, *we show (Section 3.5.1.1)*

$$D_L(f) = 2\iint_{\mathbb{R}^6} |\tilde{\nabla}\sqrt{ff_*}|^2 = \sup_{\psi}\left\{4\iint_{\mathbb{R}^6}(\tilde{\nabla}\sqrt{ff_*})\cdot\tilde{\nabla}\psi - 2\iint_{\mathbb{R}^6}|\tilde{\nabla}\psi|^2\right\}.$$

The specific set of test functions ψ for which the supremum is taken will be specified in later sections.

When the limit f satisfies (**A2.1**) from the previous Chapter 2, we can apply the results there to obtain that the Landau dissipation is a strong upper gradient, see Definition 2.4. In the spirit of Sandier-Serfaty [109, 113], we can readily show that the solution f^ϵ converges strongly to f. This is the content of our next result.

Corollary 3.1. *Suppose $(f^\epsilon)_{\epsilon>0}$ is a family of H-gradient flows for Boltzmann with equality in (EDI_ϵ) satisfying assumptions (**A3.1**) and (**A3.2**). Assume further that D_L is a strong upper gradient for the limit curve f obtained in Theorem 3.2. Then for every $t\in[0,T]$, we have*

$$\mathcal{H}[f^\epsilon(t)] \overset{\epsilon\downarrow 0}{\to} \mathcal{H}[f(t)].$$

Moreover, we also obtain

$$D_B^\epsilon(f^\epsilon)\to D_L(f), \quad |\dot{f}^\epsilon|_\epsilon^2 \to |\dot{f}|_L^2, \quad in\ L^1_{loc.}(0,T).$$

Proof. This proof follows the gradient flow Γ-convergence arguments of Sandier-Serfaty [109, 113]. We fix $t\in[0,T]$. Repeating the passage to the limit $\epsilon\downarrow 0$ from the proof of Theorem 3.2, we have

$$\liminf_{\epsilon\downarrow 0}(-\mathcal{H}[f^\epsilon(t)]) \geq -\mathcal{H}[f_0] + \frac{1}{2}\int_0^t D_L(f(s)) + |\dot{f}|_L^2(s)ds.$$

By Young's inequality and the assumption that D_L is a strong upper gradient for f, we have

$$\frac{1}{2}\int_0^t D_L(f(s)) + |\dot{f}|_L^2(s)ds \geq -\int_0^t \sqrt{D_L(f(s))}|\dot{f}|_L(s)ds \geq \mathcal{H}[f_0] - \mathcal{H}[f(t)].$$

These previous inequalities yield

$$\liminf_{\epsilon\downarrow 0} -\mathcal{H}[f^\epsilon(t)] \geq -\mathcal{H}[f(t)],$$

which, together with the lower semi-continuity of $\mathcal{H}$, gives the strong convergence $\mathcal{H}[f^\epsilon(t)] \overset{\epsilon\downarrow 0}{\to} \mathcal{H}[f(t)]$. This forces all of the inequalities to be equalities and the rest of the proof proceeds exactly the same as in [109, 113]. □

3.2 Notations and formulation of grazing collision limit

We define the Boltzmann collision operator for a fixed collision kernel B acting on test functions ψ by

$$\langle Q_B(f,f),\psi\rangle := -\frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}[f'f'_* - ff_*](\psi' + \psi'_* - \psi - \psi_*)B(|v-v_*|,\theta)\,d\sigma dv_* dv. \tag{3.6}$$

The pre-post collision quantities are defined as follows for v, $v_* \in \mathbb{R}^3$ and $\sigma \in \mathbb{S}^2$.

$$k = \frac{v-v_*}{|v-v_*|}, \quad \cos\theta = k\cdot\sigma, \quad v' = \frac{v+v_*}{2} + \frac{|v-v_*|}{2}\sigma, \quad v'_* = \frac{v+v_*}{2} - \frac{|v-v_*|}{2}\sigma.$$

The typical example of the kernel B satisfying (**A3.2**) is recalled and discussed in Section 3.2.1. In Section 3.2.2, we construct a coordinate system parametrising $\sigma \in \mathbb{S}^2$ by $\theta \in [0,\pi/2]$, the polar angle of $\mathbb{S}^2$, and $\phi \in [0,2\pi]$, the azimuthal angle, which makes the grazing collision computations more explicit in the rest of this paper. We gather these notations and recall the formal grazing collision computations in Section 3.2.3 as an example of many similar computations in this chapter.

3.2.1 Comments on the assumptions of the kernel

For $\gamma \in [-4,0)$, $\nu \in (0,2)$, we recall (**A3.2**) where the form of the kernel is

$$B(r,\theta)\sin\theta = r^\gamma\beta(\theta), \quad \beta(\theta) \gtrsim \theta^{-1-\nu}. \tag{3.7}$$

We keep γ, ν decoupled however the physically relevant case [118, 67] is given by

$$-3 \le \gamma = \frac{s-5}{s-1}, \quad \nu = \frac{2}{s-1}, \quad s \ge 2.$$

The ϵ-collision kernel B^ϵ is defined through the relation in (3.7) where β is extended to $(0,+\infty)$ by zero and we consider

$$\beta^\epsilon(\theta) = \frac{\pi^3}{\epsilon^3}\beta\left(\frac{\pi\theta}{\epsilon}\right), \quad \theta \in (0,\epsilon/2).$$

Remark 3.8. *The finite angular momentum transfer* (3.2) *is minimal for a mathematical theory [118, 4, 2] in the non-cutoff Boltzmann equation.*

- *Clearly, the scaling power of ϵ localizes the singularity in β^ϵ around $\theta = 0$. As well, the choice of ϵ^{-3} preserves* (3.2) *so that*

$$\int_0^{\epsilon/2}\theta^2\beta^\epsilon(\theta)d\theta = \int_0^{\pi/2}\theta^2\beta(\theta)d\theta = \frac{8}{\pi}, \quad \forall\epsilon > 0.$$

- *The strong form of the Landau collision operator $Q_L(f,f)$ is a second-order derivative on f while, at first glance, $Q_B^\epsilon(f,f)$ evaluates no derivatives of f. The finite angular moment transfer allows the interpretation of Q_B^ϵ as a* ***second-order difference quotient "in the angular direction"*** *[2, 4] on f^ϵ.*

We will denote quantities with sub or superscript B meaning that the choice of collision kernel is arbitrary modulo (3.2). Quantities with sub or superscript ϵ will specifically reference the ϵ-collision kernel B^ϵ described above. Based on these notations, we record the most physically relevant parameters and the corresponding behaviour of the angular part of the kernel

	s	γ	ν	$\beta(\theta) \overset{\theta\downarrow 0}{\sim} \theta^{-1-\nu}$
Maxwellian	5	0	1/2	$\theta^{-3/2}$
Coulomb	2	-3	2	θ^{-3}

.

In all these examples mentioned, notice that the following lack of integrability always holds

$$\int_0^{\pi/2} \beta(\theta) d\theta = +\infty.$$

This can be interpreted as an 'affluence of grazing collisions' [64, 66].

3.2.2 Spherical coordinates

According to $\cos\theta = k \cdot \sigma$, we describe the construction of a new coordinate system for which $\sigma \in \mathbb{S}^2$ can be parameterised by $(\theta,\, \phi) \in [0,\, \pi/2] \times [0,\, 2\pi]$ where θ is the polar angle and ϕ is the azimuthal angle. In terms of the integration, this allows us to write the change of variables formula

$$\int_{\mathbb{S}^2} d\sigma = \int_0^{\pi/2} d\theta \sin\theta \int_0^{2\pi} d\phi,$$

so that the concentrated scaling $\theta \sim \epsilon$ is easier to treat. As we shall see, these coordinates allow to identify different mechanisms in the grazing collision limit. The average over the azimuthal angle ϕ induces the second-order differentiation in the orthogonal direction of $v - v_*$ seen in the Landau operator while the integration over the polar angle θ treats the singular kernel $\beta(\theta)$.

Without loss of generality, we can assume $B(|z|,\, \cdot)$ is supported in $[0,\, \pi/2]$ due to standard symmetrising. This is because any configuration of pre-post collision velocities for $\theta \in [\pi/2,\, \pi]$ corresponds to $\theta - \pi/2 \in [0,\, \pi/2]$ when switching $v \leftrightarrow v_*$, see Figure 3.1.

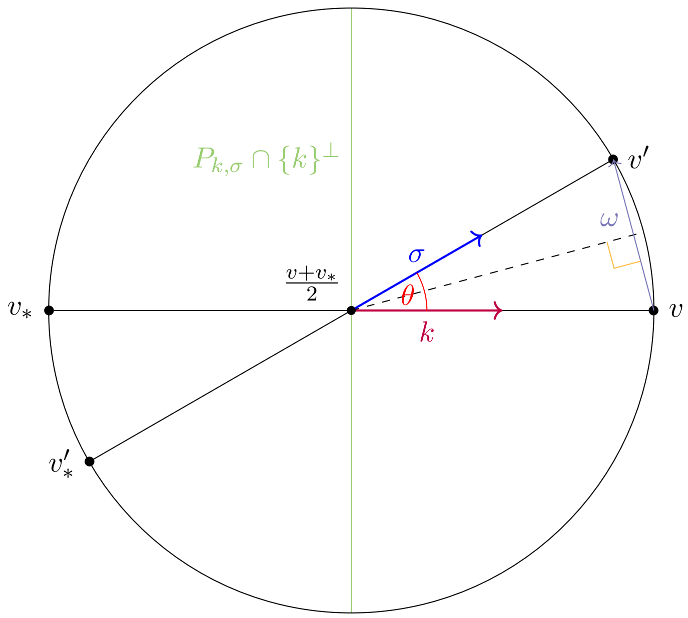


Figure 3.1: Geometry of elastic collision in σ-representation.

That the mid-point/momentum of the velocities is conserved is a consequence of only considering elastic collisions

$$\frac{v+v_*}{2} = \frac{v'+v'_*}{2}.$$

Let us refer to the plane spanned by σ and k by $P_{k,\sigma}$. So Figure 3.1 gives a perspective of $P_{k,\sigma}$ with its normal vector coming directly out of the page. Consider the line obtained by $P_{k,\sigma} \cap \{k\}^{\perp}$. Upon intersection with $\mathbb{S}^2$ (centred at $\frac{v+v_*}{2}$), this line reduces to two vectors which differ by a sign; $\mathbb{S}^2 \cap P_{k,\sigma} \cap \{k\}^{\perp} = \{p_1, p_2\}$ where $p_1 = -p_2$ and we assign $p = p_1$ the '+' choice in the decomposition

$$\sigma = \cos\theta\, k + \sin\theta\, p.$$

Note this is nothing but an orthogonal decomposition of σ with respect to k and $\{k\}^{\perp}$ with a specific sign choice. In general, we will abuse notation for this coordinate transformation by referring to $\mathbb{S}^2 \simeq \partial B_1(\frac{v+v_*}{2})$. We also introduce the following notation which is drawn in Figure 3.3

$$\mathbb{S}^1_{k^{\perp}} := \mathbb{S}^2 \cap \{k\}^{\perp} \simeq \mathbb{S}^1.$$

A three-dimensional perspective of Figure 3.1 is depicted in Figure 3.2.

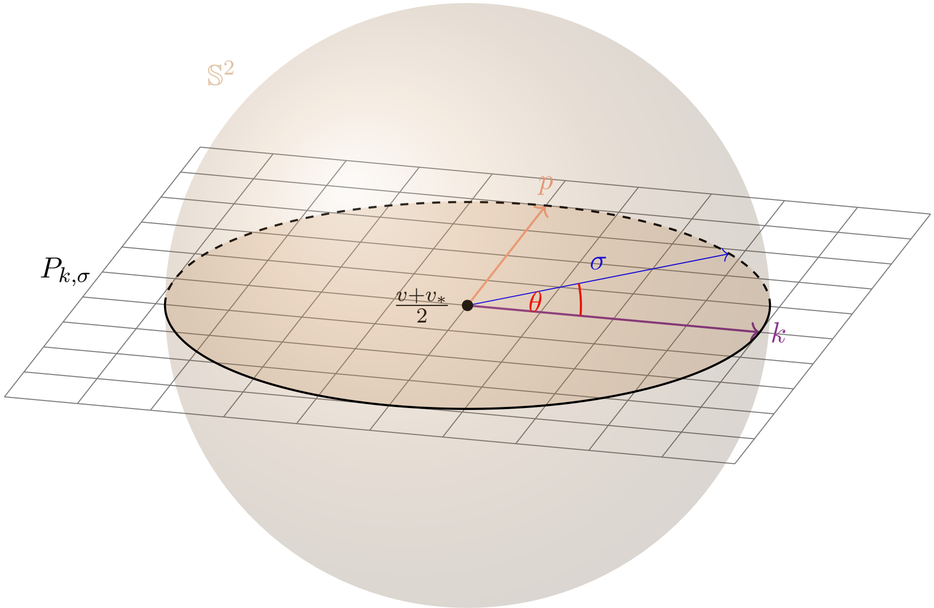


Figure 3.2: Three-dimensional perspective of $\mathbb{S}^2$ with the plane spanned by k, σ given by $P_{k,\sigma}$.

We introduce one final parameterisation for p, see Figure 3.3. Consider fixed k and $\theta \in [0, \pi/2]$, and distinct $\sigma_1,\, \sigma_2 \in \mathbb{S}^2$ such that

$$\cos\theta = k \cdot \sigma_i, \quad i = 1, 2.$$

Following the construction described earlier, this uniquely defines $p_i \in \mathbb{S}^2 \cap P_{k,\sigma_i} \cap \{k\}^\perp$ such that

$$\sigma_i = \cos\theta\, k + \sin\theta\, p_i, \quad i = 1, 2.$$

Notice that both $p_1,\, p_2 \in \mathbb{S}^1_{k^\perp}$. Thus, given orthonormal vectors $\{h,\, i\} \subset \mathbb{S}^1_{k^\perp}$, we can express

$$p = \cos\phi\, h + \sin\phi\, i, \quad \text{for some } \phi \in [0,\, 2\pi].$$

This leads to the following change of variables; given $k \in \mathbb{S}^2$ (determined by $v,\, v_* \in \mathbb{R}^3$), we have

$$\int_{\mathbb{S}^2} d\sigma = \int_0^{\pi/2} d\theta \sin\theta \int_{\mathbb{S}^1_{k^\perp}} dp = \int_0^{\pi/2} d\theta \sin\theta \int_0^{2\pi} d\phi.$$

We shall refer to both changes of variables to $(\theta,\, p)$ and $(\theta,\, \phi)$ as spherical coordinates. With these notations, we will also use the following expressions for the post-collision velocities as perturbations of the pre-collision velocities

$$v' = v + \frac{1}{2}|v - v_*|(\sigma - k)$$
$$v'_* = v_* - \frac{1}{2}|v - v_*|(\sigma - k).$$

The post-collision velocity expressions in this form are especially useful in the grazing collision limit $|\sigma - k| \to 0$ (see (3.11)). Using the spherical coordinate system described, we have

$$\begin{aligned}\sigma &= \cos\theta\, k + \sin\theta\, p, \\ v' &= v - \frac{1}{2}|v - v_*|(1-\cos\theta)k + \frac{1}{2}|v - v_*|\sin\theta\, p, \\ v'_* &= v_* + \frac{1}{2}|v - v_*|(1-\cos\theta)k - \frac{1}{2}|v - v_*|\sin\theta\, p.\end{aligned} \tag{3.8}$$

Thus, in spherical coordinates, an equivalent form of (3.6) is

$$\langle Q(f,f), \psi\rangle = -\frac{1}{4}\iint_{\mathbb{R}^6}\int_{\theta=0}^{\pi/2}\int_{\mathbb{S}^1_{k^\perp}} [f'f'_* - ff_*](\psi' + \psi'_* - \psi - \psi_*)B(|v-v_*|,\theta)\sin\theta\, dpd\theta dv_* dv. \tag{3.9}$$

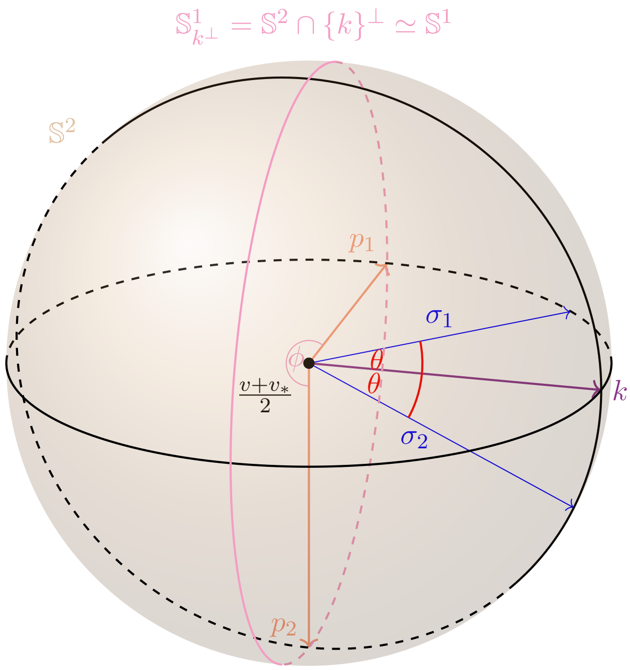


Figure 3.3: Different configurations of σ_i, p_i for $i = 1, 2$ given fixed k, θ such that $\cos\theta = k \cdot \sigma_i$.

The general principle we want to fix is that the grazing collision limit is more easily treated in the spherical coordinate representation of (3.9). We notice the identity

$$|\sigma - k|^2 = 2(1 - k\cdot\sigma) = 2(1-\cos\theta) \tag{3.10}$$

which leads to the following useful estimates

$$\theta^2 \lesssim 1 - \cos\theta \lesssim 1 - k\cdot\sigma \lesssim |\sigma - k|^2 \lesssim \theta^2. \tag{3.11}$$

More precisely, we recall

$$\begin{aligned} \frac{2}{\pi^2}\theta^2 \le 1-\cos\theta \le \frac{1}{2}\theta^2, \quad \theta\in[0,\pi], \\ \frac{2}{\pi}\theta \le \sin\theta \le \theta, \quad \theta\in[0,\pi/2]. \end{aligned} \tag{3.12}$$

Remark 3.9. *Using* (3.11) *and the spherical coordinate transformation we just described, the finite angular momentum transfer* (3.2) *can be equivalently expressed as*

$$\int_{\mathbb{S}^2}(1-k\cdot\sigma)b(cos^{-1}(k\cdot\sigma))d\sigma<+\infty, \quad b(\theta)=\sin\theta\beta(\theta).$$

This notation is sometimes used for example in [2].

3.2.3 The formal grazing collision limit

We dedicate this subsection to formally recalling the grazing collision limit. This sketch is in Lemma 3.3 which relies on some technical results (Lemmas 3.4–3.6) which are postponed until after the formal proof. Following Erbar's convention [61], for a function $\psi:\mathbb{R}^3\to\mathbb{R}$ we define the discrete gradient operator

$$\bar{\nabla}\psi(v,v_*,\sigma):=\psi(v')+\psi(v'_*)-\psi(v)-\psi(v_*).$$

To compress some expressions further, we extend Erbar's discrete gradient to functions $\psi:(v,v_*)\in\mathbb{R}^6\mapsto\psi(v,v_*)\in\mathbb{R}$ by

$$\bar{\nabla}\psi(v,v_*,\sigma):=\psi(v',v'_*)+\psi(v'_*,v')-\psi(v,v_*)-\psi(v_*,v).$$

In this way, the action of $Q_B^\epsilon(f,f)$ on test functions ψ reads (using B^ϵ as the kernel in (3.6)) either

$$\langle Q_B^\epsilon(f,f),\psi\rangle = -\frac{1}{8}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}[\bar{\nabla}(ff_*)][\bar{\nabla}\psi]B^\epsilon(|v-v_*|,\theta)\,d\sigma dv_*dv, \tag{3.13}$$

$$= \frac{1}{2}\iint_{\mathbb{R}^6}ff_*\int_{\mathbb{S}^2}\bar{\nabla}\psi B^\epsilon\,d\sigma dv_*dv, \tag{3.14}$$

where the second expression comes from changing variables in the pre-post collision velocities from (3.6).

We define the Landau collision operator $Q_L(f,f)$ acting on test functions ψ by

$$\langle Q_L(f,f),\psi\rangle := -\frac{1}{2}\iint_{\mathbb{R}^6}|v-v_*|^{2+\gamma}[(\nabla-\nabla_*)(ff_*)]^T\Pi[v-v_*](\nabla\psi-\nabla_*\psi_*)dv_*dv. \tag{3.15}$$

We have already introduced the $\tilde{\nabla}$ operator in Chapter 2. We recall this here for functions of one variable $\psi(v)$ and extend it to functions of two variables $\psi(v, v_*)$ by

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

This leads to the abbreviation of (3.15) for $Q_L(f, f)$ acting on test functions ψ by either

$$\langle Q_L(f, f), \psi\rangle = -\frac{1}{2}\iint_{\mathbb{R}^6}[\tilde{\nabla}(ff_*)]^T\tilde{\nabla}\psi\, dv_* dv, \tag{3.16}$$
$$= \frac{1}{2}\iint_{\mathbb{R}^6} ff_*\tilde{\nabla}\cdot\tilde{\nabla}\psi\, dv_* dv. \tag{3.17}$$

(3.17) is formally obtained by integrating by parts $\tilde{\nabla}$ from (3.16). The operator $\tilde{\nabla}\cdot$ is notation for the adjoint to $\tilde{\nabla}$ (the version acting on functions of two variables) meaning that for a vector field $A(v, v_*) \in \mathbb{R}^3$, it reads

$$[\tilde{\nabla}\cdot A](v, v_*) := (\nabla - \nabla_*)\cdot(\Pi[v - v_*]|v - v_*|^{1+\frac{\gamma}{2}}A(v, v_*)).$$

To be explicit, $\tilde{\nabla}\cdot\tilde{\nabla}\psi$ should just be thought of as notation for

$$\tilde{\nabla}\cdot\tilde{\nabla}\psi = |v - v_*|^{2+\gamma}(\nabla - \nabla_*)(\Pi[v - v_*](\nabla - \nabla_*)\psi),$$

where we have used $(\nabla - \nabla_*)|v - v_*| \in \ker\Pi[v - v_*]$.

Lemma 3.3 (Grazing Collision Limit). *For fixed sufficiently smooth f, ψ we have*

$$\langle Q_B^\epsilon(f, f), \psi\rangle \overset{\epsilon\downarrow 0}{\to} \langle Q_L(f, f), \psi\rangle.$$

Formal proof. We work at the level of (3.14) converging to (3.17). We change variables to spherical coordinates while factoring B^ϵ into its kinetic and angular parts, $B^\epsilon(v - v_*, \theta)\sin\theta = |v - v_*|^\gamma\beta^\epsilon(\theta)$;

$$\langle Q_B^\epsilon(f, f), \psi\rangle = \frac{1}{2}\iint_{\mathbb{R}^6} ff_*|v - v_*|^\gamma\int_{\theta=0}^{\epsilon/2}\beta^\epsilon(\theta)\int_{\mathbb{S}^1_{k^\perp}}\bar{\nabla}\psi\, dpd\theta dv_* dv$$
$$= -\frac{\pi^2}{2}\iint_{\mathbb{R}^6} ff_*|v - v_*|^\gamma\int_{\chi=0}^{\pi/2}\beta(\chi)\frac{1}{\epsilon^2}\int_{\mathbb{S}^1_{k^\perp}}\bar{\nabla}\psi\, dpd\chi dv_* dv.$$

The last line uses the rescaling of $\theta = \epsilon\chi/\pi$. According to Lemma 3.6 (in the particular case of functions

$\psi = \psi(v)$), we have

$$\frac{1}{\epsilon^2\chi^2}\int_{\mathbb{S}^1_{k^\perp}} \bar{\nabla}\psi\, dp \overset{\epsilon\downarrow 0}{\to} \frac{1}{8\pi}(\nabla - \nabla_*)\cdot \left[|v-v_*|^2\Pi[v-v_*](\nabla\psi - \nabla_*\psi_*)\right].$$

Returning to the computations, the formal passage of the limit $\epsilon \downarrow 0$ inside the integral gives

$$\begin{aligned}\langle Q^\epsilon_B(f,f_*),\psi\rangle &\overset{\epsilon\downarrow 0}{\to} \frac{\pi}{16}\iint_{\mathbb{R}^6} ff_*|v-v_*|^{2+\gamma}\left(\int_0^{\pi/2}\chi^2\beta(\chi)d\chi\right)(\nabla-\nabla_*)\cdot[\Pi[v-v_*](\nabla-\nabla_*)\psi]dv_*dv\\ &= \frac{1}{2}\iint_{\mathbb{R}^6} ff_*|v-v_*|^{2+\gamma}(\nabla-\nabla_*)\cdot[\Pi[v-v_*](\nabla-\nabla_*)\psi]dv_*dv,\end{aligned}$$

where the last line follows from Lemma A.4 and we have used the finite angular momentum transfer (3.2) with the normalization $C_\beta = \frac{\pi}{8}\int_0^{\frac{\pi}{2}}\chi^2\beta(\chi)d\chi = 1$. Notationally, we recognise $\tilde{\nabla}\cdot\tilde{\nabla}\psi$ in the integrand.

□

I invite the reader to (formally) verify the same passage of the grazing collision limit starting from the 'first order' formulations of the collision operators; (3.13) converging to (3.16). Let us stress that the averaged quantity $\int_{\mathbb{S}^1_{k^\perp}}\bar{\nabla}\psi\, dp$ **behaves like a second order derivative of** ψ which we shall make precise in Lemma 3.6. Consequently, the term to control is the angular momentum transfer (3.2). Furthermore, this approach involves **no derivatives of** f and is even amenable to weak-strong convergence pairs $\left(f^\epsilon, \frac{1}{\epsilon^2}\int_{\mathbb{S}^1_{k^\perp}}\bar{\nabla}\psi\, dp\right) \to \left(f, \tilde{\nabla}\cdot\tilde{\nabla}\psi\right)$. By weak-strong convergence, we mean that $f^\epsilon \rightharpoonup f$ weakly whereas we shall consider sufficiently smooth fixed ψ so that, up to multiplicative constants, $\frac{1}{\epsilon^2}\int_{\mathbb{S}^1_{k^\perp}}\bar{\nabla}\psi\, dp$ converges strongly to $\tilde{\nabla}\cdot\tilde{\nabla}\psi$.

We now provide the supplementary estimates that prove, up to multiplicative constants, $\bar{\nabla} \overset{\epsilon\downarrow 0}{\to} \tilde{\nabla}$ which we already used in the proof of Lemma 3.3. Given the notation with $\bar{\nabla}$, $\tilde{\nabla}$ we consider the more general case for these operators acting on functions $\psi(v, v_*)$ which will be used later; the specific case of functions of a single variable $\psi = \psi(v)$ follows in the same way. To better facilitate this, consider the following classical volume-preserving change of variables to momentum and relative velocity coordinates also considered by Bobylev [21] and Villani [118], for instance. Define

$$x = \frac{v-v_*}{2}, \quad y = \frac{v+v_*}{2},$$

and we also note in particular that $k = x/|x|$ which gives $\sigma - k = \sigma - \frac{x}{|x|}$. Recalling the definitions of the post-collision velocities, for fixed $\sigma \in \mathbb{S}^2$, this leads to

$$\begin{cases} v' &= y + |x|\sigma \\ v'_* &= y - |x|\sigma \end{cases} \implies x' = |x|\sigma, \quad y' = y.$$

Moreover, the differentiation in $\tilde{\nabla}$ really only sees the x direction in the sense that $\nabla_v - \nabla_{v_*} = \nabla_x$. Thus, abusing notation, we can write

$$\bar{\nabla}\psi(x,y) = \psi(|x|\sigma, y) + \psi(-|x|\sigma, y) - \psi(x,y) - \psi(-x,y), \quad \tilde{\nabla}\psi = |2x|^{1+\frac{\gamma}{2}}\Pi[x]\nabla_x\psi(x,y), \tag{3.18}$$

where we have identified $\psi = \psi(v,v_*) = \psi(x,y)$. Using this change of variables, we have

Lemma 3.4 (Estimates for $\bar{\nabla}\psi$ adapted from [118])**.** *Fix $\psi \in C_c^\infty(\mathbb{R}^6)$, we have the first order estimate in x and $\sigma - k$*

$$|\bar{\nabla}\psi| \leq Lip_x(\psi)|2x||\sigma - k|,$$

as well as the second order estimate in x but still first order in $\sigma - k$

$$|\bar{\nabla}\psi| \leq \|D_x^2\psi\|_{L^\infty}|2x|^2|\sigma - k|.$$

By averaging over the circle $p \in \mathbb{S}^1_{k^\perp}$, we can obtain a second order estimate in $\sigma - k$

$$\left|\frac{1}{2\pi}\int_{\mathbb{S}^1_{k^\perp}} \bar{\nabla}\psi\, dp\right| \leq \|D_x^2\psi\|_{L^\infty}|2x|^2|\sigma - k|^2,$$

where the geometric meaning of $\mathbb{S}^1_{k^\perp}$ and p given k and σ are from Figures 3.1–3.3.

Proof. We use (3.18) and the Fundamental Theorem of Calculus to write

$$\begin{aligned}\bar{\nabla}\psi &= \psi(|x|\sigma) + \psi(-|x|\sigma) - \psi(x) - \psi(-x)\\ &= \int_0^1 \frac{d}{dt}\left\{\psi(t|x|\sigma + (1-t)x) + \psi(-[t|x|\sigma + (1-t)x])\right\}dt\\ &= (|x|\sigma - x)\cdot\int_0^1 \nabla_x\psi(t|x|\sigma + (1-t)x) - \nabla_x\psi(-[t|x|\sigma + (1-t)x])dt.\end{aligned}$$

The first estimate is directly obtained from here. For the second estimate, we continue by using Taylor's formula to replace the integrand

$$\begin{aligned}\nabla_x\psi(t|x|\sigma + (1-t)x) =& \nabla_x\psi(x) + t(|x|\sigma - x)\cdot\int_0^1 D_x^2\psi(s[t|x|\sigma + (1-t)x] + (1-s)x)ds,\\ \nabla_x\psi(-[t|x|\sigma + (1-t)x]) =& \nabla_x\psi(-x) - t(|x|\sigma - x)\cdot\int_0^1 D_x^2\psi(-\{s[t|x|\sigma + (1-t)x] + (1-s)x\})ds.\end{aligned}$$

Thus, we have

$$\bar{\nabla}\psi = |x|(\sigma - k)\cdot(\nabla_x\psi(x) - \nabla_x\psi(-x))$$
$$+ |x|^2(\sigma-k)\otimes(\sigma-k) :$$
$$: \underbrace{\int_0^1 t\int_0^1 D_x^2\psi(s[t|x|\sigma + (1-t)x] + (1-s)x) + D_x^2\psi(-\{s[t|x|\sigma + (1-t)x] + (1-s)x\})dsdt}_{\leq \|D_x^2\psi\|_{L^\infty}}\,. \tag{3.19}$$

Here, the colon ':' refers to the trace inner product between square matrices. The second term is bounded by $\|D_x^2\psi\|_{L^\infty}|x|^2|\sigma-k|^2 \leq 2\|D_x^2\psi\|_{L^\infty}|x|^2|\sigma-k|$ (recalling $|\sigma-k|^2 = 2(1-\cos\theta)$ and we have restricted $\theta\in[0,\pi/2]$) which gives the right contribution for the second estimate. Thus, it suffices to look at the first term which yields the right estimate after an application of the Mean Value Theorem.

Finally, for the third estimate which improves the order with respect to $\sigma - k$, we take the average of (3.19) over $\mathbb{S}^1_{k^\perp}$. Again, the second term is bounded in the right way ($\sigma-k$ depends on p, but $|\sigma-k|$ does not), so we only focus on the first term

$$\frac{1}{2\pi}\int_{\mathbb{S}^1_{k^\perp}} |x|(\sigma-k)\cdot(\nabla_x\psi(x)-\nabla_x\psi(-x))dp = \frac{|x|}{2\pi}\int_{\mathbb{S}^1_{k^\perp}} ((\cos\theta-1)k + \sin\theta p)\cdot(N_k\,k + N_p\,p)\,dp$$
$$= \frac{|x|}{2\pi}\int_{\mathbb{S}^1_{k^\perp}} (\cos\theta-1)N_k + \sin\theta N_p\,dp.$$

where we have decomposed the vectors in the inner product under this geometry, see (3.8), as

$$N_k = (\nabla_x\psi(x)-\nabla_x\psi(-x))\cdot k,\quad N_p = (\nabla_x\psi(x)-\nabla_x\psi(-x))\cdot p.$$

Continuing, notice that the N_p term is a linear combination of $\cos\phi$ and $\sin\phi$ (ϕ being the azimuthal angle in Figure 3.3) and that the integral can be written as

$$\int_{\mathbb{S}^1_{k^\perp}} dp = \int_0^{2\pi} d\phi.$$

Hence, the second term integrates to zero and we focus on the first term. The Mean Value Theorem gives $|N_k| = |k\cdot(\nabla_x\psi(x)-\nabla_x\psi(-x))| \leq |2x|\|D_x^2\psi\|_{L^\infty}$ which yields

$$\left|\frac{|x|}{2\pi}\int_{\mathbb{S}^1_{k^\perp}} (\sigma-k)\cdot(\nabla_x\psi(x)-\nabla_x\psi(-x))dp\right| \leq 2|x|^2\|D_x^2\psi\|_{L^\infty}|1-\cos\theta| = \|D_x^2\psi\|_{L^\infty}|x|^2|\sigma-k|^2.$$

□

Lemma 3.5 (Behaviour of $\bar{\nabla}\psi/\epsilon$)**.** *Under the previous notations, for $\psi \in C_c^\infty(\mathbb{R}^6)$, $\epsilon > 0$, $\chi = \pi\theta/\epsilon$, where $\cos\theta = k \cdot \sigma$ and $\theta \in [0, \epsilon/2]$, we have*

$$\frac{1}{\epsilon}\bar{\nabla}\psi = \frac{|x|}{\epsilon}\left(\left(\cos\frac{\epsilon\chi}{\pi} - 1\right)k + \sin\frac{\epsilon\chi}{\pi}\,p\right) \cdot \nabla_x[\psi(x) + \psi(-x)] + \mathcal{O}(\|D_x^2\psi\|_{L^\infty}\epsilon|x|^2).$$

In particular, we have the convergence

$$\frac{1}{\epsilon}\bar{\nabla}\psi \overset{\epsilon\downarrow 0}{\to} \frac{\chi}{\pi}|x|\,p \cdot \nabla_x[\psi(x) + \psi(-x)], \quad \textit{pointwise } v,\, v_* \in \mathbb{R}^3.$$

For functions $\psi = \psi(v)$, this is equivalent to

$$\frac{1}{\epsilon}\bar{\nabla}\psi \overset{\epsilon\downarrow 0}{\to} \frac{\chi}{2\pi}|v - v_*|\,p \cdot (\nabla - \nabla_*)\psi.$$

Proof. Firstly, we recall the size estimates of $\sigma - k$ from (3.10) and (3.11) giving

$$|\sigma - k|^2 = 2(1 - \cos\theta) \leq \theta^2 = \epsilon^2\frac{\chi^2}{\pi^2}.$$

Using again $\sigma - k = (\cos\theta - 1)k + \sin\theta\, p$ and the substitution $\theta = \epsilon\chi/\pi$, we obtain the first identity starting from (3.19).

Passing to the limit $\epsilon \downarrow 0$ is a matter of recalling Taylor expansions. □

Lemma 3.6 (Behaviour of $\frac{1}{\epsilon^2}\int_{\mathbb{S}^1_{k^\perp}} \bar{\nabla}\psi\, dp$)**.** *Under the previous notations, for $\psi \in C_c^\infty(\mathbb{R}^6)$, $\epsilon > 0$, $\cos\theta = k \cdot \sigma$, $\theta \in [0, \epsilon/2]$, $\chi = \pi\theta/\epsilon$, we have*

$$\frac{1}{\epsilon^2\chi^2}\int_{\mathbb{S}^1_{k^\perp}} \bar{\nabla}\psi dp \overset{\epsilon\downarrow 0}{\to} \frac{1}{2\pi}\nabla_x \cdot (|x|^2\Pi[x]\nabla_x[\psi(x, y) + \psi(-x, y)]).$$

For functions $\psi = \psi(v)$, this is equivalent to

$$\frac{1}{\epsilon^2\chi^2}\int_{\mathbb{S}^1_{k^\perp}} \bar{\nabla}\psi dp \overset{\epsilon\downarrow 0}{\to} \frac{1}{8\pi}(\nabla - \nabla_*) \cdot (|v - v_*|^2\Pi[v - v_*](\nabla - \nabla_*)\psi).$$

Proof. We start from (3.19), the terms there are recalled

$$
\begin{aligned}
\bar{\nabla}\psi = &\overbrace{|x|(\sigma-k)\cdot(\nabla_x\psi(x)-\nabla_x\psi(-x))}^{=:T_1} \\
&+|x|^2(\sigma-k)\otimes(\sigma-k): \\
&:\int_0^1 t\int_0^1 D_x^2\psi(s[t|x|\sigma+(1-t)x]+(1-s)x)+D_x^2\psi(-\{s[t|x|\sigma+(1-t)x]+(1-s)x\})dsdt \\
&=: T_1+T_2.
\end{aligned}
$$

The idea is to individually take the limits

$$
\lim_{\epsilon\downarrow 0}\frac{1}{\epsilon^2}\int_{\mathbb{S}^1_{k^\perp}} T_i dp, \quad i=1,\,2.
$$

Let us start with T_2 which will involve a Dominated Convergence argument. By the triangle inequality and (3.10)

$$
|T_2|\le \|D_x^2\psi\|_{L^\infty}|x|^2|\sigma-k|^2=\epsilon^2\|D_x^2\psi\|_{L^\infty}|x|^2\frac{\chi^2}{\pi^2}.
$$

This is an integrable majorant when multiplied against ϵ^{-2} so we obtain

$$
\lim_{\epsilon\downarrow 0}\int_{\mathbb{S}^1_{k^\perp}}\frac{T_2}{\epsilon^2}dp=\frac{\chi^2}{2\pi^2}|x|^2\int_{\mathbb{S}^1_{k^\perp}} p\otimes p dp:(D_x^2\psi(x)+D_x^2\psi(-x)).
$$

This is because, recalling the expression of σ with respect to k, p in (3.8) and the scaling $\theta\sim\epsilon$, we see

$$
\sigma-k=\underbrace{(\cos\theta-1)}_{\sim\epsilon^2}\,k+\underbrace{\sin\theta}_{\sim\epsilon}\,p.
$$

Therefore, the remaining contribution is $\lim_{\epsilon\downarrow 0}\epsilon^{-1}(\sigma-k)=p$. Using Lemma A.4, we can compute $\int_{\mathbb{S}^1_{k^\perp}} p\otimes p\,dp=\pi\Pi[k]$ so that we obtain

$$
\lim_{\epsilon\downarrow 0}\int_{\mathbb{S}^1_{k^\perp}}\frac{T_2}{\epsilon^2\chi^2}dp=\frac{|x|^2}{2\pi}\Pi[x]:(D_x^2\psi(x)+D_x^2\psi(-x)).
$$

Turning to the T_1 term, we directly integrate similar to the proof of Lemma 3.4. Copying the notation there, we have

$$
\begin{aligned}
\int_{\mathbb{S}^1_{k^\perp}}\frac{T_1}{\epsilon^2}dp&=\frac{1}{\epsilon^2}\int_{\mathbb{S}^1_{k^\perp}}|x|(\sigma-k)\cdot(\nabla_x\psi(x)-\nabla_x\psi(-x))dp=\frac{|x|N_k}{\epsilon^2}\int_{\mathbb{S}^1_{k^\perp}}(\cos\theta-1)dp \\
&=2\pi(x\cdot(\nabla_x\psi(x)-\nabla_x\psi(-x)))\frac{\cos\theta-1}{\epsilon^2}
\end{aligned}
$$

recalling $N_k = (\nabla_x\psi(x) - \nabla_x\psi(-x)) \cdot k$ (independent of p) and the other contribution vanishes. Changing variables $\theta = \epsilon\chi/\pi$, we use the fact that

$$\frac{\cos(\epsilon\chi/\pi)-1}{\epsilon^2} \overset{\epsilon\downarrow 0}{\to} -\frac{\chi^2}{2\pi^2} \qquad \text{to obtain} \qquad \lim_{\epsilon\downarrow 0}\int_{\mathbb{S}^1_{k^\perp}} \frac{T_1}{\epsilon^2\chi^2}dp = -\frac{1}{\pi}x\cdot(\nabla_x\psi(x) - \nabla_x\psi(-x)).$$

Putting both terms together, we have

$$\frac{1}{\epsilon^2\chi^2}\int_{\mathbb{S}^1_{k^\perp}} \bar{\nabla}\psi dp \overset{\epsilon\downarrow 0}{\to} \frac{|x|^2}{2\pi}\Pi[x] : (D_x^2\psi(x) + D_x^2\psi(-x)) - \frac{1}{\pi}x\cdot(\nabla_x\psi(x) - \nabla_x\psi(-x)).$$

In d dimensions, a direct computation shows $\nabla_x \cdot |x|^2\Pi[x] = -(d-1)x$, which in this case for $d=3$ allows one to recognise the product rule and conclude. $\square$

3.3 Gradient flow structure of the Boltzmann equation [61]

This section recalls Erbar's gradient flow perspective of the Boltzmann equation [61]. We refer to Section 2.1.2 for the necessary gradient flow concepts needed in this chapter. Of course, the most comprehensive reference for gradient flows in metric spaces remains [6]. The content of this section should be compared with the analogous exposition for Landau in Section 2.2. We denote $\mathcal{M}_B = \mathcal{M}(\mathbb{R}^6\times\mathbb{S}^2)$ the space of scalar signed Radon measures on $\mathbb{R}^6\times\mathbb{S}^2$ endowed with the weak-* topology as members of the dual of $C_c(\mathbb{R}^6\times\mathbb{S}^2)$. $\mathcal{M}_L = \mathcal{M}^3(\mathbb{R}^6)$ will denote the space of $\mathbb{R}^3$-valued signed Radon measures on $\mathbb{R}^6$ with the weak-* topology as members of the dual of $C_c(\mathbb{R}^6;\mathbb{R}^3)$.

Definition 3.7 (Collision Rate Equation)**.** For Borel curves in time $t\in[0,T]$, $\mu : t\mapsto \mu_t \in \mathscr{P}$ and $M : t \mapsto M_t \in \mathcal{M}_B$ we say $(\mu,\, M) \in CRE_T$ (or CRE if $T=1$), if they satisfy Erbar's *collision rate equation* [61]

$$\partial_t\mu_t + \frac{1}{4}\bar{\nabla}\cdot M_t = 0,$$

in the distributional sense. By this, we mean that for any $\phi\in C_c^\infty((0,T)\times\mathbb{R}^3)$,

$$\int_0^T\int_{\mathbb{R}^3}\partial_t\phi(t,v)d\mu_t(v)dt + \frac{1}{4}\int_0^T\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\bar{\nabla}\phi(t,v,v_*,\sigma)dM_t(v,v_*,\sigma)dt = 0.$$

Furthermore, the second moment is finite and non-increasing

$$\int_{\mathbb{R}^3}|v|^2d\mu_t(v) \le \int_{\mathbb{R}^3}|v|^2d\mu_s(v) \qquad \text{for any } 0\le s\le t\le T.$$

We restrict our attention to probability measures with densities $\mu = f\mathcal{L}$. As well, if $M\in\mathcal{M}_B$ has a density against Lebesgue measure on $\mathbb{R}^6\times\mathbb{S}^2$, we will identify $dM(v,v_*,\sigma) = M(v,v_*,\sigma)dvdv_*d\sigma$. For

$(\mu, M) \in CRE_T$, consider $\tau = \mu(dv)\mu(dv_*) + |M| \in \mathcal{M}(\mathbb{R}^6 \times \mathbb{S}^2)$ that dominates both $\mu \otimes \mu$ and M. Taking $N\tau = M$ and $g\tau = \mu(dv)\mu(dv_*)$, the densities of M and $\mu \otimes \mu$ with respect to τ, we can define the *Boltzmann action* of the curve μ by

$$\mathcal{A}_B(\mu, M) := \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{\alpha(N, g, g')}{B(|v - v_*|, \theta)} d\sigma dv_* dv, \quad \alpha(u, s, t) := \begin{cases} \frac{|u|^2}{4L(s,t)}, & L(s,t) \neq 0, \\ 0, & L(s,t) = 0,\, u = 0 \\ +\infty, & L(s,t) = 0,\, u \neq 0 \end{cases}.$$

Here, $L(s, t)$ is the logarithmic mean between $s, t > 0$ (c.f. Lemma A.3) and B is a fixed collision kernel. The function α is jointly convex, lower semi-continuous, and 1-homogeneous (see [61] and compare with Lemma 2.14). The appearance of the collision kernel B is free and our choice of its location is consistent with Erbar so that the collision rate equation (as well as $\bar{\nabla}$) may be written without the collision kernel. In the case where both M and μ have densities with respect to Lebesgue (denoting again $\mu = f\mathcal{L}$), we write

$$\Lambda = \Lambda(f) = \frac{f' f'_* - f f_*}{\log f' f'_* - \log f f_*}.$$

For $\mu^\epsilon = f^\epsilon \mathcal{L}$ corresponding to curves with respect to the collision kernel B^ϵ, we will write $\Lambda^\epsilon = \Lambda(f^\epsilon)$. In exactly the same way as Definition 2.19 for the Landau metric, d_L, the Boltzmann (pseudo)-metric d_B is defined by

$$d_B^2(\lambda_1, \lambda_2) := \inf \left\{ T \int_0^T \mathcal{A}_B(\mu_t, M_t) dt \,\middle|\, \begin{array}{c} (\mu, M) \in CRE_T, \\ \mu(0) = \lambda_1, \quad \mu(T) = \lambda_2 \end{array} \right\}, \quad \lambda_1, \lambda_2 \in \mathscr{P}_2.$$

Analogously, for $(\mu, M) \in GCE_T$, we already defined the *Landau action of the curve* μ in Section 2.2.2 by (2.5). We will write d_{B^ϵ} or d_ϵ for the Boltzmann metric corresponding to the collision kernel B^ϵ. We have already used the notation in Definition 3.1, but we make precise here that $|\dot{f}|_\epsilon, |\dot{f}|_L$ refer to the d_ϵ, d_L-metric derivatives for a curve $\mu_t = f_t \mathcal{L}$, respectively. We now quote the Boltzmann analogue of Proposition 2.22.

Lemma 3.8 (Propositions A.9, A.11, and Corollary A.10 of [61]). *For a fixed collision kernel B, a curve $t \in [0, T] \mapsto f_t \mathcal{L} \in \mathscr{P}$ is absolutely continuous with respect to the Boltzmann distance d_B if and only if there is a family of mobilities $M_t^B \in \mathcal{M}_B$ such that $(f,\, M^B) \in CRE_T$ with finite total action*

$$\int_0^T \mathcal{A}_B(f, M_B) dt = \frac{1}{4} \int_0^T \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{|M^B|^2}{\Lambda(f) B} d\sigma dv_* dv dt < +\infty.$$

Moreover, there is a unique $\tilde{M}^B \in \mathcal{M}_B$ *such that*

$$|\dot{f}|_B^2(t) = \mathcal{A}_B(f, \tilde{M}_B) = \frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} \frac{|\tilde{M}^B|^2}{\Lambda(f)B} d\sigma dv_* dv.$$

Furthermore, under the class of admissible $M^B \in \mathcal{M}_B$ *(i.e.* $(f, M^B) \in CRE_T$*),* $\tilde{M}^B$ *is characterised by the minimisation property*

$$\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} \frac{|\tilde{M}^B|^2}{\Lambda(f)B} d\sigma dv_* dv \leq \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} \frac{|\tilde{M}^B + \eta|^2}{\Lambda(f)B} d\sigma dv_* dv.$$

for any $\eta \in \mathcal{M}_B$ *which is* $\bar{\nabla}\cdot$*-free, that is to say*

$$\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} \bar{\nabla}\xi \, d\eta = 0, \quad \forall \xi \in C_c^\infty(\mathbb{R}^3).$$

If a measure $M \in \mathcal{M}_B$ *satisfies the minimisation property, then* $M = U\Lambda(f)B$ *where*

$$U \in \overline{\{\bar{\nabla}\phi \,|\, \phi \in C_c^\infty(\mathbb{R}^3)\}}^{L^2(\Lambda(f)B d\sigma dv_* dv)}.$$

We end this section with some remarks concerning the gradient flow concepts in our proof of Theorem 3.2 without going into the details.

Remark 3.10. *We recall the uniform metric derivative integrability* (3.5) *in the proof of Theorem 3.2*

$$\sup_{\epsilon > 0} \int_0^T |\dot{f}^\epsilon|_\epsilon^2(t) dt < +\infty.$$

This estimate yields the following which we shall revisit in Sections 3.4 and 3.6.

- *(Regularity) - Each curve we consider* $t \mapsto f^\epsilon(t)$ *is absolutely continuous with respect to* d_ϵ*. Moreover, this property is uniform in* $\epsilon > 0$*.*
- *(Compactness) - Furthermore, from Lemma 3.8, we can evaluate the metric derivative as the action of a unique collision rate* M^ϵ*:* $|\dot{f}^\epsilon|_\epsilon^2(t) = \mathcal{A}_B(f^\epsilon(t), M^\epsilon(t))$*. This representation has two consequences - firstly, our assumptions allow us to prove compactness of* $(f^\epsilon, M^\epsilon)_{\epsilon > 0}$ *in Section 3.4 to some limit* (f^0, M^0)*. Secondly, it is easier to work with the jointly convex integrands of* $\mathcal{A}_B$ *and* $\mathcal{A}_L$ *to show Step 3 from the proof of Theorem 3.2;*

$$\liminf_{\epsilon \downarrow 0} |\dot{f}^\epsilon|_\epsilon^2(t) = \liminf_{\epsilon \downarrow 0} \mathcal{A}_B(f^\epsilon, M^\epsilon) \geq \mathcal{A}_L(f^0, M^0) \geq |\dot{f}|_L^2(t).$$

Remark 3.11. *Given a weak solution f of the Landau equation (respectively, f^ϵ of the Boltzmann equation) that dissipates entropy as in Remark 3.2, one immediately obtains an estimate for the metric derivative;*

$$|\dot{f}|_L^2(t) \leq D_L(f(t)), \quad (\textit{respectively, } |\dot{f^\epsilon}|_\epsilon^2(t) \leq D_B^\epsilon(f^\epsilon(t))).$$

This is owed to the fact that one can take the admissible grazing (respectively, collision) rate

$$M = -ff_*\tilde{\nabla}\log f, \quad (\textit{respectively, } M_B^\epsilon = -\Lambda(f^\epsilon)B^\epsilon\bar{\nabla}\log f^\epsilon),$$

giving an upper bound for the square of the Landau metric

$$d_L^2(f_0, f_1) \leq \int_0^1 \mathcal{A}_L(f_t, M)dt = \frac{1}{2}\int_0^1 \iint_{\mathbb{R}^6} ff_*|\tilde{\nabla}\log f|^2 dv_* dv dt < +\infty.$$

An analogous estimate holds for the Boltzmann metric. In this way, the dissipation of entropy implies the Energy Dissipation Inequality, hence renormalized solutions are H-gradient flow solutions.

3.4 Compactness of curves

The aim of this section is to deduce general compactness results of curves $(f^\epsilon)_{\epsilon>0}$ (not necessarily solutions in any sense) subject to the uniform moments and metric derivative bounds in (3.5) from the proof of Theorem 3.2. The first half of this section establishes Theorem 3.9 and Proposition 3.10 which are the main compactness results. The second half of this section, Section 3.4.1, builds on these compactness results by confirming the passage of the grazing collision limit at the level of the generalized continuity equations; CRE to GCE. Here, we have taken $T = 1$ for simplicity but the results work for CRE_T and GCE_T with the appropriate time scaling.

Theorem 3.9 (Compactness of f^ϵ)**.** *Let $f^\epsilon : [0,1] \to \mathscr{P}_2$ be curves satisfying the uniform moments and metric derivative bounds* (3.5)*. Then there exists $f : [0,1] \to \mathscr{P}_2$ obtained by a convergent subsequence such that*

$$f^\epsilon(t) \xrightarrow{\sigma} f(t), \quad \forall t \in [0,1].$$

Moreover, in the case of $\gamma \in [-2, 0]$ (respectively, $\gamma \in [-4,-2)$), f is continuous in duality against 1-Lipschitz functions (respectively, continuous in duality against bounded functions with second derivatives bounded by 1).

Remark 3.12. *By the well-known Kantorovich-Rubinstein duality, the continuity in duality against 1-Lipschitz functions in the case of $\gamma \in [-2, 0]$ is equivalent to continuity in the 1-Wasserstein met-*

ric [121]. This was noticed by Erbar [61] in the case $\gamma = 0$, whose proof we have generalised using the finite angular momentum transfer (3.2).

Proof of Theorem 3.9. This is an application of a general Ascoli-Arzelà compactness result [6, Proposition 3.3.1] together with Lemma 3.12, which says that there exists $C > 0$ an explicit constant depending on the finite momentum transfer (3.2), and the uniform moments from (3.5) such that

$$\left| \int_{\mathbb{R}^3} \psi(v)(f_t^\epsilon(v) - f_s^\epsilon(v))dv \right| \le C d_\epsilon(f^\epsilon(t), f^\epsilon(s)), \quad \forall s,\, t \in [0,1],$$

for any function ψ with Lipschitz semi-norm bounded by 1 in the case $\gamma \in [-2, 0]$ (respectively, second derivative bounded by 1 in the case $\gamma \in [-4, -2)$).

By the absolute continuity of f^ϵ and the uniform L^2 integrability of the metric derivative, we obtain

$$\sup_{\epsilon>0} \sup_{\psi} \left| \int_{\mathbb{R}^3} \psi(v)(f_t^\epsilon(v) - f_s^\epsilon(v))dv \right| \le C|t-s|^{\frac{1}{2}}.$$

This estimate with the basic weak compactness of $(f^\epsilon)_{\epsilon>0} \subset \mathscr{P}$ by the moment bounds in (3.5) satisfies the conditions to apply a version of Ascoli-Arzelà in this setting [6, Proposition 3.3.1]. □

Proposition 3.10. *Let $(f^\epsilon)_{\epsilon>0}$ be curves satisfying the uniform moment and metric derivative bounds in* (3.5)*. We consider the subsequence of f^ϵ that converges to f given by Theorem 3.9. Assume further that there exists $t_0 \in [0,1]$ such that*

$$\sup_{\epsilon>0} D_B^\epsilon(f^\epsilon(t_0)) < \infty,$$

then

$$\sqrt{f^\epsilon(t_0)} \to \sqrt{f(t_0)}, \text{ in } L^2_{loc.}.$$

Remark 3.13. *This result is reminiscent, but weaker, than those in [96, 4]. There, the stronger convergence $\sqrt{f^\epsilon} \to \sqrt{f}$ in $L^2_{t,v}$ is achieved by exploiting the extra time (and space, in the case of inhomogeneity) regularity from velocity averaging methods [71] on renormalized solutions to the Boltzmann equation [58].*

Proof of Proposition 3.10. The argument is standard after recalling the main estimate of Appendix C so we shall quickly sketch the main ideas. For brevity, we fix and suppress t_0. Using that f^ϵ are probability densities, we immediately obtain (up to a subsequence) the weak convergence $\sqrt{f^\epsilon} \rightharpoonup g$ in L^2 for some $g \in L^2$. According to Appendix C, we obtain the estimate

$$\sup_{\epsilon>0} \left\| \sqrt{f_R^\epsilon} \right\|_{\dot{H}^{\frac{\nu}{2}}} \le C_R, \quad \forall R > 1,$$

where $f^\epsilon_R = f^\epsilon \chi_R$ is a smooth cut-off approximation of f^ϵ vanishing outside B_{R+1} and $C_R > 0$ is a constant depending on R and the value of $\sup_{\epsilon>0} D^\epsilon_B(f^\epsilon)$. This upgrades the convergence so that $\sqrt{f^\epsilon} \to g$ strongly in $L^2_{loc.}$. In particular, along a further subsequence, we have $f^\epsilon \to g^2$ pointwise almost every $v \in \mathbb{R}^3$. By Theorem 3.9, this identifies $g^2 = f$ and we are done. $\square$

Curves in CRE and GCE are pairs of measures (f, M). Assuming the bounds (3.5), we have established compactness for the first component of these curves, f^ϵ. We now state and prove the compactness result for the second component, M^ϵ.

Proposition 3.11 ((Scaled) compactness of M^ϵ)**.** *Suppose $(f^\epsilon, M^\epsilon) \in CRE$ is a pair of curves where $(f^\epsilon)_{\epsilon>0}$ satisfies the uniform moment and metric derivative bounds (3.5). Assume M^ϵ is the optimal collision rate given by Lemma 3.8. Then, for any $\tilde{q} \in \left[-\frac{\gamma}{2}, 1-\frac{\gamma}{2}\right]$ the family $\{|v-v_*|^{\tilde{q}}\theta M^\epsilon\}_{\epsilon>0}$ is a bounded set of Radon measures in the weak-* topology against C_c functions. In particular, choosing*

$$q_\gamma := \begin{cases} 1, & \gamma \in [-2,0) \\ 2, & \gamma \in [-4,-2) \end{cases}, \quad 0 < \delta < \delta_\gamma := \begin{cases} -\frac{\gamma}{2}, & \gamma \in [-2,0) \\ -\frac{\gamma}{2} - 1, & \gamma \in [-4,-2) \end{cases},$$

we have that the family

$$\left\{|v-v_*|^{q_\gamma}\left(1+\left[|v|^2+|v_*|^2\right]\right)^{\frac{\delta}{2}}\theta M^\epsilon\right\}_{\epsilon>0}$$

is compact in the set of Radon measures.

Proof. We will only show the uniform bound. The compactness statement uses the same argument because the choices of $q = q_\gamma$ and δ depending on γ ensure $q + \delta \in \left[-\frac{\gamma}{2}, 1-\frac{\gamma}{2}\right)$. Fix $\Psi \in C_c([0,1] \times \mathbb{R}^6 \times \mathbb{S}^2)$ non-negative and $\tilde{q} \in \left[-\frac{\gamma}{2}, 1-\frac{\gamma}{2}\right]$; we use Corollary A.2 to estimate

$$\begin{aligned}\int_0^1 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \Psi|v-v_*|^{\tilde{q}}|\theta||M^\epsilon| &= \int_0^1 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{|M^\epsilon|}{\sqrt{\Lambda(f^\epsilon)B^\epsilon}}|v-v_*|^{\tilde{q}}|\theta|\Psi\sqrt{\Lambda(f^\epsilon)B^\epsilon} \\ &\leq \left(\int_0^1 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{|M^\epsilon|^2}{\Lambda(f^\epsilon)B^\epsilon}\right)^{\frac{1}{2}} \left(\int_0^1 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |v-v_*|^{2\tilde{q}}|\theta|^2\Psi^2\Lambda(f^\epsilon)B^\epsilon\right)^{\frac{1}{2}}.\end{aligned}$$

Here, we have multiplied and divided by $\sqrt{\Lambda^\epsilon B^\epsilon}$ and then applied Cauchy-Schwarz. This reveals precisely the ϵ-action in the first term, which by the metric derivative bound in (3.5) and Lemma 3.8, is bounded. Focusing on the second term, we use Corollary A.2 to estimate

$$\int_0^1 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |v-v_*|^{2\tilde{q}}|\theta|^2\Psi^2\left[\frac{f^{\epsilon\prime}f^{\epsilon\prime}_* + f^\epsilon f^\epsilon_*}{2}\right]B^\epsilon d\sigma dv_* dv dt.$$

By symmetry, we can pass the post-collision velocity evaluations of $f^{\epsilon\prime}f^{\epsilon\prime}_*$ onto Ψ. We develop $B^\epsilon \sin\theta =$

$|v-v_*|^\gamma\beta^\epsilon$ and continue the estimate

$$\begin{aligned}
&\int_0^1 \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} \Psi|v-v_*|^{\tilde{q}}|\theta||M^\epsilon| d\sigma dv_* dv dt \\
&\quad \leq C\left(\int_0^1 \iint_{\mathbb{R}^6}\int_{\theta=0}^{\epsilon/2}\int_{\mathbb{S}^1_{k\perp}} |v-v_*|^{2\tilde{q}+\gamma}|\theta|^2(\Psi^2+\Psi'^2) f^\epsilon f^\epsilon_* \frac{\pi^3}{\epsilon^3}\beta\left(\frac{\pi\theta}{\epsilon}\right) dp d\theta dv_* dv dt\right)^{\frac{1}{2}} \\
&\quad = \sqrt{2\pi}C\left(\int_0^1 \iint_{\mathbb{R}^6}\int_0^{\pi/2} \pi^2|v-v_*|^{2\tilde{q}+\gamma}\chi^2(\Psi^2+\Psi'^2) f^\epsilon f^\epsilon_*\beta(\chi) d\chi dv_* dv dt\right)^{\frac{1}{2}}.
\end{aligned}$$

The final expression is uniformly bounded in ϵ by the assumptions (notice $2\tilde{q}+\gamma\in[0,2]$) and finite angular momentum transfer (3.2). □

Corollary 3.2. *Consider the setting of Proposition 3.11 and denote $\mathscr{M}\in\mathcal{M}((0,1)\times\mathbb{R}^6\times\mathbb{S}^2)$ a limit of the family $\{|v-v_*|^{\tilde{q}}\theta M^\epsilon\}_\epsilon$. Then, $\mathscr{M}$ can be disintegrated with respect to Lebesgue measure on t.*

Remark 3.14. *This disintegration in time is the Boltzmann analogue of the same statement for Proposition 2.18.*

Proof. We repeat the proof of Proposition 3.11 but fix a test function $\Psi(v,\, v_*,\, \sigma,\, t)\in C_c((0,1)\times\mathbb{R}^6\times\mathbb{S}^2)$ now so that its time dependence is an indicator function, i.e.

$$\Psi(v,\, v_*,\, \sigma,\, t)=\psi(v,\, v_*,\, \sigma)\chi_{[a,b]}(t), \quad \psi\in C_c(\mathbb{R}^6\times\mathbb{S}^2).$$

We continue from the last line of the previous proof to obtain

$$\begin{aligned}
&\int_a^b \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} |\psi||v-v_*|^{\tilde{q}}|\theta||M^\epsilon_t| d\sigma dv_* dv dt \\
&\quad \leq \sqrt{2\pi}C\left(\int_a^b \iint_{\mathbb{R}^6}\int_0^{\pi/2} \pi^2|v-v_*|^{2\tilde{q}+\gamma}\chi^2(\psi^2+\psi'^2) f^\epsilon f^\epsilon_*\beta(\chi) d\chi dv_* dv dt\right)^{\frac{1}{2}}.
\end{aligned}$$

The finite angular momentum transfer (3.2) is independent of time, moreover the zeroeth to second moments of f^ϵ are bounded uniformly in ϵ and t from (3.5) so absorbing these terms into a constant leaves

$$\int_a^b \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} |\psi||v-v_*|^{\tilde{q}}|\theta||M^\epsilon_t| d\sigma dv_* dv dt \leq C|b-a|^{\frac{1}{2}}.$$

This estimate holds in the limit $\epsilon\downarrow 0$ as well, so the measure $\mathscr{M}$ can also be disintegrated with respect to Lebesgue measure on $t\in[0,1]$. □

From now on, we take for granted that the limits in Proposition 3.11 are also families in time of signed measures on $\mathbb{R}^6\times\mathbb{S}^2$.

Lemma 3.12 (Comparison of certain topologies against the Boltzmann metric)**.** *Let $\mu_0, \mu_1 \in \mathscr{P}_2(\mathbb{R}^3)$ be probability measures that are absolutely continuous with respect to Lebesgue. There exists a constant $C > 0$ depending only on the finite angular momentum transfer* (3.2) *and the second moment of μ_0 such that the following holds:*

1. *In the case $\gamma \in [-2, 0]$, we have*

$$\left| \int_{\mathbb{R}^3} \psi(v) d(\mu_0 - \mu_1)(v) \right| \leq C d_B^\epsilon(\mu_0, \mu_1)$$

   for any function ψ with Lipschitz semi-norm bounded by 1.

2. *In the case $\gamma \in [-4, -2)$, we have*

$$\left| \int_{\mathbb{R}^3} \psi(v) d(\mu_0 - \mu_1)(v) \right| \leq C d_B^\epsilon(\mu_0, \mu_1)$$

   for any function ψ with second derivative bounded by 1.

Proof. We will only show the proof of the first estimate in the case $\gamma \in [-2, 0]$. The proof of the second estimate differs only by estimating $\bar{\nabla}\psi$ using the second estimate of Lemma 3.4 instead of the first. Without loss of generality we can assume that $d_B^\epsilon(\mu_0, \mu_1) < \infty$. We take M the optimal collision rate in the sense of Lemma 3.8. Fix Lipschitz $\psi : \mathbb{R}^3 \to \mathbb{R}$ with $\mathrm{Lip}(\psi) \leq 1$ and we recall the first estimate of Lemma 3.4

$$\begin{aligned} |\bar{\nabla}\psi| &\leq \left| \psi\left(v + \frac{1}{2}|v - v_*|(\sigma - k)\right) - \psi(v) \right| + \left| \psi\left(v_* - \frac{1}{2}|v - v_*|(\sigma - k)\right) - \psi(v_*) \right| \\ &\leq |v - v_*||\sigma - k|. \end{aligned}$$

Using ψ as a test function in the collision rate equation (which we justify at the end) connecting μ_0 to μ_1 we have

$$\begin{aligned} &\left| \int_{\mathbb{R}^d} \psi d\mu_0 - \int_{\mathbb{R}^d} \psi d\mu_1 \right| = \frac{1}{4} \left| \int_0^1 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \bar{\nabla}\psi M d\sigma dv_* dv dt \right| \\ &\leq \frac{1}{4} \int_0^1 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |v - v_*||\sigma - k||M| d\sigma dv_* dv dt \\ &\leq \frac{1}{4} \left(\int_0^1 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{|M|^2}{\Lambda(f) B^\epsilon} d\sigma dv_* dv dt \right)^{\frac{1}{2}} \left(\int_0^1 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |v - v_*|^2 |\sigma - k|^2 \Lambda(f) B^\epsilon d\sigma dv_* dv dt \right)^{\frac{1}{2}}. \end{aligned}$$

At this point, we recognise the first term as the time integrated ϵ-Boltzmann action. In the second term, we can apply Corollary A.2. Since M is optimal we can estimate the previous expression by

$$\frac{1}{4}d_B^\epsilon(\mu_0,\mu_1)\left(\int_0^1\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|v-v_*|^2|\sigma-k|^2\left(\frac{f'f'_*+ff_*}{2}\right)B^\epsilon d\sigma dv_* dvdt\right)^{\frac{1}{2}}.$$

By Figure 3.1 or directly from the definitions, we have $|\sigma'-k'|=|k-\sigma|$. Hence, the arithmetic mean is just ff_* upon symmetrisation (B^ϵ is invariant when $(v,v_*)\leftrightarrow(v',v'_*)$). This leads to

$$\left|\int_{\mathbb{R}^d}\psi d\mu_0-\int_{\mathbb{R}^d}\psi d\mu_1\right|\leq\frac{1}{4}d_B^\epsilon(\mu_0,\mu_1)\left(\int_0^1\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|v-v_*|^2|\sigma-k|^2B^\epsilon ff_* d\sigma dv_* dvdt\right)^{\frac{1}{2}}.$$

Now, we change representation from σ-representation to (θ,ϕ)-representation, see Section 3.2. We recall from (3.10) that $|\sigma-k|^2=2(1-\cos\theta)$. Substituting this leads to the further estimate

$$\begin{aligned}&\left|\int_{\mathbb{R}^d}\psi d\mu_0-\int_{\mathbb{R}^d}\psi d\mu_1\right|\\&\leq\frac{\sqrt{2}}{4}d_B^\epsilon(\mu_0,\mu_1)\left(\int_0^1\iint_{\mathbb{R}^6}\int_{\theta=0}^{\epsilon/2}\int_{\phi=0}^{2\pi}(1-\cos\theta)|v-v_*|^{2+\gamma}\beta^\epsilon(\theta)ff_* d\theta d\phi dv_* dvdt\right)^{\frac{1}{2}}\\&=\frac{\sqrt{2}}{4}d_B^\epsilon(\mu_0,\mu_1)\left(\int_0^1\iint_{\mathbb{R}^6}\int_{\theta=0}^{\epsilon/2}\int_{\phi=0}^{2\pi}(1-\cos\theta)|v-v_*|^{2+\gamma}\frac{\pi^3}{\epsilon^3}\beta\left(\frac{\pi\theta}{\epsilon}\right)ff_* d\theta d\phi dv_* dvdt\right)^{\frac{1}{2}}.\end{aligned}$$

We perform the change of variables $\chi=\pi\theta/\epsilon$ and directly compute the ϕ integral to get

$$\begin{aligned}&\left|\int_{\mathbb{R}^d}\psi d\mu_0-\int_{\mathbb{R}^d}\psi d\mu_1\right|\\&\leq\frac{\sqrt{\pi}}{2}d_B^\epsilon(\mu_0,\mu_1)\left(\int_0^1\iint_{\mathbb{R}^6}\int_{\chi=0}^{\pi/2}(1-\cos\frac{\epsilon\chi}{\pi})|v-v_*|^{2+\gamma}\frac{\pi^2}{\epsilon^2}\beta(\chi)ff_* d\chi dv_* dvdt\right)^{\frac{1}{2}}.\end{aligned}$$

We eliminate the factor of $1/\epsilon^2$ by the inequality $1-\cos x\leq\frac{1}{2}x^2$ in (3.12) when $x\in[0,\pi]$ to give

$$\left|\int_{\mathbb{R}^d}\psi d\mu_0-\int_{\mathbb{R}^d}\psi d\mu_1\right|\leq\frac{\sqrt{2\pi}}{4}d_B^\epsilon(\mu_0,\mu_1)\left(\int_0^1\iint_{\mathbb{R}^6}\int_{\chi=0}^{\pi/2}\chi^2|v-v_*|^{2+\gamma}\beta(\chi)ff_* d\chi dv_* dvdt\right)^{\frac{1}{2}}.$$

The integral decouples and the proof is complete recalling the finite angular momentum transfer (3.2).

We now address the use of time-independent Lipschitz-bounded functions as test functions in the collision rate equation for $\gamma\in[-2,0]$. Fix $\phi\in C_c^\infty((0,1)\times\mathbb{R}^3)$ and repeat the previous estimates. In particular by the first estimate of Lemma 3.4, notice that the drift term can be estimated as follows

$$\begin{aligned}\left|\int_0^1\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\bar{\nabla}\phi M\right|&\leq\int_0^1\mathrm{Lip}_v(\phi)\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|v-v_*||\sigma-k||M|\\&\leq\frac{1}{2}\int_0^1\mathrm{Lip}_v(\phi)\left\{\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\frac{|M|^2}{\Lambda(f)B^\epsilon}+\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|v-v_*|^2|\sigma-k|^2\Lambda(f)B^\epsilon\right\}.\end{aligned}$$

The finite metric derivative and second moment bounds from (3.5) and the finite angular momentum (3.2) give estimates on both the integrals in the curly brackets. This leaves only the dependence on the Lipschitz semi-norm of the test function so by density, one can take test functions with bounded Lipschitz dependence in $v \in \mathbb{R}^3$. The argument is similar for $\gamma \in [-4,-2)$ by using instead the second estimate of Lemma 3.4 to restrict the class of test functions to those with bounded second derivative. The time independence can be treated by considering test functions $\phi_k(t,v) = \eta_k(t)\zeta(v)$ for $\eta_k \in C_c^\infty((0,1))$ and $\zeta \in C_c^\infty(\mathbb{R}^3)$ where η_k is a smooth approximation of the indicator function on (0,1) for $k \in \mathbb{N}$. $\square$

3.4.1 Grazing collision limit of the continuity equations

Now that we understand compactness for (f^ϵ, M^ϵ), we need to verify that the limit points actually satisfy the *GCE*. Here, we only consider the distributional formulation of these equations without regard to, for example, the moment condition (2.3). Given $\psi \in C_c^\infty((0,T)\times\mathbb{R}^3)$ a test function, we recall the generalised continuity equations for the ϵ-Boltzmann ($(f^\epsilon, M^\epsilon) \in CRE$) and Landau ($(f,M) \in GCE$) equations

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \psi f^\epsilon dv dt + \frac{1}{4}\int_0^T \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} \bar{\nabla}\psi M^\epsilon d\sigma dv_* dv dt = 0, \tag{3.20}$$

and

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \psi f dv dt + \frac{1}{2}\int_0^T \iint_{\mathbb{R}^6} \tilde{\nabla}\psi \cdot M dv_* dv dt = 0. \tag{3.21}$$

We can directly compare the first terms from the weak convergence $f^\epsilon \rightharpoonup f$ from Theorem 3.9. It remains to understand precisely the convergence in the transport term. Recalling Proposition 3.11, we know that $\left(|v-v_*|^q \left(1 + \left[|v|^2 + |v_*|^2\right]^{\frac{\delta}{2}}\right)\theta M^\epsilon\right)_{\epsilon>0}$ is compact for q, δ satisfying

$$q = \begin{cases} 1, & \gamma \in [-2,0) \\ 2, & \gamma \in [-4,-2) \end{cases}, \quad 0 < \delta < \begin{cases} -\frac{\gamma}{2} & \gamma \in [-2,0) \\ -\frac{\gamma}{2} - 1, & \gamma \in [-4,-2) \end{cases}.$$

We define a 'lift' mapping whose use will soon be clear

$$L_{q,\delta} : M \in \mathcal{M}_B \mapsto \frac{|v-v_*|^{-\frac{\gamma}{2}-q}}{4\left(1 + \left[|v|^2 + |v_*|^2\right]^{\frac{\delta}{2}}\right)} \int_{\mathbb{S}^2} M\, p\, d\sigma \in \mathcal{M}_L.$$

Recall $p \in \mathbb{S}^1_{k^\perp}$ as defined in Section 3.2.2 in the integral over $\mathbb{S}^2$. The motivation for this comes from looking at the formal grazing collision limit of (3.20). Along the subsequence of convergence in Proposition 3.11, let us write

$$|v-v_*|^q\left(1+\left[|v|^2+|v_*|^2\right]^{\frac{\delta}{2}}\right)\theta M^\epsilon \stackrel{*}{\rightharpoonup} M^0_{q,\delta},\quad \epsilon\downarrow 0.$$

This suggests to multiply and divide by θ within the integral in (3.20). By Lemma 3.4, we obtain

$$\frac{1}{\theta}\bar{\nabla}\psi \stackrel{\epsilon\downarrow 0}{\to} \frac{1}{2}|v-v_*|(\nabla\psi-\nabla_*\psi_*)\cdot p=\frac{1}{2}|v-v_*|^{-\frac{\gamma}{2}}p\cdot\tilde{\nabla}\psi.$$

Multiplying and dividing by $|v-v_*|^q\left(1+\left[|v|^2+|v_*|^2\right]^{\frac{\delta}{2}}\right)\theta$ in the transport term of (3.20) and omitting the time integral, we have

$$\frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\bar{\nabla}\psi M^\epsilon=\frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\frac{|v-v_*|^q\left(1+\left[|v|^2+|v_*|^2\right]^{\frac{\delta}{2}}\right)\theta M^\epsilon}{\theta|v-v_*|^q\left(1+[|v|^2+|v_*|^2]^{\frac{\delta}{2}}\right)}\bar{\nabla}\psi d\sigma dv_* dv.$$

In order to apply weak-strong convergence, we need to ensure that $|v-v_*|^{-q}\left(1+\left[|v|^2+|v_*|^2\right]^{\frac{\delta}{2}}\right)^{-1}\frac{\bar{\nabla}\psi}{\theta}$ decays when v, $v_*\to\infty$ uniformly in $\epsilon>0$. Recall the meaning of weak-strong convergence; if ψ_n converges strongly to ψ in the space of continuous and decaying functions on $\mathbb{R}^6\times\mathbb{S}^2$ and the sequence of signed Radon measures M_n converges weakly-* to M in $\mathcal{M}(\mathbb{R}^6\times\mathbb{S}^2)$, then $\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\psi_n dM_n\to\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\psi dM$. By our choice of q depending on $\gamma\in[-4,0)$ and Lemma 3.4, we can estimate

$$|v-v_*|^{-q}\frac{\bar{\nabla}\psi}{\theta}\lesssim\begin{cases}\mathrm{Lip}(\psi) & q=1,\ \gamma\in[-2,0)\\ \|D^2\psi\|_{L^\infty} & q=2,\ \gamma\in[-4,-2)\end{cases}.$$

By the convergence result in Lemma 3.5, we obtain

$$|v-v_*|^{-q}\frac{\bar{\nabla}\psi}{\theta}\to\frac{1}{2}|v-v_*|^{1-q}p\cdot(\nabla-\nabla_*)\psi=\frac{1}{2}|v-v_*|^{-q-\frac{\gamma}{2}}p\cdot\tilde{\nabla}\psi.$$

Therefore, we can pass to the limit $\epsilon\downarrow 0$,

$$\begin{aligned}\frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\bar{\nabla}\psi M^\epsilon &\stackrel{\epsilon\downarrow 0}{\to}\frac{1}{8}\iint_{\mathbb{R}^6}\frac{|v-v_*|^{-q-\frac{\gamma}{2}}}{\left(1+[|v|^2+|v_*|^2]^{\frac{\delta}{2}}\right)}\tilde{\nabla}\psi\cdot\int_{\mathbb{S}^2}M^0_{q,\delta}p\,d\sigma dv_* dv\\ &=\frac{1}{2}\iint_{\mathbb{R}^6}\tilde{\nabla}\psi\cdot L_{q,\delta}(M^0_{q,\delta})dv_* dv.\end{aligned}$$

This shows that $(f,L_{q,\delta}(M^0_{q,\delta}))$ is an admissible pair for the grazing continuity equation coming from the Boltzmann sequence (f^ϵ,M^ϵ).

Remark 3.15. *We describe more precisely the various changes necessary to generalise our assumptions* *on the collision kernel from the discussion in Remark 3.4. Repeating the proof of Proposition 3.11, we can show that the family*

$$\left\{\frac{\theta M^\epsilon}{\sqrt{T^\epsilon(v,v_*)}}\right\}_{\epsilon>0}$$

has uniformly bounded moments up to first order. To prove a similar result as Lemma 3.12, one has to utilise both the Lipschitz (for large $|v-v_*|\gg 1$) and Hessian (for local $|v-v_*|\leq 1$) estimates in Lemma 3.4 to obtain

$$\left|\int_{\mathbb{R}^3}\psi(v)d(\mu_0-\mu_1)(v)\right|\lesssim d_B^\epsilon(\mu_0,\mu_1),$$

for test functions ψ with Lipschitz semi-norm and second derivative bounded by 1.

Finally, we discuss the grazing collision limit at the level of the generalised continuity equations. We need good estimates for the pairing of $\bar{\nabla}\psi$ against M^ϵ to show the grazing collision limit:

$$\int_0^1\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\bar{\nabla}\psi M^\epsilon=\int_0^1\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\frac{\bar{\nabla}\psi}{\theta}\sqrt{T^\epsilon}\left[\frac{\theta M^\epsilon}{\sqrt{T^\epsilon}}\right]\to\int_0^1\iint_{\mathbb{R}^6}\tilde{\nabla}\psi(v,v_*)\cdot\int_{\mathbb{S}^2}M^0p\,d\sigma dv_*dv,$$

where

$$\tilde{\nabla}\psi(v,v_*)=|v-v_*|\sqrt{T(|v-v_*|)}\Pi[v-v_*](\nabla\psi-\nabla_*\psi_*).$$

The convergence above uses again the Lipschitz and Hessian estimates in Lemma 3.4, the uniform bound for T^ϵ (3.3)*, and the convergence of T^ϵ to T as in* (3.4)*.*

3.5 Lower semicontinuous convergence of the dissipations in the grazing collision limit

Throughout this section we consider $(f^\epsilon)_{\epsilon>0}$ a sequence of probability densities with uniformly bounded second moment and entropy

$$\sup_{\epsilon>0}\int_{\mathbb{R}^3}|v|^2f^\epsilon(v)dv<+\infty,\quad \sup_{\epsilon>0}\int_{\mathbb{R}^3}f^\epsilon\log f^\epsilon dv<+\infty. \tag{3.22}$$

such that $f^\epsilon\overset{\sigma}{\rightharpoonup}f$ for some probability density f. We wish to show the lower semicontinuous convergence of the dissipation.

Proposition 3.13. *Assume $f^\epsilon\overset{\sigma}{\rightharpoonup}f$ with uniform second moment and entropy bounds* (3.22)*. Then we have*

$$\liminf_{\epsilon\downarrow 0}D_B^\epsilon(f^\epsilon)\geq D_L(f).$$

Proof. We first reduce to the case

$$\sup_{\epsilon>0} D_B^\epsilon(f^\epsilon) < +\infty.$$

Indeed, without loss of generality, we may assume that $\liminf_{\epsilon\downarrow 0} D_B^\epsilon(f^\epsilon) < +\infty$ otherwise there is nothing to show. There is a subsequence $(\epsilon_n)_{n\in\mathbb{N}}$ such that $\epsilon_n \downarrow 0$ as $n \to \infty$ for which

$$\sup_{n\in\mathbb{N}} D_B^{\epsilon_n}(f^{\epsilon_n}) \leq \liminf_{\epsilon\downarrow 0} D_B^\epsilon(f^\epsilon) + 1 < +\infty.$$

In this uniformly bounded dissipation setting, we can further say $\sqrt{f^\epsilon} \to \sqrt{f}$ in L^2_{loc} by Proposition 3.10.

Collecting the results of Section 3.5.1, we obtain the following estimates for the dissipation

$$\begin{aligned} D_B^\epsilon(f^\epsilon) &\geq \sup_{\psi\in DS_c^\infty} \left\{-2\iint_{\mathbb{R}^6} \sqrt{f^\epsilon f^\epsilon_*}\left(\int_{\mathbb{S}^2} \bar{\nabla}\psi B^\epsilon\right) - \frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} |\bar{\nabla}\psi|^2 B^\epsilon\right\}, \\ D_L(f) &\leq \sup_{\psi\in DS_c^\infty} \left\{-4\iint_{\mathbb{R}^6} \sqrt{ff_*}\tilde{\nabla}\cdot\tilde{\nabla}\psi - 2\iint_{\mathbb{R}^6} |\tilde{\nabla}\psi|^2\right\}, \end{aligned}$$

where the test functions ψ belong to the following class of smooth functions

$$DS_c^\infty := \left\{ \psi = \psi(v,v_*) \in C_c^\infty(\mathbb{R}^6;\mathbb{R}) \ : \ \begin{array}{l} \psi(v,v_*) = \psi(v_*,v)\, \forall v, v_* \in \mathbb{R}^3, \\ \exists \delta_\psi > 0 \text{ s.t. } \psi(v,v_*) = 0 \ \forall |v-v_*| \leq \delta_\psi \end{array} \right\}.$$

Up to some constants (consistent in both expressions below), we apply the results of Section 3.5.2 and Lemmas 3.5 and 3.6 which state

$$\iint_{\mathbb{R}^6} \sqrt{f^\epsilon f^\epsilon_*}\left(\int_{\mathbb{S}^2} \bar{\nabla}\psi B^\epsilon\right) \overset{\epsilon\downarrow 0}{\to} 2\iint_{\mathbb{R}^6} \sqrt{ff_*}\tilde{\nabla}\cdot\tilde{\nabla}\psi \qquad \text{and} \qquad \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} |\bar{\nabla}\psi|^2 B^\epsilon \overset{\epsilon\downarrow 0}{\to} 8\iint_{\mathbb{R}^6} |\tilde{\nabla}\psi|^2,$$

for any fixed $\psi \in DS_c^\infty$. □

3.5.1 Affine representation of dissipations

In this subsection, we seek to show

$$\begin{aligned} D_B^\epsilon(f^\epsilon) &\geq \sup_{\psi\in DS_c^\infty} \left\{-2\iint_{\mathbb{R}^6} \sqrt{f^\epsilon f^\epsilon_*}\left(\int_{\mathbb{S}^2} \bar{\nabla}\psi B^\epsilon\right) - \frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} |\bar{\nabla}\psi|^2 B^\epsilon\right\}, \\ D_L(f) &\leq \sup_{\psi\in DS_c^\infty} \left\{-4\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} \sqrt{ff_*}\tilde{\nabla}\cdot\tilde{\nabla}\psi - 2\iint_{\mathbb{R}^6} |\tilde{\nabla}\psi|^2\right\}, \end{aligned} \tag{3.23}$$

where DS_c^∞ was introduced in the proof of Proposition 3.13. For ease of notation, we will drop the ϵ superscripts for the Boltzmann dissipation. Recall, the Boltzmann and Landau dissipations can be written as

$$D_B(f) = \frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}[f'f'_* - ff_*]\log\frac{f'f'_*}{ff_*}B,$$
$$D_L(f) = 2\iint_{\mathbb{R}^6}|v-v_*|^{2+\gamma}|\Pi[v-v_*](\nabla-\nabla_*)\sqrt{ff_*}|^2.$$

The proof of the upper bound for the Landau dissipation in (3.23) is the content of the subsequent Sections 3.5.1.1 and 3.5.1.2 while the lower bound for the Boltzmann dissipation in (3.23) is the content of Section 3.5.1.3.

3.5.1.1 Landau dissipation

We begin with a preliminary expression of the Landau dissipation.

Proposition 3.14. *We can express*

$$D_L(f) = \sup_{\xi\in C_c^\infty(\mathbb{R}^6;\mathbb{R}^3)}\left\{-4\iint_{\mathbb{R}^6}\sqrt{ff_*}|v-v_*|^{1+\frac{\gamma}{2}}(\nabla-\nabla_*)\cdot(\Pi[v-v_*]\xi) - 2\iint_{\mathbb{R}^6}|\xi|^2\right\}. \tag{3.24}$$

Proof. Let us denote the right-hand side of (3.24) by $I_L(f)$; we want to show equality $D_L(f) = I_L(f)$.

Showing $I_L(f) \le D_L(f)$ is straight-forward because if $D_L(f) < +\infty$, we can integrate by parts the differential operator $|v-v_*|^{1+\frac{\gamma}{2}}(\nabla-\nabla_*)\cdot\Pi[v-v_*]$ onto $\sqrt{ff_*}$ (since finite dissipation implies $\tilde{\nabla}\sqrt{ff_*}\in L^2$) and then apply Cauchy-Schwarz and Young's inequality in the following way for fixed $\xi\in C_c^\infty(\mathbb{R}^6;\mathbb{R}^3)$

$$\begin{aligned}
&-4\iint_{\mathbb{R}^6}\sqrt{ff_*}|v-v_*|^{1+\frac{\gamma}{2}}(\nabla-\nabla_*)\cdot(\Pi[v-v_*]\xi) - 2\iint_{\mathbb{R}^6}|\xi|^2\\
&\qquad= 4\iint_{\mathbb{R}^6}|v-v_*|^{1+\frac{\gamma}{2}}\Pi[v-v_*]\left[(\nabla-\nabla_*)\sqrt{ff_*}\right]\cdot\xi - 2\iint_{\mathbb{R}^6}|\xi|^2\\
&\qquad\le 4\left(\iint_{\mathbb{R}^6}|v-v_*|^{2+\gamma}|\Pi[v-v_*](\nabla-\nabla_*)\sqrt{ff_*}|^2\right)^{\frac{1}{2}}\left(\iint_{\mathbb{R}^6}|\xi|^2\right)^{\frac{1}{2}} - 2\iint_{\mathbb{R}^6}|\xi|^2\\
&\qquad\le D_L(f).
\end{aligned}$$

Turning to the other direction, we wish to show that $D_L(f)\le I_L(f)$. We assume here that $I_L(f) < +\infty$ or else there is nothing to show. Define the linear operator acting on $\xi\in C_c^\infty(\mathbb{R}^6;\mathbb{R}^3)$ given by

$$F:\xi\mapsto -4\iint_{\mathbb{R}^6}\sqrt{ff_*}|v-v_*|^{1+\frac{\gamma}{2}}(\nabla-\nabla_*)\cdot(\Pi[v-v_*]\xi).$$

Since we assume $I_L(f) < +\infty$, we have

$$\sup_{\xi \in C_c^\infty, \|\xi\|_{L^2}=1} \{\langle F, \xi\rangle - 2\} \le C < +\infty,$$

So by density, F extends uniquely to a bounded linear operator on $L^2(\mathbb{R}^6;\mathbb{R}^3)$. We consider now the continuous, coercive, symmetric, and bilinear form

$$a(\xi,\eta) = 2\iint_{\mathbb{R}^6} \xi\cdot\eta, \quad \forall \xi,\eta \in L^2(\mathbb{R}^6;\mathbb{R}^3).$$

By Lax-Milgram/Riesz representation, there is a unique Riesz representative $u \in L^2(\mathbb{R}^6;\mathbb{R}^3)$ such that

$$2\iint_{\mathbb{R}^6} u\cdot\xi = \langle F,\xi\rangle, \quad \forall \xi \in L^2(\mathbb{R}^6;\mathbb{R}^3).$$

Moreover, $u = \tilde\nabla\sqrt{ff_*}$ in L^2 and this characterises the minimisation problem

$$\frac{1}{2}a(u,u) - \langle F,u\rangle = \min_{\xi\in L^2}\left\{\frac{1}{2}a(\xi,\xi) - \langle F,\xi\rangle\right\}.$$

Using the definitions of a, ϕ, and $u = \tilde\nabla\sqrt{ff_*}$ as well as density, the above gives

$$-2\iint_{\mathbb{R}^6} |\tilde\nabla\sqrt{ff_*}|^2 = \inf_{\xi\in C_c^\infty}\left\{2\iint_{\mathbb{R}^6}|\xi|^2 + 8\iint_{\mathbb{R}^6}\sqrt{ff_*}|v-v_*|^{1+\frac{\gamma}{2}}(\nabla-\nabla_*)\cdot(\Pi[v-v_*]\xi)\right\}.$$

Applying one sign change gives precisely $D_L(f) = I_L(f)$. □

Intuitively, a near optimal $\xi \in C_c^\infty(\mathbb{R}^6;\mathbb{R}^3)$ in the right-hand side of (3.24) needs to approximate $\tilde\nabla\sqrt{ff_*}$. We have formulated (3.24) in a big space of test functions without taking advantage of the anti-symmetry of variable swapping $v \leftrightarrow v_*$. To take advantage of symmetries, we define

$$AS = \{V = V(v,v_*) \in L^2(\mathbb{R}^6;\mathbb{R}^3) \,|\, V(v,v_*) = -V(v_*,v) \text{ a.e. } v, v_* \in \mathbb{R}^3\},$$

together with the smooth and compactly supported approximations

$$AS_c^\infty = \left\{ V = V(v,v_*) \in C_c^\infty(\mathbb{R}^6;\mathbb{R}^3) \;\middle|\; \begin{array}{l} V(v,v_*) = -V(v_*,v)\, \forall v,\, v_* \in \mathbb{R}^3, \\ \exists\delta>0 \text{ s.t. } V(v,v_*) = 0\ \forall |v-v_*| \le \delta \end{array} \right\}.$$

The anti-symmetry allows to write the following identity

$$V(v,v_*) = \frac{1}{2}(V(v,v_*) - V(v_*,v)), \quad \forall V \in AS. \tag{3.25}$$

To shorten some notation, we will write

$$\overset{\leftrightarrow}{V}(v, v_*) := V(v_*, v), \quad \forall v, v_* \in \mathbb{R}^3.$$

Using (3.25), it is easy to see that AS is a closed subspace of $L^2(\mathbb{R}^6; \mathbb{R}^3)$ and hence is a Hilbert space with the L^2 inner product. Moreover, we have density of the smooth compactly supported approximations.

Lemma 3.15. *AS_c^∞ is dense in AS with respect to the L^2 topology.*

We skip the proof of Lemma 3.15. Using density we can show the next characterisation.

Proposition 3.16. *The Landau dissipation can be written as*

$$D_L(f) = \sup_{V \in AS_c^\infty} \left\{ -4 \iint_{\mathbb{R}^6} \sqrt{f f_*} |v - v_*|^{1+\frac{\gamma}{2}} (\nabla - \nabla_*) \cdot (\Pi[v - v_*]V) - 2 \iint_{\mathbb{R}^6} |\Pi[v - v_*]V|^2 \right\}. \quad (3.26)$$

Proof. Replace $L^2(\mathbb{R}^6; \mathbb{R}^3)$ and $C_c^\infty(\mathbb{R}^6; \mathbb{R}^3)$ with AS and AS_c^∞, respectively and follow the same proof of Proposition 3.14 taking advantage of (3.25). This would lead to the majorant

$$\sup_{V \in AS_c^\infty} \left\{ -4 \iint_{\mathbb{R}^6} \sqrt{f f_*} |v - v_*|^{1+\frac{\gamma}{2}} (\nabla - \nabla_*) \cdot (\Pi[v - v_*]V) - 2 \iint_{\mathbb{R}^6} |V|^2 \right\}.$$

Since Π is a projection, we have $|\Pi[v - v_*]V| \le |V|$. Estimating the second term of the affine representation in this way completes the proof. ☐

3.5.1.2 Projecting a vector field onto the image of $\tilde{\nabla}$

Starting from (3.26), our goal now is to replace the vector field $V \in AS_c^\infty$ by $\tilde{\nabla}\psi$ for some $\psi \in DS_c^\infty$. In this section, the role of DS_c^∞ which we introduced in the proof of Proposition 3.13 will be clarified. Given any $V \in AS_c^\infty$, we find ψ such that

$$\begin{aligned} -4 \iint_{\mathbb{R}^6} \sqrt{f f_*} |v - v_*|^{1+\frac{\gamma}{2}} (\nabla - \nabla_*) \cdot (\Pi[v - v_*]V) - 2 \iint_{\mathbb{R}^6} |\Pi[v - v_*]V|^2 \\ \le -4 \iint_{\mathbb{R}^6} \sqrt{f f_*} \tilde{\nabla} \cdot \tilde{\nabla}\psi - 2 \iint_{\mathbb{R}^6} |\tilde{\nabla}\psi|^2. \end{aligned}$$

Said ψ can be characterised by the projection of V (or equivalently $\Pi[v - v_*]V$) onto the image of $\tilde{\nabla}$. More explicitly, ψ will be obtained as the solution to the following minimisation problem

$$\min_{g \in DS_c^\infty} \iint_{\mathbb{R}^6} |\tilde{\nabla} g - \Pi[v - v_*]V|^2 = \iint_{\mathbb{R}^6} |\tilde{\nabla}\psi - \Pi[v - v_*]V|^2. \quad (3.27)$$

We begin by investigating the solvability of the first order condition of this convex minimisation problem which is given by the PDE in the following result.

Lemma 3.17. *For $V \in AS_c^\infty$, there exists a unique smooth solution $\psi \in DS_c^\infty$ to the following equation*

$$\tilde{\nabla} \cdot \tilde{\nabla}\psi = \tilde{\nabla} \cdot (\Pi[v - v_*]V). \tag{3.28}$$

Moreover, ψ solves (3.27) *and we have*

$$\|\Pi[v - v_*]V\|_{L^2_{v,v_*}}^2 = \|\tilde{\nabla}\psi\|_{L^2_{v,v_*}}^2 + \|\Pi[v - v_*]V - \tilde{\nabla}\psi\|_{L^2_{v,v_*}}^2. \tag{3.29}$$

Proof. After changing variables, we will construct ψ as a superposition of solutions to the Laplace-Beltrami equation on spheres $|v - v_*| = r$. Recalling some of the notations from Section 3.2.3, for fixed $v, v_* \in \mathbb{R}^3$ we consider the smooth and volume preserving coordinate transformation

$$(v, v_*) \mapsto \Big(\frac{v - v_*}{2}, \frac{v + v_*}{2}\Big) =: (x, y).$$

Given vector fields $V = V(v, v_*)$ and scalar functions $\psi = \psi(v, v_*)$, we will use the same symbols to denote their versions under this and future coordinate transformations $V = V(v, v_*) = V(x, y)$ and similarly for ψ. It is readily checked that $\nabla - \nabla_* = \nabla_x$ and similarly for the divergence. Notice that when $V \in AS_c^\infty$, (3.28) reads

$$|2x|^{2+\gamma}\nabla_x \cdot (\Pi[x]\nabla_x\psi(x, y)) = |2x|^{1+\frac{\gamma}{2}}\nabla_x \cdot (\Pi[x]V(x, y)),$$

which is an equation in the $x = \frac{1}{2}(v - v_*)$ variable only. Henceforth, we consider fixed $y = \frac{1}{2}(v + v_*)$ as a parameter to the problem above. To further specify the problem, the compact support and anti-symmetry of $V \in AS_c^\infty$ translate into compactness in both the x, y variables, and moreover V vanishes in a neighbourhood of $\{x = 0\}$. So for some $0 < \delta \leq R$ depending on the support of V, but uniform in the $y = \frac{1}{2}(v+v_*)$ variable, we consider the following elliptic PDE with homogeneous Dirichlet boundary conditions

$$\begin{cases} |2x|^{1+\frac{\gamma}{2}}\nabla_x \cdot (\Pi[x]\nabla_x\psi(x, y)) = \nabla_x \cdot (\Pi[x]V(x, y)), & 0 < \delta \leq |2x| \leq R \\ \psi(x, y) = 0, & |2x| \in \{\delta, R\} \end{cases}. \tag{3.30}$$

The weight $|2x|^{1+\frac{\gamma}{2}}$ on the left-hand side is well-behaved, since we avoid a neighbourhood of the singularity $x = 0$. To reiterate, we will solve (3.30) in x for fixing the value of y as a parameter. Since the dependence on y of V is smooth, it will follow any solution ψ of (3.30) is also smooth in y. In terms of solvability of (3.30), we make one further change of variables. Having fixed y as a parameter, we consider the spherical decomposition of

$$x = rk, \quad r = |x|, \quad k = \frac{x}{|x|} \in \mathbb{S}^2.$$

Under these coordinates, we again identify $\psi = \psi(x, y) = \psi(k, r, y)$ and similarly for V. By the identities of Lemma A.5 and Corollary A.3, we can consider $r \in [\delta, R]$ as another parameter so that (3.30) becomes an equation in the spherical variable k

$$\begin{cases} \Delta_{\mathbb{S}^2}\psi(k, r, y) = 2^{-1-\frac{\gamma}{2}} r^{-\frac{\gamma}{2}} \nabla_k \cdot (\Pi[k]V(k, r, y)) & 0 < \delta \leq 2r \leq R, k \in \mathbb{S}^2 \\ \psi(k, r, y) = 0 & 2r \in \{\delta, R\} \end{cases}. \tag{3.31}$$

The interpretation of (3.31) is that, at every level set of $|x|$, (3.30) is actually the Poisson problem for the Laplace-Beltrami operator in $\mathbb{S}^2$ for the $k = \frac{x}{|x|}$ variable. Noticing that the right-hand side is a divergence of a smooth function, the integral over $\mathbb{S}^2$ of the right-hand side vanishes, which is a necessary condition for solvability of the Poisson problem in a compact manifold. Using the usual method of the Lax-Milgram Theorem combined with the Poincaré inequality on the sphere (see [81]), for each fixed $r \in (\delta, R)$ and $y \in \mathbb{R}^3$ (3.31) admits a unique weak solution $\psi(\cdot, r, y) \in H^1(\mathbb{S}^2)$ which is also smooth by standard elliptic regularity arguments. Finally, since V is anti-symmetric when swapping $v \leftrightarrow v_*$ (which means reflecting $x \leftrightarrow -x$, or $k \leftrightarrow -k$), uniqueness of solutions gives that ψ is symmetric

$$\psi(v, v_*) = \psi(v_*, v) \iff \psi(x, y) = \psi(-x, y) \iff \psi(k, r, y) = \psi(-k, r, y).$$

Based on the regularity and symmetries of $V \in AS_c^\infty$, the previous discussion implies $\psi \in DS_c^\infty$. Returning to the minimisation problem of (3.27), we can deduce from convexity and our discussion on the solubility of the first order conditions (3.28) that there exists a unique $\psi \in DS_c^\infty$ such that

$$\inf_{g \in DS_c^\infty} \|\tilde{\nabla} g - \Pi[v - v_*]V\|^2_{L^2_{v,v_*}} = \|\tilde{\nabla}\psi - \Pi[v - v_*]V\|^2_{L^2_{v,v_*}}.$$

We can interpret the solution operator for (3.31) as the orthogonal projection map of V and $\Pi[v - v_*]V$ to the image of $\tilde{\nabla}$. To see (3.29), we add and subtract $\tilde{\nabla}\psi$ in the L^2 norm of $\Pi[v - v_*]V$ to get

$$\begin{aligned} \|\Pi[v - v_*]V\|^2 = & \iint_{\mathbb{R}^6} |\Pi[v - v_*]V - \tilde{\nabla}\psi + \tilde{\nabla}\psi|^2 \\ = & \|\tilde{\nabla}\psi\|^2 + \|\Pi[v - v_*]V - \tilde{\nabla}\psi\|^2 + 2 \iint_{\mathbb{R}^6} \tilde{\nabla}\psi \cdot (\Pi[v - v_*]V - \tilde{\nabla}\psi) \\ = & \|\tilde{\nabla}\psi\|^2 + \|\Pi[v - v_*]V - \tilde{\nabla}\psi\|^2 - 2 \iint_{\mathbb{R}^6} \psi \underbrace{(\tilde{\nabla} \cdot (\Pi[v - v_*]V) - \tilde{\nabla} \cdot \tilde{\nabla}\psi)}_{=0}. \end{aligned}$$

The last line is obtained by integrating by parts the differential operator $\tilde{\nabla}$ using the smoothness and compact support of ψ and V. Of course, by our construction of ψ, the cross term contributes nothing owing to (3.28). $\square$

Using Lemma 3.17, we can further majorise the Landau dissipation from (3.26)

$$D_L(f) \leq \sup_{\psi \in DS_c^\infty} \left\{ -4 \iint_{\mathbb{R}^6} \sqrt{ff_*}\tilde{\nabla} \cdot \tilde{\nabla}\psi - 2 \iint_{\mathbb{R}^6} |\tilde{\nabla}\psi|^2 \right\}. \tag{3.32}$$

3.5.1.3 Boltzmann dissipation

Before directly manipulating the Boltzmann dissipation, we insist on the appearance of a finite difference of $\sqrt{ff_*}$. Using Corollary A.2, we can lower bound the Boltzmann dissipation by

$$D_B(f) = \frac{1}{4} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} [f'f'_* - ff_*] \log \frac{f'f'_*}{ff_*} B \geq \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |\sqrt{f'f'_*} - \sqrt{ff_*}|^2 B.$$

Let us refer to the lower bound as the reduced Boltzmann dissipation

$$D_B^R(f) := \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |\sqrt{f'f'_*} - \sqrt{ff_*}|^2 B \, d\sigma dv_* dv.$$

We want a similar representation for the reduced Boltzmann dissipation as we had for the Landau dissipation.

Lemma 3.18. *The reduced Boltzmann dissipation can be expressed as*

$$D_B^R(f) = \sup_{\tilde{\psi} \in L^2(\mathbb{R}^6 \times \mathbb{S}^2; Bd\sigma dv_* dv)} \left\{ 2 \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} (\sqrt{f'f'_*} - \sqrt{ff_*})\tilde{\psi} B d\sigma dv_* dv - \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |\tilde{\psi}|^2 B d\sigma dv_* dv \right\}. \tag{3.33}$$

Proof. This proof is analogous to the proof of Proposition 3.14 in the Hilbert framework of L^2 with respect to the collision kernel B. $\square$

As with the Landau dissipation, we seek to pass all the difference structure onto the test function. Taking advantage of various pre-post collision velocity symmetries, we extend Lemma 3.18 to

Lemma 3.19. *The reduced Boltzmann dissipation can be minorised by*

$$D_B^R(f) \geq \sup_{\psi \in DS_c^\infty} \left\{ -2 \iint_{\mathbb{R}^6} \sqrt{ff_*} \left(\int_{\mathbb{S}^2} \bar{\nabla}\psi B \right) - \frac{1}{4} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |\bar{\nabla}\psi|^2 B \right\}.$$

Proof. We are interested in showing the inequality, hence we only need to show that each $\psi \in DS_c^\infty$ induces an admissible $\tilde{\psi} \in L^2(\mathbb{R}^6 \times \mathbb{S}^2; Bd\sigma dv_* dv)$ and then de-symmetrise the first term in (3.33).

More specifically, for each $\psi \in DS_c^\infty$

$$\frac{1}{2}\bar{\nabla}\psi = \psi(v', v'_*) - \psi(v, v_*) =: \tilde{\psi}(v, v_*, \sigma) \in L^2(\mathbb{R}^6 \times \mathbb{S}^2; Bd\sigma dv_* dv)$$

is an admissible test function for (3.33). This is a consequence of Lemma 3.4 which provides the estimate

$$|\psi' - \psi|^2 \lesssim \mathbb{1}_{0<\delta\leq|v-v_*|\leq R}|v - v_*|^2|\sigma - k|^2, \tag{3.34}$$

where $\delta, R > 0$ are the inner and outer radii of the support of $\psi \in DS_c^\infty$ with respect to $v - v_*$.

Next, we de-symmetrise the first term of (3.33)

$$\begin{aligned} 2\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}(\sqrt{f'f'_*} - \sqrt{ff_*})\tilde{\psi}B &= 2\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\sqrt{f'f'_*}\int_{\mathbb{S}^2}(\psi' - \psi)B - 2\iint_{\mathbb{R}^6}\sqrt{ff_*}\int_{\mathbb{S}^2}(\psi' - \psi)B \\ &= -4\iint_{\mathbb{R}^6}\sqrt{ff_*}\int_{\mathbb{S}^2}(\psi' - \psi)B, \end{aligned}$$

which is justified as each of the integrals above are absolutely convergent. This follows from the estimate (3.34), $|\sigma - k|^2 \sim \theta^2$ from (3.11), and the finite angular momentum transfer assumption (3.2). The desired inequality now follows. □

3.5.2 Boltzmann gradient converges to Landau gradient

The aim of this section is to show that for arbitrary $\psi \in DS_c^\infty$, we obtain

$$\iint_{\mathbb{R}^6}\sqrt{f^\epsilon f^\epsilon_*}\left(\int_{\mathbb{S}^2}\bar{\nabla}\psi B^\epsilon\right) \overset{\epsilon\downarrow 0}{\to} 2\iint_{\mathbb{R}^6}\sqrt{ff_*}\tilde{\nabla}\cdot\tilde{\nabla}\psi \tag{3.35}$$

$$\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|\bar{\nabla}\psi|^2B^\epsilon \overset{\epsilon\downarrow 0}{\to} 8\iint_{\mathbb{R}^6}|\tilde{\nabla}\psi|^2. \tag{3.36}$$

These limits are the final pieces needed to finish the proof of Proposition 3.13. Recall there that we had reduced to the case with bounded dissipation for a fixed time

$$\sup_{\epsilon>0} D_B^\epsilon(f^\epsilon) < +\infty,$$

which implies the local L^2_{loc} convergence $\sqrt{f^\epsilon} \to \sqrt{f}$ which we will use in the proof below.

The key to these limits is understanding the limiting behaviour of $\bar{\nabla}$ on DS_c^∞ functions. Naturally, the crucial ingredients are Lemmas 3.5 and 3.6 which state

$$\forall\psi \in DS_c^\infty, \quad \begin{cases} \frac{1}{\epsilon}\bar{\nabla}\psi & \overset{\epsilon\downarrow 0}{\to} & \frac{\chi}{\pi}|v - v_*|p\cdot(\nabla - \nabla_*)\psi \\ \frac{1}{\epsilon^2}\int_{\mathbb{S}^1_{k^\perp}}\bar{\nabla}\psi dp & \overset{\epsilon\downarrow 0}{\to} & \frac{\chi^2}{4\pi}(\nabla - \nabla_*)\cdot(|v - v_*|^2\Pi[v - v_*](\nabla - \nabla_*)\psi) \end{cases}. \tag{3.37}$$

Recall the notation that $\chi = \pi\theta/\epsilon$ is the rescaled angle of collision and $p \in \mathbb{S}^2$ is the orthogonal vector to $k = \frac{v-v_*}{|v-v_*|}$ shown in Figures 3.1 to 3.3. These limits hold in the pointwise sense.

Proof of (3.35) *and* (3.36). We begin with showing (3.35). Since we know $\sqrt{f^\epsilon} \to \sqrt{f}$, it remains to show that

$$\int_{\mathbb{S}^2} \bar{\nabla}\psi B^\epsilon = |v-v_*|^\gamma \int_{\theta=0}^{\epsilon/2} \beta^\epsilon(\theta) \int_{\mathbb{S}^1_{k^\perp}} \bar{\nabla}\psi \, dpd\theta \overset{\epsilon\downarrow 0}{\to} 2\tilde{\nabla}\cdot\tilde{\nabla}\psi.$$

Since $\psi \in DS_c^\infty$, we are in the nice situation of avoiding all problems in the v, v_* variables because ψ is supported in

$$|v| + |v_*| \leq R, \quad 0 < \delta \leq |v-v_*| \leq R,$$

for some $\delta, R > 0$. Combining this localisation with the third estimate of Lemma 3.4,

$$\left| \int_{\mathbb{S}^1_{k^\perp}} \bar{\nabla}\psi dp \right| \lesssim_\psi |v-v_*|^2\theta^2,$$

we get domination in L^2 recalling finite angular momentum transfer (3.2), so we only have to show pointwise a.e. convergence of the previous limit. By rescaling $\theta = \epsilon\chi/\pi$ and applying (3.37), we have

$$\begin{aligned} \int_{\mathbb{S}^2} \bar{\nabla}\psi B^\epsilon &= |v-v_*|^\gamma \int_0^{\pi/2} \beta(\chi)\frac{\pi^2}{\epsilon^2} \int_{\mathbb{S}^1_{k^\perp}} \bar{\nabla}\psi dp d\chi \\ \overset{\epsilon\downarrow 0}{\to} |v-v_*|^\gamma \frac{\pi}{4} \int_0^{\pi/2} \chi^2\beta(\chi)(\nabla-\nabla_*)\cdot(|v-v_*|^2\Pi[v-v_*](\nabla-\nabla_*)\psi) &= 2\tilde{\nabla}\cdot\tilde{\nabla}\psi, \end{aligned}$$

again recalling the normalisation $\int_0^{\pi/2} \chi^2\beta(\chi)d\chi = 8/\pi$.

We turn to showing (3.36). Arguing as in the proof of (3.35), since ψ is compactly supported in v, v_*, and in $\{|v-v_*| \geq \delta > 0\}$, we only need to show

$$\int_{\mathbb{S}^2} |\bar{\nabla}\psi|^2 B^\epsilon = |v-v_*|^\gamma \int_{\theta=0}^{\epsilon/2} \beta^\epsilon(\theta) \int_{\mathbb{S}^1_{k^\perp}} |\bar{\nabla}\psi|^2 dpd\theta \overset{\epsilon\downarrow 0}{\to} 8|\tilde{\nabla}\psi|^2, \quad \text{a.e. } v, v_* \in \mathbb{R}^3.$$

This is because of the first estimate of Lemma 3.4 which again gives the right majorant against B^ϵ

$$|\bar{\nabla}\psi|^2 \lesssim_\psi |v-v_*|^2\theta^2.$$

By the same rescaling $\theta = \epsilon\chi/\pi$, we have

$$\begin{aligned}
\int_{\mathbb{S}^2} |\bar\nabla\psi|^2 B^\epsilon &= |v-v_*|^\gamma \int_0^{\pi/2} \beta(\chi)\pi^2 \int_{\mathbb{S}^1_{k\perp}} \left|\frac{1}{\epsilon}\bar\nabla\psi\right|^2 dpd\chi \\
&\overset{\epsilon\downarrow 0}{\to} |v-v_*|^{2+\gamma} \int_0^{\pi/2} \beta(\chi)\chi^2 \int_{\mathbb{S}^1_{k\perp}} |p\cdot(\nabla-\nabla_*)\psi|^2 dpd\chi \\
&= \frac{8}{\pi}|v-v_*|^{2+\gamma} \int_{\mathbb{S}^1_{k\perp}} p^i p^j (\nabla-\nabla_*)^i\psi(\nabla-\nabla_*)^j\psi dp \\
&= 8|v-v_*|^{2+\gamma}|\Pi[v-v_*](\nabla-\nabla_*)\psi|^2.
\end{aligned}$$

In the last line, we have used Lemma A.4 for the computation of $\int_{\mathbb{S}^1_{k\perp}} p\otimes pdp = \pi\Pi[k]$, while remembering that since $\Pi[k]$ is a projection matrix, the quadratic form satisfies $z^T\Pi[k]z = |\Pi[k]z|^2$, for all $z\in\mathbb{R}^3$. □

3.6 Lower semicontinuous convergence of metric derivatives in the grazing collision limit

We consider a sequence of curves $(f^\epsilon)_{\epsilon>0}$ satisfying the uniform moment and metric derivative bounds (3.5). In particular, we know that a subsequence converging to f exists by Theorem 3.9. Along this subsequence and in parallel to the general affine representation strategy in Section 3.5, we seek to prove

Proposition 3.20. *Consider a sequence of curves $(f^\epsilon)_{\epsilon>0}$ satisfying the uniform bounds in second moment, entropy, dissipation, and metric derivative* (3.5)*. Along possibly a further subsequence for which $f^\epsilon \overset{\sigma}{\rightharpoonup} f$ from Theorem 3.9, we have*

$$\liminf_{\epsilon\downarrow 0} |\dot f^\epsilon|^2_\epsilon(t) \ge |\dot f|^2_L(t), \quad a.e.\ t\in[0,T].$$

As in Section 3.5.1, we will prove Proposition 3.20 by looking at the affine representations of the metric derivatives. Without loss of generality, we assume $\sup_{\epsilon>0}|\dot f^\epsilon|^2_\epsilon(t) < +\infty$ by taking a subsequence, if necessary.

Lemma 3.21. *We consider $(f^\epsilon, M^\epsilon_B)$ and (f,M) curves for CRE (in the sense of Erbar [61]) and GCE (in the sense of Chapter 2), respectively. We assume that M^ϵ_B and M are optimal collision and grazing rates in the sense that their associated metric derivatives can be written as the respective action of these curves*

$$|\dot f^\epsilon|^2_\epsilon(t) = \frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} \frac{|M^\epsilon_B|^2}{\Lambda(f^\epsilon)B^\epsilon} d\sigma dv_* dv \qquad and \qquad |\dot f|^2_L(t) = \frac{1}{2}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\frac{|M|^2}{ff_*}dv_*dv.$$

Then, we have the following affine representations.

$$
\begin{aligned}
|\dot{f}^\epsilon|_\epsilon^2(t) &= \sup_{\psi \in C_c^\infty(\mathbb{R}^3)} \left\{ \frac{1}{2} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} M_B^\epsilon \bar{\nabla}\psi - \frac{1}{4} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |\bar{\nabla}\psi|^2 \Lambda(f^\epsilon) B^\epsilon \right\}, \\
|\dot{f}|_L^2(t) &= \sup_{\psi \in C_c^\infty(\mathbb{R}^3)} \left\{ \iint_{\mathbb{R}^6} M \cdot \tilde{\nabla}\psi - \frac{1}{2} \iint_{\mathbb{R}^6} |\tilde{\nabla}\psi|^2 f f_* \right\}.
\end{aligned}
\tag{3.38}
$$

Proof. We only show the proof for the first of these claimed equations since the argument for the Landau metric derivative is analogous (using Proposition 2.22 instead of Lemma 3.8). We shall also drop the superscript ϵ and abbreviate $\Lambda = \Lambda(f^\epsilon)$ for ease of notation. Arguing as in Section 3.5.1.1, we have

$$
\frac{1}{4} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{|M_B|^2}{\Lambda B} = \sup_{\xi \in L^2(\Lambda B)} \left\{ \frac{1}{2} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} M_B \xi - \frac{1}{4} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |\xi|^2 \Lambda B \right\}
\tag{3.39}
$$

which follows from similar lines of reasoning as Proposition 3.14 and Lemma 3.18.

We now want to replace the test function $\xi \in L^2(\Lambda B)$ by $\bar{\nabla}\psi$ for $\psi \in C_c^\infty(\mathbb{R}^3)$. Using the optimality of M_B, the finite action assumption, and Lemma 3.8, we can write $M_B = U \Lambda B$ where the density U can be approximated by

$$
U \in \overline{\{\bar{\nabla}\psi \,|\, \psi \in C_c^\infty(\mathbb{R}^3)\}}^{L^2(\Lambda B)}.
$$

Take $(\psi_n)_{n\in\mathbb{N}} \subset C_c^\infty(\mathbb{R}^3)$ an approximating sequence so that $\bar{\nabla}\psi_n \to U$ in $L^2(\Lambda B)$ as $n \to \infty$. The expression inside the supremum with $\xi = \bar{\nabla}\psi_n$ of (3.39) has the following limit

$$
\begin{aligned}
&\iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{M_B}{\Lambda B} \bar{\nabla}\psi_n \Lambda B - \frac{1}{2} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |\bar{\nabla}\psi_n|^2 \Lambda B \\
\overset{n\to\infty}{\to} &\iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{M_B}{\Lambda B} U - \frac{1}{2} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |U|^2 \Lambda B \\
= &\iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{|M_B|^2}{2\Lambda B}.
\end{aligned}
$$

This establishes the first equation of (3.38). As mentioned before, the second equation of (3.38) follows analogously. □

With Lemma 3.21, we are in a position to prove Proposition 3.20.

Proof of Proposition 3.20. Without loss of generality, take M_B^ϵ optimal collision rates in the sense of Lemma 3.8 so that the Boltzmann metric derivative is the action of $(f^\epsilon, M_B^\epsilon)$

$$
|\dot{f}^\epsilon|_\epsilon^2(t) = \frac{1}{4} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{|M_B^\epsilon|^2}{\Lambda(f^\epsilon) B^\epsilon} d\sigma dv_* dv.
$$

By the computations of Section 3.4.1,

$$\text{for } q = \begin{cases} 1, & \gamma \in [-2,0) \\ 2, & \gamma \in [-4,-2) \end{cases}, \text{ sufficiently small } 0 < \delta < \begin{cases} -\frac{\gamma}{2}, & \gamma \in [-2,0) \\ -\frac{\gamma}{2} - 1, & \gamma \in [-4,-2) \end{cases},$$

and along some subsequence $\epsilon \downarrow 0$, we have that $(f, L_{q,\delta}(M_{q,\delta}))$ is an admissible pair in the grazing continuity equation where

$$|v - v_*|^q \left(1 + \left[|v|^2 + |v_*|^2\right]^{\frac{\delta}{2}}\right) \theta M^\epsilon \overset{\epsilon\downarrow 0}{\rightharpoonup} M_{q,\delta}, \quad L_{q,\delta}(M_{q,\delta}) = \frac{|v - v_*|^{-\gamma/2-q}}{4\left(1 + \left[|v|^2 + |v_*|^2\right]^{\frac{\delta}{2}}\right)} \int_{\mathbb{S}^2} M_{q,\delta} p d\sigma.$$

As well, since $L_{q,\delta}(M_{q,\delta})$ is admissible, there exists (by Lemma 3.8) a unique grazing rate $M \in \mathcal{M}_L$ so that we have the following inequality for the Landau metric derivative

$$|\dot{f}|_L^2(t) = \frac{1}{2} \iint_{\mathbb{R}^6} \frac{|M|^2}{f f_*} dv_* dv \leq \frac{1}{2} \iint_{\mathbb{R}^6} \frac{|L_{q,\delta}(M_{q,\delta})|^2}{f f_*} dv_* dv.$$

Recall the affine representations of the metric derivatives from Lemma 3.21

$$\frac{1}{4} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \frac{|M_B^\epsilon|^2}{\Lambda(f^\epsilon) B^\epsilon} = \sup_{\psi \in C_c^\infty(\mathbb{R}^3)} \left\{ \frac{1}{2} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} M_B^\epsilon \bar{\nabla} \psi - \frac{1}{4} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} |\bar{\nabla}\psi|^2 \Lambda(f^\epsilon) B^\epsilon \right\}$$

and

$$\frac{1}{2} \iint_{\mathbb{R}^6} \frac{|M|^2}{f f_*} = \sup_{\psi \in C_c^\infty(\mathbb{R}^3)} \left\{ \iint_{\mathbb{R}^6} M \cdot \tilde{\nabla}\psi - \frac{1}{2} \iint_{\mathbb{R}^6} |\tilde{\nabla}\psi|^2 f f_* \right\}.$$

It remains to establish that, for the $L_{q,\delta}(M_{q,\delta})$ rate (coming from the sequence $\{M_B^\epsilon\}_\epsilon$ in CRE) and the corresponding unique optimal rate M (coming directly GCE), we must have

$$\begin{aligned} &\sup_{\psi \in C_c^\infty(\mathbb{R}^3)} \left\{ \iint_{\mathbb{R}^6} M \cdot \tilde{\nabla}\psi - \frac{1}{2} \iint_{\mathbb{R}^6} |\tilde{\nabla}\psi|^2 f f_* \right\} \\ &= \sup_{\psi \in C_c^\infty(\mathbb{R}^3)} \left\{ \iint_{\mathbb{R}^6} L_{q,\delta}(M_{q,\delta}) \cdot \tilde{\nabla}\psi - \frac{1}{2} \iint_{\mathbb{R}^6} |\tilde{\nabla}\psi|^2 f f_* \right\}. \end{aligned} \tag{3.40}$$

This follows because

$$M = L_{q,\delta}(M_{q,\delta}) + \underbrace{(M - L_{q,\delta}(M_{q,\delta}))}_{\tilde{\nabla}\cdot\text{-free}}.$$

Recall from Proposition 2.22 '$\tilde{\nabla}\cdot$-free' means that for any $\psi \in C_c^\infty(\mathbb{R}^3)$, we have

$$\iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} \tilde{\nabla}\psi \cdot d(M - L_{q,\delta}(M_{q,\delta})) = 0,$$

which follows since both M and $L_{q,\delta}(M_{q,\delta})$ solve the grazing continuity equation with the same f

$$\frac{d}{dt}\int_{\mathbb{R}^d}\psi f_t(v)dv + \frac{1}{2}\iint_{\mathbb{R}^6}\tilde{\nabla}\psi\cdot dM_t(v,v_*) = \frac{d}{dt}\int_{\mathbb{R}^d}\psi f_t(v)dv + \frac{1}{2}\iint_{\mathbb{R}^6}\tilde{\nabla}\psi\cdot d(L_{q,\delta}(M_{q,\delta}))(v,v_*),$$

for all $\psi\in C_c^\infty(\mathbb{R}^3)$. Looking at the level of the dual formulations, for fixed $\psi\in C_c^\infty(\mathbb{R}^3)$ we claim

$$\frac{1}{2}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} M_B^\epsilon\bar{\nabla}\psi \overset{\epsilon\downarrow 0}{\to} \iint_{\mathbb{R}^6} L_{q,\delta}(M_{q,\delta})\cdot\tilde{\nabla}\psi, \tag{3.41}$$

$$\limsup_{\epsilon\downarrow 0}\frac{1}{4}\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|\bar{\nabla}\psi|^2\Lambda(f^\epsilon)B^\epsilon \le \frac{1}{2}\iint_{\mathbb{R}^6}|\tilde{\nabla}\psi|^2 ff_*. \tag{3.42}$$

Once (3.41) and (3.42) are proven for fixed $\psi\in C_c^\infty$, this establishes

$$\liminf_{\epsilon\downarrow 0}|\dot{f}^\epsilon|_B^2(t) \ge \sup_{\psi\in C_c^\infty(\mathbb{R}^3)}\left\{\iint_{\mathbb{R}^6} L_{q,\delta}(M_{q,\delta})\cdot\tilde{\nabla}\psi - \frac{1}{2}\iint_{\mathbb{R}^6}|\tilde{\nabla}\psi|^2 ff_*\right\}.$$

Recalling (3.40), the right-hand side is the affine representation of $|\dot{f}|_L^2(t)$ we can conclude the proof.

To prove (3.41), we use the estimates for $\bar{\nabla}\psi$ from Lemma 3.4 combined with the appropriate choice of $q\in\left[-\frac{\gamma}{2},1-\frac{\gamma}{2}\right)$ and small $\delta>0$ depending on γ. We recall the computations from Section 3.4.1 - the first step is to reveal the correct sequence of measures from the scaled compactness Proposition 3.11;

$$\begin{aligned}
&\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} M_B^\epsilon\bar{\nabla}\psi \\
&= \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\left\{|v-v_*|^q\left(1+\left[|v|^2+|v_*|^2\right]^{\frac{\delta}{2}}\right)\theta M_B^\epsilon\right\}|v-v_*|^{-q}\left(1+\left[|v|^2+|v_*|^2\right]^{\frac{\delta}{2}}\right)^{-1}\frac{\bar{\nabla}\psi}{\theta} \\
&\overset{\epsilon\downarrow 0}{\to}\frac{1}{2}\iint_{\mathbb{R}^6}|v-v_*|^{-q-\frac{\gamma}{2}}\left(1+\left[|v|^2+|v_*|^2\right]^{\frac{\delta}{2}}\right)^{-1}\tilde{\nabla}\psi\cdot\int_{\mathbb{S}^2}M_{q,\delta}p\,d\sigma dv_*dv \\
&= 2\iint_{\mathbb{R}^6}\tilde{\nabla}\psi\cdot L_{q,\delta}(M_{q,\delta})dv_*dv.
\end{aligned}$$

The justification for the weak-strong convergence is also found in Section 3.4.1.

Turning to the proof of (3.42), we use Corollary A.2 and symmetry to write

$$\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|\bar{\nabla}\psi|^2B^\epsilon\Lambda(f^\epsilon) \le \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|\bar{\nabla}\psi|^2B^\epsilon\left[\frac{f^{\epsilon\prime}f^{\epsilon\prime}_* + f^\epsilon f^\epsilon_*}{2}\right] = \iint_{\mathbb{R}^6}f^\epsilon f^\epsilon_*\left(\int_{\mathbb{S}^2}|\bar{\nabla}\psi|^2B^\epsilon\right).$$

We first provide an integrable majorant (uniformly against the measure $f^\epsilon f^\epsilon_*$) for the term integrated over $\mathbb{S}^2$ which is again provided by Lemma 3.4. We have

$$\int_{\mathbb{S}^2} |\bar{\nabla}\psi|^2 B^\epsilon = \pi^2 |v-v_*|^\gamma \int_0^{\pi/2} \beta(\chi) \int_{\mathbb{S}^1_{k^\perp}} \left|\frac{1}{\epsilon}\bar{\nabla}\psi\right|^2 dp d\chi$$

$$\lesssim \int_0^{\pi/2} \chi^2 \beta(\chi) d\chi \begin{cases} \mathrm{Lip}(\psi)|v-v_*|^{2+\gamma}, & \gamma \in [-2,0) \\ \|D^2\psi\|_{L^\infty} |v-v_*|^{4+\gamma}, & \gamma \in [-4,-2) \end{cases} \lesssim_{\psi,\beta} 1 + |v-v_*|^l.$$

Here $0 < l < 2$ is some power which gives strictly subquadratic growth. By the uniformly bounded moments assumption (3.5), this is uniformly integrable against $f^\epsilon f^\epsilon_*$ so we can pass to the weak-strong limit. According to Lemma 3.5, we have the pointwise limit

$$\int_{\mathbb{S}^2} |\bar{\nabla}\psi|^2 B^\epsilon = \pi^2 |v-v_*|^\gamma \int_0^{\pi/2} \beta(\chi) \int_{\mathbb{S}^1_{k^\perp}} \left|\frac{1}{\epsilon}\bar{\nabla}\psi\right|^2 dp d\chi$$

$$\overset{\epsilon\downarrow 0}{\to} \frac{1}{4}|v-v_*|^{2+\gamma} \int_0^{\pi/2} \chi^2 \beta(\chi) \int_{\mathbb{S}^1_{k^\perp}} |p\cdot(\nabla - \nabla_*)\psi|^2 dp d\chi$$

$$= 2|v-v_*|^{2+\gamma} |\Pi[v-v_*](\nabla-\nabla_*)\psi|^2 = 2|\tilde{\nabla}\psi|^2.$$

Here, we have used Lemma A.4 and the usual finite angular momentum transfer (3.2). □

Chapter 4

A particle method for the homogeneous Landau equation

The content of this chapter is based on joint work with José A. Carillo, Jingwei Hu, and Li Wang. It has been published [33] in the Journal of Computational Physics: X volume 7, page 100066 in June 2020.

We propose a novel deterministic particle method to numerically approximate the Landau equation for plasmas. In particular, by fixing the mollifying sequence ψ_ε (scaled Maxwellian distributions as in (ψ_ε)), this induces the regularised entropy functional we consider

$$\mathcal{H}_{2,\varepsilon}[f] := \int_{\mathbb{R}^d} [f * \psi_\varepsilon](v) \log [f * \psi_\varepsilon](v) dv. \tag{4.1}$$

Moreover, the regularised equation studied here is the $\mathcal{H}_{2,\varepsilon}$ Landau equation (in the sense of (2.1))

$$\partial_t f = \nabla \cdot \left(f \int_{\mathbb{R}^d} f_* |v - v_*|^{2+\gamma} \Pi[v - v_*] \left(\nabla \psi_\varepsilon * \log [f * \psi_\varepsilon] - \nabla_* \psi_\varepsilon * \log [f_* * \psi_\varepsilon]\right)\right). \tag{4.2}$$

This regularisation suppresses the parabolic behaviour of the Landau equation which allows for particle solutions; the collision term is well-defined when f is a combination of Dirac measures so particles remain particles. These particle solutions solve a large coupled ODE system that retains all the important properties of the Landau operator, namely the conservation of mass, momentum and energy, and the decay of $\mathcal{H}_{2,\varepsilon}$. We illustrate our new method by showing its performance in several test cases including the physically relevant case of the Coulomb interaction. The comparison to the exact solution and the spectral method is strikingly good maintaining 2nd order accuracy. Moreover, an efficient implementation of the method via the treecode is explored. This gives a proof of concept for the practical use of our method when coupled with the classical particle-in-cell (PIC) method for the

Vlasov equation.

We remark here that ψ_ε and $G^{2,\varepsilon}$ as defined in Section 2.1.1 are both Maxwellian approximations of the identity, albeit with slightly different scalings as $\varepsilon \downarrow 0$. Nevertheless, the idea here is complementary to Theorem 2.8 using a Maxwellian distribution instead of an exponential distribution as the mollifier. Although the equivalence of gradient flow solutions and classical H-solutions to the $\mathcal{H}_{s,\varepsilon}$ Landau equation was only proved in Theorem 2.8 for $s = 1$ (exponential and not Maxwellian tails), this work presents numerical evidence that the result also holds for $s = 2$.

Section 4.1 focuses on the theoretical aspects of this regularisation. The particle method is formulated in Section 4.2, similar structural properties are proven here at the level of spatial and temporal discretisations. Finally, Section 4.3 explores the numerical experimentation of the particle method with comparisons to known cases.

4.1 Regularised Landau equation: basic properties and kernel

In this section, we explore some theoretical properties associated to the homogeneous Landau equation with the regularised entropy functional $\mathcal{H}_{2,\varepsilon}$. The nonlinearity of Q_L makes it difficult to directly regularise f in a structure preserving way. Instead, the regularisation is introduced at the level of the entropy functional which then modifies the homogeneous Landau equation. Notice that the regularised entropy is well-defined on L^1_+ and its first variation with respect to constant mass densities f gives

$$\frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f} = \psi_\varepsilon * \log(f * \psi_\varepsilon)\,, \quad \nabla \frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f} = (\nabla \psi_\varepsilon) * \log(f * \psi_\varepsilon)\,, \tag{4.3}$$

after some computations, see [30] for details.

The aim of this section is to show that (4.2) preserves important structural properties as with the original homogeneous Landau equation. To fix ideas, we introduce a preliminary notion of a weak solution which we can refine after proving the standard conservation properties. For $p > 0$, we recall the notations $\langle v \rangle$ and $L^1_p(\mathbb{R}^d)$ defined by

$$L^1_p(\mathbb{R}^d) := \left\{ g \in L^1(\mathbb{R}^d) \,\middle|\, \int_{\mathbb{R}^d} \langle v \rangle^p \, |g(v)| \mathrm{d}v < \infty \right\}, \quad \langle v \rangle^2 := 1 + |v|^2, \quad v \in \mathbb{R}^d.$$

Let us define

$$\kappa(\gamma) = \begin{cases} 4 + \gamma, & -2 \le \gamma \le 0 \\ 6 + \frac{\gamma}{2}, & -4 \le \gamma < -2 \end{cases}.$$

We also abbreviate the collision kernel and the projection matrix with

$$A[z] := |z|^{2+\gamma}\Pi[z].$$

Definition 4.1 (Weak ε-solution)**.** We say that a nonnegative $f \in C([0,T]; L^1_{\kappa(\gamma)}(\mathbb{R}^d))$ is a *weak ε-solution to* (4.2) if for every $\phi \in C_0^\infty((0,T)\times\mathbb{R}^d)$ we have

$$\begin{aligned}&\int_0^T\int_{\mathbb{R}^d}\partial_t\phi f dvdt = \\ &\frac{1}{2}\int_0^T\iint_{\mathbb{R}^{2d}}(\nabla\phi-\nabla_*\phi_*)\cdot A[v-v_*]\left(\nabla\frac{\delta\mathcal{H}_{2,\varepsilon}}{\delta f}-\nabla_*\frac{\delta\mathcal{H}_{2,\varepsilon,*}}{\delta f_*}\right)ff_*dvdv_*dt.\end{aligned} \tag{4.4}$$

Let us discuss the meaning of the weighted L^1_κ requirement on f. We claim this is sufficient to make sense of the triple integral in (4.4). Here, we are mainly concerned with the soft potentials given by $-4\le\gamma\le 0$. In particular, since $\kappa\ge 2$ we have

$$\sup_{t\in[0,T]}\int_{\mathbb{R}^d}\langle v\rangle^2 f(t,v)\,\mathrm{d}v<\infty,$$

which ensures the following bound

$$\sup_{t\in[0,T]}|\log\left[f(t,\cdot)*\psi_\varepsilon\right](v)|\le C_\varepsilon\langle v\rangle^2, \tag{4.5}$$

where $C_\varepsilon>0$ is a uniform constant depending only on $\varepsilon>0$. The estimate (4.5) is obtained by computations similar to Lemma 2.24. Now let us investigate $B^\varepsilon_{v,v_*} := \nabla\frac{\delta\mathcal{H}_{2,\varepsilon}}{\delta f}-\nabla_*\frac{\delta\mathcal{H}_{2,\varepsilon,*}}{\delta f_*}$. By (4.3), this has the form

$$B^\varepsilon_{v,v_*}=C_\varepsilon\int_{\mathbb{R}^d}\big((v-v')\psi_\varepsilon(v-v')-(v_*-v')\psi_\varepsilon(v_*-v')\big)\log(f*\psi_\varepsilon)(v')dv'.$$

Applying estimate (4.5) gives

$$|B^\varepsilon_{v,v_*}|\le C_\varepsilon\int_{\mathbb{R}^d}\big|(v-v')\psi_\varepsilon(v-v')-(v_*-v')\psi_\varepsilon(v_*-v')\big|\left\langle v'\right\rangle^2dv'.$$

Consider first the easier moderately soft potential case $\gamma\ge-2$. For every test function, we have the bound

$$\begin{aligned}&\left|(\nabla\phi-\nabla_*\phi_*)\cdot A[v-v_*]\left(\nabla\frac{\delta\mathcal{H}_{2,\varepsilon}}{\delta f}-\nabla_*\frac{\delta\mathcal{H}_{2,\varepsilon,*}}{\delta f_*}\right)\right|\\ &\quad\le C_{\varepsilon,\phi,d,\gamma}(|v|^{2+\gamma}+|v_*|^{2+\gamma})\int_{\mathbb{R}^d}(|(v-v')\psi_\varepsilon(v-v')|+|(v_*-v')\psi_\varepsilon(v_*-v')|)(1+|v'|^2)dv'.\end{aligned}$$

By the change of variables $v' \mapsto v - v'$ and $v' \mapsto v_* - v'$, we have the following estimate

$$\int_{\mathbb{R}^d} (|(v-v')\psi_\varepsilon(v-v')| + |(v_*-v')\psi_\varepsilon(v_*-v')|)(1+|v'|^2)dv' \le C_\varepsilon(1+|v|^2+|v_*|^2).$$

This can be used to estimate the triple integral of (4.4) by

$$\begin{aligned}\int_0^T \iint_{\mathbb{R}^{2d}} &\left|(\nabla\phi - \nabla_*\phi_*)^T A[v-v_*]\left(\nabla\frac{\delta\mathcal{H}_{2,\varepsilon}}{\delta f} - \nabla_*\frac{\delta\mathcal{H}_{2,\varepsilon,*}}{\delta f_*}\right)\right| ff_* dv dv_* dt \\ &\le C_{\varepsilon,\phi,d,\gamma}\int_0^T \iint_{\mathbb{R}^{2d}} (|v|^{2+\gamma}+|v_*|^{2+\gamma})(1+|v|^2+|v_*|^2) ff_* dv dv_* dt.\end{aligned}$$

In this case, the $4+\gamma$ weight becomes clear to ensure absolute integrability.

Let us now turn to the very soft potential case $-4 \le \gamma < -2$. The same trick above will not work because the weight $|v-v_*|^{2+\gamma}$ is singular. Instead, split $|v-v_*|^{2+\gamma} = |v-v_*|^{1+\frac{\gamma}{2}}|v-v_*|^{1+\frac{\gamma}{2}}$ so that we have

$$\begin{aligned}&\left|(\nabla\phi - \nabla_*\phi_*)^T A[v-v_*]\left(\nabla\frac{\delta\mathcal{H}_{2,\varepsilon}}{\delta f} - \nabla_*\frac{\delta\mathcal{H}_{2,\varepsilon,*}}{\delta f_*}\right)\right| \\ &\qquad \le C_{\varepsilon,d}\frac{|\nabla\phi - \nabla_*\phi_*|}{|v-v_*|^{-(1+\frac{\gamma}{2})}}\int_{\mathbb{R}^d} |v-v_*|^{1+\frac{\gamma}{2}}\left|(v-v')\psi_\varepsilon(v-v') - (v_*-v')\psi_\varepsilon(v_*-v')\right|(1+|v'|^2)dv'.\end{aligned}$$

Splitting the weight allows us to see that $-(1+\frac{\gamma}{2}) \in (0,1]$ in the very soft potential case so that

$$\frac{|\nabla\phi - \nabla_*\phi_*|}{|v-v_*|^{-(1+\frac{\gamma}{2})}} \le \|\phi\|_{C^{1,-(1+\frac{\gamma}{2})}}.$$

For the remaining $|v-v_*|^{1+\frac{\gamma}{2}}$ term within the integral over v' we use the Mean Value theorem with $|(v-v')\psi_\varepsilon(v-v') - (v_*-v')\psi_\varepsilon(v_*-v')|$ to smother the singularity. Indeed, due to the form of ψ_ε, we have that

$$|v-v_*|^{1+\frac{\gamma}{2}}\left|(v-v')\psi_\varepsilon(v-v') - (v_*-v')\psi_\varepsilon(v_*-v')\right| \le C_\varepsilon|v-v_*|^{2+\frac{\gamma}{2}}(1+|\xi-v'|^2)|\psi_\varepsilon(\xi-v')|,$$

where $\xi \in [v, v_*]$. Substitute this inequality back and use the change of variables $v' \mapsto \xi - v'$ to obtain

$$\begin{aligned}&\left|(\nabla\phi - \nabla_*\phi_*)^T A[v-v_*]\left(\nabla\frac{\delta\mathcal{H}_{2,\varepsilon}}{\delta f} - \nabla_*\frac{\delta\mathcal{H}_{2,\varepsilon,*}}{\delta f_*}\right)\right| \\ &\qquad \le C_{\varepsilon,d,\phi}|v-v_*|^{2+\frac{\gamma}{2}}\int_{\mathbb{R}^d}|v'|^2\psi_\varepsilon(v')(1+|\xi-v'|^2)dv'.\end{aligned}$$

The integral produces a term that has growth bounded (up to a multiplicative constant depending on ε) by $(1+|\xi|^4)$. Since $\xi \in [v, v_*]$, we can estimate $|\xi|^4 \le C(|v|^4+|v_*|^4)$. Inserting this back into the triple integral yields finally

$$\int_0^T \iint_{\mathbb{R}^{2d}} \left| (\nabla\phi - \nabla_*\phi_*)^T A[v-v_*] \left(\nabla \frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f} - \nabla_* \frac{\delta \mathcal{H}_{2,\varepsilon,*}}{\delta f_*} \right) \right| f f_* dv dv_* dt$$
$$\leq C_{\varepsilon,d,\phi} \int_0^T \iint_{\mathbb{R}^{2d}} (1 + |v|^{6+\frac{\gamma}{2}} + |v_*|^{6+\frac{\gamma}{2}}) f f_* dv dv_* dt.$$

Equation (4.4) can be tested against more general functions ϕ. As in [6, Remark 8.1.1], an equivalent expression of (4.4) is

$$\frac{d}{dt}\int_{\mathbb{R}^d} \phi f \, dv = -\frac{1}{2}\iint_{\mathbb{R}^{2d}} (\nabla\phi - \nabla_*\phi_*)^T A[v-v_*] B^{\varepsilon}_{v,v_*} f f_* dv dv_*, \quad \forall \phi \in C_0^\infty(\mathbb{R}^d). \tag{4.6}$$

Furthermore, [6, Lemma 8.1.2] allows us to refine the solution to be weakly continuous $t \in [0,T] \mapsto f(t,\cdot) \in L^1_\kappa(\mathbb{R}^d)$ such that whenever $\phi \in C_0^2((0,T)\times\mathbb{R}^d)$,

$$\begin{aligned} &\int_{t_1}^{t_2}\int_{\mathbb{R}^d} \partial_t \phi f(t,v) dv dt - \frac{1}{2}\int_{t_1}^{t_2}\iint_{\mathbb{R}^{2d}} (\nabla\phi - \nabla_*\phi_*)^T A[v-v_*] B^{\varepsilon}_{v,v_*} f f_* dv dv_* dt \\ &\quad = \int_{\mathbb{R}^d} \phi(t_2,v) f(t_2,v) dv - \int_{\mathbb{R}^d} \phi(t_1,v) f(t_1,v) dv, \quad \forall - \leq t_1 \leq t_2 \leq T. \end{aligned} \tag{4.7}$$

Lemma 4.2. *Let ϕ be an admissible test function and f be a weak ε-solution to* (4.2). *Assume further that*

$$\nabla\phi - \nabla_*\phi_* \in \ker A[v-v_*],$$

then

$$\frac{d}{dt}\int_{\mathbb{R}^d} \phi(v) f(t,v) dv = 0$$

holds, and therefore $\int_{\mathbb{R}^d} \phi(v) f(t,v) \mathrm{d}v$ is a conserved quantity.

Proof. We begin with the formal computations. Differentiating in time, we get

$$\frac{d}{dt}\int_{\mathbb{R}^d} \phi(v) f(t,v)\, \mathrm{d}v = -\frac{1}{2}\iint_{\mathbb{R}^{2d}} (\nabla\phi(v) - \nabla_*\phi(v_*))^T A[v-v_*] B^{\varepsilon}_{v,v_*} f f_* dv dv_* = 0\,.$$

To justify these formal computations, we appeal to approximation arguments by admissible test functions using (4.7) to compare $\int_{\mathbb{R}^d} \phi(v) f(0,v) dv$ with $\int_{\mathbb{R}^d} \phi(v) f(t,v) dv$. A specific cut-off choice can be found in the gradient flow formulation of the Boltzmann equation [61, Lemma 3.8 (version 2)]. □

Since the kernel of the matrix $\Pi[z]$ is spanned by z, a direct consequence of the previous result is that the mass, momentum, and energy of weak ε-solutions are conserved, i.e.,

$$\frac{d}{dt}\left(\int_{\mathbb{R}^d} f(t,v) dv, \int_{\mathbb{R}^d} v f(t,v) dv, \int_{\mathbb{R}^d} |v|^2 f(t,v) dv \right) = 0\,. \tag{4.8}$$

In this way, we define the mass, momentum, and energy of f for all times by the constants ρ, u, T related in the following way

$$\rho = \int_{\mathbb{R}^d} f(t,v)dv, \quad \rho u = \int_{\mathbb{R}^d} v f(t,v)dv, \quad \rho u^2 + \rho d T = \int_{\mathbb{R}^d} |v|^2 f(t,v)dv. \tag{4.9}$$

For general constants ρ, u, T, we denote the Maxwellian satisfying (4.9) by

$$\mathcal{M}_{\rho,u,T}(v) := \frac{\rho}{(2\pi T)^{\frac{d}{2}}} \exp\left\{-\frac{|v-u|^2}{2T}\right\}.$$

As promised, we can refine the notion of weak ε-solution. We add a finite dissipation property which is a mild assumption but yields theoretical and numerical advantages in the spirit of Villani's H-solution [118]. One example of the analytic benefits is in [61, Corollary B.3] where Erbar recovers a strong upper gradient notion for the Boltzmann equation.

Definition 4.3 (Dissipative ε-solution)**.** We say that $f \in C([0,T]; L^1_\kappa(\mathbb{R}^d))$ is a *dissipative ε-solution with moments* $(\rho, u, T) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}_+$ under the relation (4.9) to (4.2) if it is a weak ε-solution in the sense of Definition 4.1 and

1. For every $\phi \in C_0^\infty(\mathbb{R}^d)$ equation (4.6) holds:

$$\frac{d}{dt}\int_{\mathbb{R}^d} \phi f(t,v)dv = -\frac{1}{2}\iint_{\mathbb{R}^{2d}} (\nabla\phi - \nabla_*\phi_*)^T A[v-v_*] B^\varepsilon_{v,v_*} f f_* dv dv_*.$$

2. The dissipation associated to the regularised equation satisfies

$$D_\varepsilon(f(t,\cdot)) := \frac{1}{2}\iint_{\mathbb{R}^{2d}} |v-v_*|^{2+\gamma} \left|\Pi[v-v_*]B^\varepsilon_{v,v_*}\right|^2 f f_* dv dv_* < \infty, \quad \forall t \in [0,T]. \tag{4.10}$$

With the notion of a dissipative ε-solution, we move onto the regularised H-theorem. We will denote by $\frac{d\mathcal{H}_{2,\varepsilon}}{dt}$ for $\frac{d\mathcal{H}_{2,\varepsilon}[f(t,\cdot)]}{dt}$ and similar derivatives in time along solutions to (4.2) from now on.

Lemma 4.4. *Let f be a dissipative ε-solution of* (4.2)*, then we have:*

$$\frac{d\mathcal{H}_{2,\varepsilon}}{dt} = -\frac{1}{2}\iint_{\mathbb{R}^{2d}} |v-v_*|^{2+\gamma} \left|\Pi[v-v_*]B^\varepsilon_{v,v_*}\right|^2 f f_* dv dv_* \le 0\,, \tag{4.11}$$

for all times $t > 0$.

Proof. Beginning with the formal computations, we appeal to (4.6) with $\phi = \psi_\varepsilon * \log(f * \psi_\varepsilon)$ as a test function for every fixed $t \in [0, T]$. By swapping convolutions and using the symmetry of ψ_ε we have

$$\begin{aligned}\frac{dE_\varepsilon}{dt} &= \frac{d}{dt}\left(\int_{\mathbb{R}^d} (f(t,\cdot) * \psi_\varepsilon) \log(f * \psi_\varepsilon) dv\right) = \frac{d}{dt}\left(\int_{\mathbb{R}^d} f(t,\cdot)\psi_\varepsilon * \log(f * \psi_\varepsilon) dv\right) \\ &= -\frac{1}{2}\iint_{\mathbb{R}^{2d}} B^\varepsilon_{v,v_*} \cdot A[v - v_*] B^\varepsilon_{v,v_*} f f_* dv dv_* \leq 0.\end{aligned}$$

Recall that since Π is a projection, the quadratic form satisfies $B^\varepsilon_{v,v_*} \cdot \Pi[v-v_*] B^\varepsilon_{v,v_*} = \left|\Pi[v - v_*] B^\varepsilon_{v,v_*}\right|^2$. To justify using $\psi_\varepsilon * \log(f * \psi_\varepsilon)$ as a test function, we note that the right-hand side of the weak $\mathcal{H}_{2,\varepsilon}$ Landau equation (4.6) makes sense due to the finite dissipation assumption (4.10). It remains to make sense of the left-hand side. Following the estimate (4.5), ϕ as chosen above has at most quadratic growth. Smooth cut-off arguments similar to Lemma 4.2 complete the validity of ϕ chosen above as a test function so the computations are justified. □

In the rest of this section, we follow the strategy of [74, Theorem 4] and [117, Lemma 3] to deduce that stationary states of the homogeneous $\mathcal{H}_{2,\varepsilon}$ Landau equation can be characterised by Maxwellians. Since we are working with weak solutions, let us be specific and define what we mean by stationary states.

Definition 4.5 (Stationary states)**.** We say that a dissipative ε-solution f is a *stationary state* to the homogeneous $\mathcal{H}_{2,\varepsilon}$ Landau equation if for every test function $\phi \in C_0^\infty(\mathbb{R}^d)$,

$$\iint_{\mathbb{R}^{2d}} (\nabla\phi - \nabla_*\phi_*) \cdot A[v - v_*] B^\varepsilon_{v,v_*} f f_* dv dv_* = 0\,, \quad \forall t \in [0, T].$$

We can use this definition with Lemma 4.4 to characterise the first variation of the entropy for a stationary state.

Lemma 4.6. *If f is a stationary solution of* (4.2) *or equivalently f is in the kernel of* (L_ε)*, then the first variation of $\mathcal{H}_{2,\varepsilon}$ is a quadratic polynomial in v, that is*

$$\frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f} = \lambda^{(0)} + \lambda^{(1)} \cdot v + \frac{\lambda^{(2)}}{2}|v|^2\,. \tag{4.12}$$

The constants λ (depending on ε) can be determined by the conserved quantities (4.8) *(see later in Lemma 4.7).*

Proof. This proof adopts the strategy of [74, Theorem 4]. Lemma 4.4 implies that the entropy dissipation, the right-hand side of (4.11), is zero. Moreover, the entropy dissipation is zero if and only if the quadratic form in the integrand of the right-hand side of (4.11) is zero. By definition of $\Pi[v - v_*]$, we must have that B^{ε}_{v,v_*} belongs to the kernel of $\Pi[v - v_*]$ which is characterised by those vectors which are linearly dependent with $v - v_*$. Thus, there exists $\lambda^{(2)} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ with the property

$$\nabla \frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f}(v) - \nabla_* \frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f}(v_*) = \lambda^{(2)}(v, v_*)(v - v_*) \,. \tag{4.13}$$

At this point, we study $\lambda^{(2)}$ and seek to show that the diagonal mapping is constant. Immediately from (4.13), we notice that $\lambda^{(2)}(v, v_*) = \lambda^{(2)}(v_*, v)$. For any $i, j \in \{1, \dots, d\}$ when looking at the j^{th} coordinate of (4.13) and then differentiating with respect to v_i (valid as the ε-regularisation grants arbitrary smoothness), we have

$$\partial_{v_i} \partial_{v_j} \frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f} = \partial_{v_i} \lambda^{(2)}(v, v_*)(v_j - v_{*j}) + \lambda^{(2)}(v, v_*)\delta_{ij} \,.$$

Set $v = v_*$ in the above equation to deduce

$$\partial_{v_i} \partial_{v_j} \frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f} = \lambda^{(2)}(v, v)\delta_{ij} \,. \tag{4.14}$$

Differentiating (4.14) again with respect to v_k for $k \in \{1, \dots, d\}$ yields

$$\partial_{v_k} \partial_{v_i} \partial_{v_j} \frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f} = \partial_{v_k} \lambda^{(2)}(v, v)\delta_{ij} \,.$$

The partial derivatives on the left hand side of the above may be freely permuted with no change to the expression. More interesting is the permutation of the associated indices i, j, and k on the right hand side. One instance of this is the following identity $\partial_{v_k} \lambda^{(2)}(v, v)\delta_{ij} = \partial_{v_i} \lambda^{(2)}(v, v)\delta_{kj}$. For arbitrary $k \in \{1, \dots, d\}$, simply take $i = j \in \{1, \dots, d\} \setminus \{k\}$ and one sees from the above that $\partial_{v_k} \lambda^{(2)}(v, v) = 0$. Since $k \in \{1, \dots, d\}$ was arbitrary, this implies that $\lambda^{(2)}(v, v)$ is actually a constant which we shall refer to as itself (dropping the dependence on v). Equipped with this information, integrating (4.14) twice confirms the claim of the lemma that the first variation of the entropy is a quadratic polynomial given by (4.12). □

Our next step is to show that if f satisfies equation (4.12) then it is a Maxwellian with explicitly computable mass, momentum, and energy.

Lemma 4.7. *If $f \in L^1_+(\mathbb{R}^d) \setminus \{0\}$ satisfies the following equation*

$$\frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta f} = \lambda^{(0)} + \lambda^{(1)} \cdot v + \frac{\lambda^{(2)}}{2}|v|^2 \,,$$

then it must be a Maxwellian. We can deduce a restriction on $\lambda^{(2)}$, specifically, that $\varepsilon|\lambda^{(2)}| < 1$. Fur-

thermore, the mass, momentum, and energy explicitly depend on $\varepsilon, \lambda^{(0)}, \lambda^{(1)}$*, and* $\lambda^{(2)}$ *in the following way:*

$$\begin{cases} \rho &= \left(\frac{2\pi}{|\lambda^{(2)}|}\right)^{\frac{d}{2}} \exp\left\{\lambda^{(0)} + \frac{\varepsilon|\lambda^{(2)}|d}{2} - \frac{\varepsilon|\lambda^{(1)}|^2}{2(1-\varepsilon|\lambda^{(2)}|)} + \frac{|\lambda^{(1)}|^2}{2|\lambda^{(2)}|(1-\varepsilon|\lambda^{(2)}|)}\right\} \\ u &= \frac{\lambda^{(1)}}{|\lambda^{(2)}|} \\ T &= \frac{1}{|\lambda^{(2)}|} - \varepsilon \end{cases} . \tag{4.15}$$

Proof. We iteratively Fourier transform equation (4.12) recalling in particular the convolution and inversion theorems to deduce the identities

$$\psi_\varepsilon * \log(f * \psi_\varepsilon) = \lambda^{(0)} + \lambda^{(1)} \cdot v + \frac{\lambda^{(2)}}{2}|v|^2 ,$$
$$\widehat{\log(f * \psi_\varepsilon)} = (2\pi\varepsilon)^{\frac{d}{2}} \frac{1}{\psi_{\frac{1}{\varepsilon}}} \left(\lambda^{(0)}\delta_0 + i\lambda^{(1)} \cdot \nabla\delta_0 - \frac{\lambda^{(2)}}{2}\Delta\delta_0\right) ,$$

and then

$$\log(f * \psi_\varepsilon) = \lambda^{(0)} - \frac{\lambda^{(2)}\varepsilon d}{2} + \lambda^{(1)} \cdot v + \frac{\lambda^{(2)}}{2}|v|^2 .$$

At this point, we remark that the sign of $\lambda^{(2)}$ can be deduced. Specifically, we claim that $\lambda^{(2)} < 0$. Firstly, $\lambda^{(2)} \leq 0$ because the Dominated Convergence Theorem yields $\lim_{|v|\to\infty} f*\psi_\varepsilon(v) = 0$. If $\lambda^{(2)} = 0$, we must have that $\lambda^{(1)} \cdot v \to -\infty$ whenever $|v| \to \infty$ by the same limiting behaviour of $f * \psi_\varepsilon$. However, this must hold for every direction v such that $|v| \to \infty$ which implies that $\lambda^{(1)}$ depends on v and this is a contradiction. Taking exponentials on both sides of the previous equation, we have

$$f * \psi_\varepsilon(v) = \exp\left\{\lambda^{(0)} - \frac{|\lambda^{(1)}|^2}{2\lambda^{(2)}} - \frac{\lambda^{(2)}\varepsilon d}{2}\right\} \exp\left\{\frac{\lambda^{(2)}}{2}\left|v + \frac{\lambda^{(1)}}{\lambda^{(2)}}\right|^2\right\} ,$$

and one more Fourier transform leads to

$$\hat{f}(\xi) = (2\pi)^d(1 - \varepsilon|\lambda^{(2)}|)^{-\frac{d}{2}} \exp\left\{\lambda^{(0)} + \frac{|\lambda^{(2)}|\varepsilon d}{2} - \frac{\varepsilon|\lambda^{(1)}|^2}{2(1-\varepsilon|\lambda^{(2)}|)}\right\} \mathcal{M}_{\left(1,-\frac{i\lambda^{(1)}}{1-\varepsilon|\lambda^{(2)}|},\frac{|\lambda^{(2)}|}{1-\varepsilon|\lambda^{(2)}|}\right)}(\xi) .$$

Here, we are using the convention that, for vectors $x, y \in \mathbb{R}^d$, $|x + iy|^2 := |x|^2 + 2ix \cdot y - |y|^2$. By the Riemann-Lebesgue lemma, we know that $|\hat{f}(\xi)| \to 0$ as $|\xi| \to \infty$. With the expression for $\hat{f}$ above, this means that the variance of $\mathcal{M}$ (the third parameter in the subscript) must be strictly positive. Hence, $1 - \varepsilon|\lambda^{(2)}| > 0$. Note that the coefficient of $(1 - \varepsilon|\lambda^{(2)}|)^{-\frac{d}{2}}$ is removable due to the cancellation with the same term coming from the variance of the Maxwellian. The Fourier transform step is justified because we previously showed that $\lambda^{(2)} < 0$ so $f * \psi_\varepsilon$ is a Maxwellian and hence belongs to L^1. One final Fourier inversion gives an expression for f as

$$f(v) = \left(\frac{2\pi}{|\lambda^{(2)}|}\right)^{\frac{d}{2}} \exp\left\{\lambda^{(0)} + \frac{\varepsilon|\lambda^{(2)}|d}{2} - \frac{\varepsilon|\lambda^{(1)}|^2}{2(1-\varepsilon|\lambda^{(2)}|)} + \frac{|\lambda^{(1)}|^2}{2|\lambda^{(2)}|(1-\varepsilon|\lambda^{(2)}|)}\right\} \mathcal{M}_{\left(1,\frac{\lambda^{(1)}}{|\lambda^{(2)}|},\frac{1}{|\lambda^{(2)}|}-\varepsilon\right)}(v)\,.$$

Reading off the constants, one confirms (4.15). Note that in the determination of ρ, u, T in equation (4.15), we have a one-to-one correspondence between (ρ, u, T) and $(\lambda^{(0)}, \lambda^{(1)}, \lambda^{(2)})$. Indeed, $\lambda^{(2)}$ is determined from T which then gives $\lambda^{(1)}$ in the equation for u. Finally, $\lambda^{(0)}$ is determined from the equation for ρ. □

The previous lemmas give the following equivalence.

Theorem 4.8. *f is a stationary solution of* (4.2) *if and only if f is a Maxwellian with parameters given by* (4.15) *depending on the quadratic polynomial in equation* (4.12).

Proof. ($\Longrightarrow$) This direction combines Lemmas 4.6 and 4.7.
($\Longleftarrow$) This direction is a computation of $\frac{\delta\mathcal{H}_{2,\varepsilon}}{\delta f} = \psi_\varepsilon * \log(f * \psi_\varepsilon)$ when f is a Maxwellian. □

Remark 4.1. *An alternative regularisation for the entropy is*

$$\tilde{\mathcal{H}}_{2,\varepsilon}(f) = \int_{\mathbb{R}^d} f \log(f * \psi_\varepsilon)\,\mathrm{d}v\,. \tag{4.16}$$

Lemma 4.6 characterising the variation of the entropy for stationary states is still valid, being the variation, see [30], given by

$$\frac{\delta\tilde{\mathcal{H}}_{2,\varepsilon}}{\delta f} = \log(f*\psi_\varepsilon) + \left(\frac{f}{f*\psi_\varepsilon}\right) * \psi_\varepsilon\,, \quad \nabla\frac{\delta\tilde{\mathcal{H}}_{2,\varepsilon}}{\delta f} = \frac{f*\nabla\psi_\varepsilon}{f*\psi_\varepsilon} + \left(\frac{f}{f*\psi_\varepsilon}\right) * \nabla\psi_\varepsilon\,. \tag{4.17}$$

However, the characterisation result as a Maxwellian is not true, even if one might expect the existence and uniqueness of a stationary solution being the conserved quantities fixed.

4.2 A particle method for the homogeneous Landau equation

The main idea is analogous to the recent work [30] for aggregation-diffusion equations. In fact, (4.2) can be viewed as a convection in v with velocity field given by the integral term in (L_ε), and thus giving access to a particle formulation. More specifically, denote

$$f^N(t,v) = \sum_{i=1}^N w_i \delta(v - v_i(t))\,, \tag{4.18}$$

with N being the total number of particles, $v_i(t)$ the velocity of particle i, and w_i the weight of particle i. Plugging (4.18) as a distributional solution to (4.2), we obtain that the evolution for the particle velocities $v_i(t)$, $1 \le i \le N$ is given by

$$\begin{aligned}\frac{dv_i(t)}{dt} &= U_\varepsilon(f^N)(t, v_i(t)) = -\sum_j w_j A[v_i - v_j] \left[\nabla \frac{\delta \mathcal{H}_{2,\varepsilon}^N}{\delta f}(v_i) - \nabla \frac{\delta \mathcal{H}_{2,\varepsilon}^N}{\delta f}(v_j)\right] \\ &= -\sum_j w_j A[v_i - v_j] \left\{ \int_{\mathbb{R}^d} \nabla \psi_\varepsilon(v_i - v) \log\left(\sum_k w_k \psi_\varepsilon(v - v_k)\right) dv \right. \\ &\qquad \left. - \int_{\mathbb{R}^d} \nabla \psi_\varepsilon(v_j - v) \log\left(\sum_k w_k \psi_\varepsilon(v - v_k)\right) dv \right\},\end{aligned} \tag{4.19}$$

with $\frac{\delta \mathcal{H}_{2,\varepsilon}^N}{\delta f} := \psi_\varepsilon * \log\bigl(f^N * \psi_\varepsilon\bigr)$ (the discrete entropy $\mathcal{H}_{2,\varepsilon}^N$ to be defined in the sequel) and therefore,

$$\nabla \frac{\delta \mathcal{H}_{2,\varepsilon}^N}{\delta f}(v_i) = \int_{\mathbb{R}^d} \nabla \psi_\varepsilon(v_i - v) \log\left(\sum_k w_k \psi_\varepsilon(v - v_k)\right) dv. \tag{4.20}$$

Let us show next that the semidiscrete particle method defined by (4.19) leads to a numerical particle approximation f^N of the solution to (L_ε) conserving mass, momentum, energy, and enjoying the regularised entropy dissipation (4.11).

Theorem 4.9. *The semidiscrete particle method* (4.19) *satisfies the following properties:*

1) *Conservation of mass, momentum, and energy:* $\frac{d}{dt}\sum_{i=1}^N w_i \phi(v_i) = 0$ *for* $\phi(v_i) = 1, v_i, |v_i|^2$.

2) *Dissipation of entropy: let*

$$\mathcal{H}_{2,\varepsilon}^N = \mathcal{H}_{2,\varepsilon}^N[f^N] = \int_{\mathbb{R}^d} (f^N * \psi_\varepsilon) \log\bigl(f^N * \psi_\varepsilon\bigr) \mathrm{d}v \tag{4.21}$$

be the discrete entropy, then $\frac{d}{dt}\mathcal{H}_{2,\varepsilon} = -D_\varepsilon^N \le 0$, *where*

$$D_\varepsilon^N = \frac{1}{2}\sum_{i,j} w_i w_j \left(\nabla \frac{\delta \mathcal{H}_{2,\varepsilon}^N}{\delta f}(v_i) - \nabla \frac{\delta \mathcal{H}_{2,\varepsilon}^N}{\delta f}(v_j)\right)^T A[v_i - v_j] \left(\nabla \frac{\delta \mathcal{H}_{2,\varepsilon}^N}{\delta f}(v_i) - \nabla \frac{\delta \mathcal{H}_{2,\varepsilon}^N}{\delta f}(v_j)\right).$$

Proof. First, we notice from (4.19) that

$$\begin{aligned}
\frac{d}{dt}\sum_i w_i\phi(v_i) &= \sum_i w_i \nabla\phi(v_i)\cdot U_\varepsilon(f^N)(t,v_i(t)) \\
&= -\sum_{i,j} w_i w_j \nabla\phi(v_i)^T A[v_i-v_j]\left(\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_i)-\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_j)\right) \\
&= -\frac{1}{2}\sum_{i,j} w_i w_j (\nabla\phi(v_i)-\nabla\phi(v_j))^T A[v_i-v_j]\left(\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_i)-\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_j)\right)
\end{aligned}$$

which vanishes with $\phi(v)=1,v,|v|^2$. Therefore, mass, momentum, and energy are preserved. Next, using (4.18), we rewrite (4.21) as

$$\mathcal{H}^N_{2,\varepsilon}=\int_{\mathbb{R}^d}\left(\sum_i w_i\psi_\varepsilon(v-v_i(t))\right)\log\left(\sum_k w_k\psi_\varepsilon(v-v_k(t))\right)dv,$$

then

$$\begin{aligned}
\frac{d}{dt}\mathcal{H}^N_{2,\varepsilon} &= \int_{\mathbb{R}^d}\sum_i w_i\nabla\psi_\varepsilon(v-v_i(t))\frac{dv_i(t)}{dt}\log\left(\sum_k w_k\psi_\varepsilon(v-v_k(t))\right)dv \\
&\quad+\int_{\mathbb{R}^d}\left(\sum_i w_i\psi_\varepsilon(v-v_i(t))\right)\frac{\sum_k w_k\nabla\psi_\varepsilon(v-v_k(t))\frac{dv_k(t)}{dt}}{\sum_k w_k\psi_\varepsilon(v-v_k(t))}dv \\
&=: I_1+I_2\,.
\end{aligned}$$

Note that I_2 can be simplified to

$$I_2=\int_{\mathbb{R}^d}\sum_k w_k\nabla\psi_\varepsilon(v-v_k(t))\frac{dv_k(t)}{dt}dv=-\frac{d}{dt}\sum_k w_k\int_{\mathbb{R}^d}\psi_\varepsilon(v-v_k(t))dv=0\,,$$

thanks to the fact that $\int_{\mathbb{R}^d}\psi_\varepsilon(v-v_k(t))dv=1$. By virtue of (4.20), I_1 has the following estimate

$$I_1=\sum_i w_i\left(\int_{\mathbb{R}^d}\nabla\psi_\varepsilon(v-v_i(t))\log\left(\sum_k w_k\psi_\varepsilon(v-v_k(t))\right)dv\right)\frac{dv_i}{dt}=\sum_i w_i\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_i)\frac{dv_i}{dt}\,.$$

Then using (4.19), it becomes

$$\begin{aligned}
I_1 &= \sum_i w_i\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_i)\cdot\left[-\sum_j w_j A[v_i-v_j]\left(\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_i)-\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_j)\right)\right] \\
&= -\frac{1}{2}\sum_{i,j} w_i w_j\left(\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_i)-\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_j)\right)^T A[v_i-v_j]\left(\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_i)-\nabla\frac{\delta\mathcal{H}^N_{2,\varepsilon}}{\delta f}(v_j)\right) \\
&\le 0,
\end{aligned}$$

and therefore, the entropy dissipation follows. □

In practical implementation of particle methods, the update of particle velocity via (4.19) is not be computed exactly, but with the integral replaced by quadrature rule. Therefore, we need to introduce a discrete in velocity particle method. The computational domain in any dimension is the square domain $[-L, L]^d$ with $L > 0$. The mesh size is defined by $h = 2L/n$ and $N = n^d$ is the total chosen number of particles. Let us denote the squares of the mesh as Q_i with $i = 1, \dots, n^d$. We will always initialise our particle method by projecting the mass of the initial data on the computational domain to a sum of Dirac deltas located at the centre of each Q_i with mass given by the mass of the initial data in Q_i, that is

$$\bar{f}^N(0, v) := \sum_{i=1}^{N} w_i \delta(v - \bar{v}_i(0)), \quad \text{with } \bar{v}_i(0) = v_i^c \text{ and } w_i = f_0(v_i^c) h^d,$$

with v_i^c denoting the center of the square Q_i. Now, we introduce the discrete in velocity particle method as

$$\bar{f}^N = \sum_{i=1}^{N} w_i \delta(v - \bar{v}_i(t))$$

where $\bar{v}_i(t)$ satisfies

$$\begin{aligned} \frac{d\bar{v}_i(t)}{dt} &= -\sum_j w_j A[\bar{v}_i - \bar{v}_j] \left\{ \sum_l h^d \nabla \psi_\varepsilon(\bar{v}_i - v_l^c) \log\left(\sum_k w_k \psi_\varepsilon(v_l^c - \bar{v}_k) \right) \right. \\ &\qquad \left. - \sum_l h^d \nabla \psi_\varepsilon(\bar{v}_j - v_l^c) \log\left(\sum_k w_k \psi_\varepsilon(v_l^c - \bar{v}_k) \right) \right\} \\ &=: -\sum_j w_j A[\bar{v}_i - \bar{v}_j] \left[\bar{F}_\varepsilon^N(\bar{v}_i) - \bar{F}_\varepsilon^N(\bar{v}_j) \right] =: \bar{U}_\varepsilon(\bar{f}^N)(t, \bar{v}_i(t)). \end{aligned} \tag{4.22}$$

Here, the function

$$\bar{F}_\varepsilon^N(\bar{v}_i) = \sum_l h^d w_l \nabla \psi_\varepsilon(\bar{v}_i - v_l^c) \log\left(\sum_k w_k \psi_\varepsilon(v_l^c - \bar{v}_k) \right) \tag{4.23}$$

is a discrete analogue of the $\nabla \frac{\delta \mathcal{H}_{2,\varepsilon}^N}{\delta f}$. One can also define the fully discrete regularised entropy as

$$\bar{\mathcal{H}}_{2,\varepsilon}^N = \sum_l h^d \left(\sum_i w_i \psi_\varepsilon(v_l^c - \bar{v}_i) \right) \log\left(\sum_k w_k \psi_\varepsilon(v_l^c - \bar{v}_k) \right). \tag{4.24}$$

Then we show that at this fully discrete level, some properties in Theorem 4.9 are inherited.

Theorem 4.10. *The discrete in velocity particle method* (4.22) *satisfies the conservation of mass, momentum, and energy. Moreover, the discrete entropy* (4.24) *almost decays in time, that is,*

$$\bar{\mathcal{H}}_{2,\varepsilon}^N(t) - \bar{\mathcal{H}}_{2,\varepsilon}^N(0) = -\int_0^t \bar{D}_\varepsilon^N dt + \mathcal{O}(h^2),$$

where

$$\bar{D}_\varepsilon^N = \frac{1}{2}\sum_{i,j} w_i w_j \left(\bar{F}_\varepsilon^N(v_i) - \bar{F}_\varepsilon^N(v_j)\right)^T A[v_i - v_j]\left(\bar{F}_\varepsilon^N(v_i) - \bar{F}_\varepsilon^N(v_j)\right) \geq 0\,.$$

Proof. Indeed, for $\phi(v) = 1, v, |v|^2$, we have

$$\begin{aligned}\frac{d}{dt}\sum_i w_i\phi(\bar{v}_i) &= \sum_i w_i \nabla\phi(\bar{v}_i)\cdot \bar{U}_\varepsilon(f^N)(t,\bar{v}_i(t))\\ &= -\sum_{i,j} w_i w_j \nabla\phi(\bar{v}_i)^T A[\bar{v}_i - \bar{v}_j]\left(\bar{F}_\varepsilon^N(\bar{v}_i) - \bar{F}_\varepsilon^N(\bar{v}_j)\right)\\ &= -\frac{1}{2}\sum_{i,j} w_i w_j (\nabla\phi(\bar{v}_i) - \nabla\phi(\bar{v}_j))^T A[\bar{v}_i - \bar{v}_j]\left(\bar{F}_\varepsilon^N(\bar{v}_i) - \bar{F}_\varepsilon^N(\bar{v}_j)\right)\\ &= 0,\end{aligned} \tag{4.25}$$

hence the conversation of mass, momentum, and energy are guaranteed. A similar computation to the entropy dissipation in the semidiscrete level leads to

$$\begin{aligned}\frac{d}{dt}\bar{\mathcal{H}}_{2,\varepsilon}^N(t) =& \sum_l h^d \sum_i w_i \nabla\psi_\varepsilon(v_l^c - \bar{v}_i(t))\frac{d\bar{v}_i(t)}{dt}\log\left(\sum_k w_k\psi_\varepsilon(v_l^c - \bar{v}_k(t))\right)\\ &+ \sum_l h^d\left(\sum_i w_i\psi_\varepsilon(v_l^c - \bar{v}_i(t))\right)\frac{\sum_k w_k\nabla\psi_k(v_l^c - \bar{v}_k(t))\frac{d\bar{v}_k(t)}{dt}}{\sum_k w_k\psi_\varepsilon(v_l^c - \bar{v}_k(t))}\\ =:& I_1 + I_2.\end{aligned}$$

By the definition of (4.23) and similarly to (4.25), I_1 can be written as

$$I_1 = \sum_i w_i\bar{F}_\varepsilon^N(\bar{v}_i)\frac{d\bar{v}_i}{dt} = -\frac{1}{2}\sum_{i,j} w_i w_j |\bar{v}_i - \bar{v}_j|^{2+\gamma}\left|\Pi[\bar{v}_i - \bar{v}_j]\left(\bar{F}_\varepsilon^N(\bar{v}_i) - \bar{F}_\varepsilon^N(\bar{v}_j)\right)\right|^2 \leq 0.$$

As before I_2 can be written as

$$I_2 = \sum_l h^d \sum_i w_i\nabla\psi_\varepsilon(v_l^c - \bar{v}_i(t))\frac{d\bar{v}_i(t)}{dt} = \frac{d}{dt}\sum_i w_i\sum_l h^d\psi_\varepsilon(v_l^c - \bar{v}_i(t))\,.$$

We reduce to showing that

$$\sum_i w_i \sum_l h^d\psi_\varepsilon(v_l^c - \bar{v}_i(t)) = \mathcal{O}(h^2)$$

which is true thanks to the fact that $\int_{\mathbb{R}^d}\psi_\varepsilon(v - v_k(t))\mathrm{d}v = 1$ and that the mid-point composite quadrature rule is of order 2 for smooth functions. Note that the constant in the error depends on ε but not on the location of the particles. □

Remark 4.2. *The particle method for the alternative regularisation for the entropy* (4.16) *has the advantage of not needing a continuous convolution and it also has the conservation and dissipative* *properties. The particle method reads as*

$$\frac{d\tilde{v}_i(t)}{dt} = -\sum_j w_j A[\tilde{v}_i - \tilde{v}_j]\left[\nabla\frac{\delta\tilde{\mathcal{H}}_{2,\varepsilon}^N}{\delta f}(\tilde{v}_i) - \nabla\frac{\delta\tilde{\mathcal{H}}_{2,\varepsilon}^N}{\delta f}(\tilde{v}_j)\right], \tag{4.26}$$

with

$$\nabla\frac{\delta\tilde{\mathcal{H}}_{2,\varepsilon}^N}{\delta f}(v) = \frac{\sum_k w_k \nabla\psi_\varepsilon(v-\tilde{v}_k)}{\sum_k w_k\psi_\varepsilon(v-\tilde{v}_k)} + \sum_k w_k \frac{\nabla\psi_\varepsilon(v-\tilde{v}_k)}{\sum_m w_m\psi_\varepsilon(\tilde{v}_k-\tilde{v}_m)},$$

according to (4.17)*. One can show that the semidiscrete particle method* (4.26) *satisfies the conservation of mass, momentum, energy, and the dissipation of entropy defined as*

$$\tilde{\mathcal{H}}_{2,\varepsilon}^N = \sum_i w_i \log\left(\sum_j w_j\psi_\varepsilon(\tilde{v}_i - \tilde{v}_j)\right),$$

then $\frac{d}{dt}\tilde{\mathcal{H}}_{2,\varepsilon}^N = -\tilde{D}_\varepsilon^N \le 0$*, where*

$$\tilde{D}_\varepsilon^N = \frac{1}{2}\sum_{i,j} w_i w_j |\tilde{v}_i - \tilde{v}_j|^{2+\gamma}\left|\Pi[\tilde{v}_i - \tilde{v}_j]\left(\nabla\frac{\delta\tilde{\mathcal{H}}_{2,\varepsilon}^N}{\delta f}(\tilde{v}_i) - \nabla\frac{\delta\tilde{\mathcal{H}}_{2,\varepsilon}^N}{\delta f}(\tilde{v}_j)\right)\right|^2.$$

This alternative regularisation will be explored elsewhere.

4.3 Numerical implementation and simulation

In order to visualise our particle solution and compare it to the exact solutions in smoother norms, we construct a blob solution, as in [30], obtained by convolving the particle solution with the mollifier,

$$\tilde{f}^N(t,v) := (\psi_\varepsilon * \bar{f}^N)(t,v) = \sum_{i=1}^N w_i\psi_\varepsilon(v - \bar{v}_i(t)), \tag{4.27}$$

with $\bar{v}_i(t)$ given by (4.22) for all $t > 0$. We measure the accuracy of our numerical method with respect to the L^1- and L^∞-norms. To compute the L^1- and L^∞-errors, we take the difference between the exact or reference solution and the blob solution (4.27) and evaluate discrete L^p- and L^∞-norms in a grid. The norms are computed in this computational mesh using the centres of the squares Q_i as

$$\|g\|_{L^p}^p = \sum_{i=1}^N h^d |g(v_i^c)|^p, \quad \|g\|_{L^\infty} = \max_i |g(v_i^c)|,$$

for any function g defined on the computational mesh, and $1 \leq p < \infty$. The quantities of interest are computed as follows: the discrete mass, momentum, and energy are defined as

$$\sum_{i=1}^{N} w_i, \quad \sum_{i=1}^{N} w_i \bar{v}_i \quad \text{and} \quad \sum_{i=1}^{N} w_i |\bar{v}_i|^2 ,$$

respectively. The discrete entropy is defined by $\bar{\mathcal{H}}_{2,\varepsilon}^N$ in (4.24).

Let us now comment on the practical implementation of the method. The time discretisation of the system of ODEs defined by the particle method (4.22) is done by the simple explicit Euler method. This choice is motivated by our main purpose: we want to illustrate the performance of this particle method by focusing on the basic properties and its capabilities even with the lowest order in time discretisation. Note that the fully discrete in time method conserves mass and momentum exactly, but the energy conservation is satisfied up to a first order error in time. We will check these issues later on in the examples. One can obviously improve some of the time discretisation errors committed by choosing higher order time approximations of the ODE system with adaptive time stepping. We leave this for future work in the scientific computing direction focusing here on the convergence analysis and error in velocity of the particle approximation (4.22).

As usual in particle methods, the regularisation parameter has to be chosen very carefully. This regularisation was already used for nonlinear diffusion and aggregation-diffusion equations in [30]. It was proven in [30, Theorem 6.1] that, for the porous medium equation with exponent larger than or equal to 2, a particle method using the regularisation strategy presented in this work is convergent by choosing $h^2 = o(\varepsilon)$ as $\varepsilon \to 0$. By choosing $h^p \simeq \varepsilon$, the previous constraint is satisfied for $0 < p < 2$. Then, it was checked heuristically that with $\varepsilon \simeq h^{1.98}$, the numerical particle scheme is a second order approximation to the solutions of all nonlinear degenerate diffusion equations of porous medium type and also for the heat equation. Notice it is more convenient to choose the largest posible h to have the least number of particles since $h = 2L/n$. For these reasons, the regularising parameter for the Landau equation is chosen as $\varepsilon = 0.64h^{1.98}$. Here the prefactor is empirical and is found by trial and error.

Finally, let us comment that this error estimate is different for transport equations as studied in [40, 43]. For the transport equation, depending on the regularity of the initial data, one gets $h^p \simeq \varepsilon$ for $0 < p < 1$, that is $h = o(\varepsilon)$ meaning that for transport equations one needs typically smaller meshes and therefore more particles than for diffusion-type equations.

4.3.1 Example 1: 2D BKW solution for Maxwell molecules

Consider the collision kernel

$$A[z] = \frac{1}{16}|z|^2 \Pi[z],$$

and an exact solution given by

$$f^{\text{ext}}(t, v) = \frac{1}{2\pi K} \exp\left(-\frac{|v|^2}{2K}\right) \left(\frac{2K-1}{K} + \frac{1-K}{2K^2}|v|^2\right), \quad K = 1 - \exp(-t/8)/2.$$

We choose $t_0 = 0$ and compute the solution until $t = 5$. The number of particles are chosen as $N = n^2$ with $n = 60, 80, 100, 120, 150$. The computational domain is $[-L, L]^2$ with $L = 4$, so the initial mesh size is $h = 2L/n$. The forward Euler with $\Delta t = 0.01$ is used for time discretisation.

We first track the relative L^2 error of the solution, see Figure 4.1 (left), from which we observe the errors remain stable over time and decrease with higher number of particles. To check the decay rate, we generate the loglog plot of the errors at a fixed time $t = 5$, see Figure 4.1 (right). Here the x-axis is h, i.e., the initial mesh size. Using the least square fitting, we can find the approximate slope of the errors which exhibits almost second order convergence.

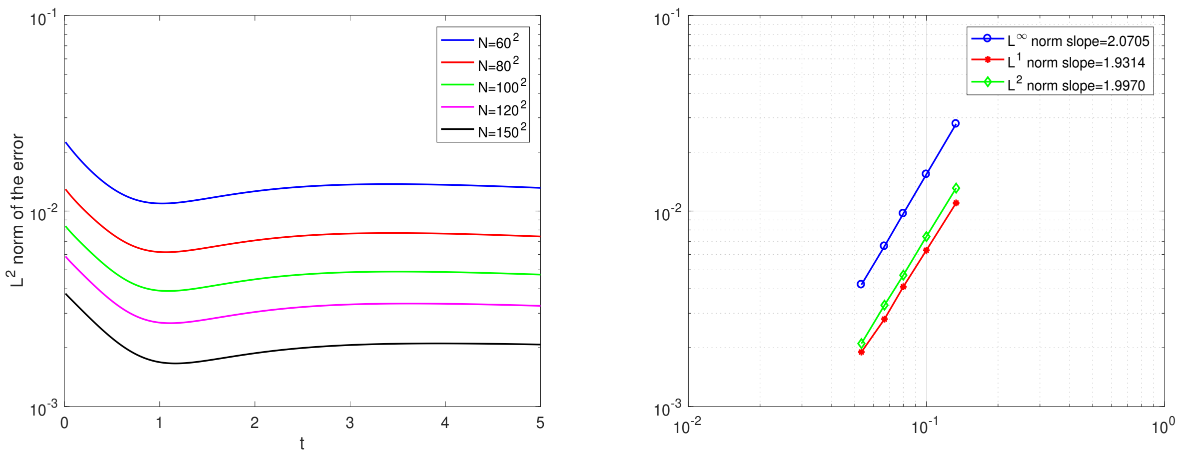


Figure 4.1: Left: Time evolution of $\|f^{\text{num}} - f^{\text{ext}}\|_{L^2}/\|f^{\text{ext}}\|_{L^2}$ with respect to different number of particles. Right: Relative L^∞, L^1, and L^2 norms of the error at time $t = 5$ with respect to different h.

To further check the conservation and entropy decay properties of the method, we plot the time evolution of the total energy and relative entropy of the system in Figure 4.2. The energy is conserved up to a very small error (this error decays when the time step decreases) while the entropy decays monotonically as expected. Analogously to equation (4.24), we define the relative entropy as

$$\sum_l h^d \left(\sum_{k=1}^N w_k \psi_\varepsilon(v_l^c - \bar{v}_k)\right) \left(\log\left(\sum_{k=1}^N w_k \psi_\varepsilon(v_l^c - \bar{v}_k)\right) + \log(2\pi) + \frac{1}{2}|v_l^c|^2\right).$$

4.3.2 Example 2: 3D BKW solution for Maxwell molecules

Consider the collision kernel

$$A[z] = \frac{1}{24}|z|^2 \Pi[z],$$

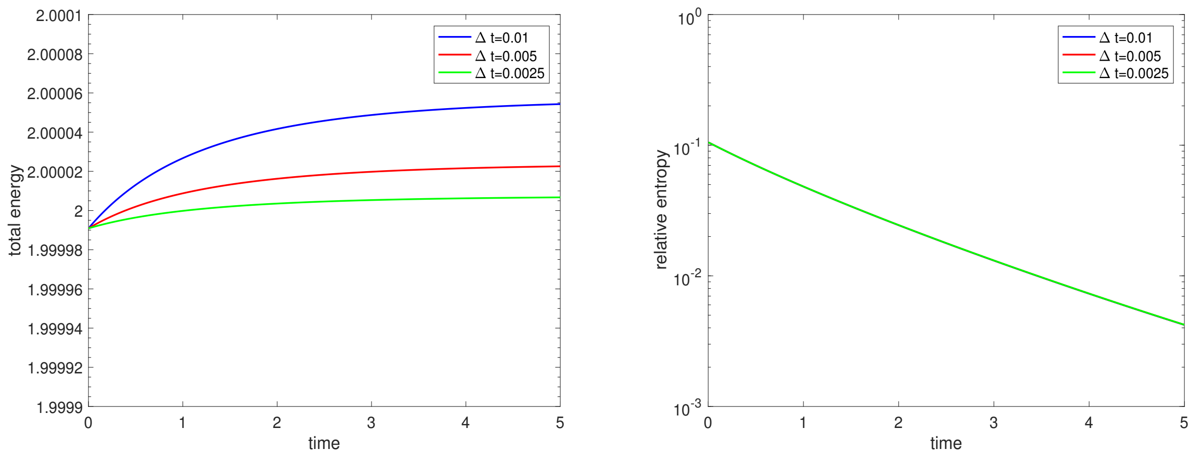


Figure 4.2: Time evolution of the total energy (left) and relative entropy (right) with respect to different time step. Particle number $N = 60^2$ is fixed.

and an exact solution given by

$$f^{\text{ext}}(t,v) = \frac{1}{(2\pi K)^{3/2}} \exp\left(-\frac{|v|^2}{2K}\right)\left(\frac{5K-3}{2K} + \frac{1-K}{2K^2}|v|^2\right), \quad K = 1 - \exp(-t/6).$$

We choose $t_0 = 5.5$ and compute the solution until $t = 6$. The number of particles are chosen as $N = n^3$ with $n = 20, 30, 40, 50, 60$. The computational domain is $[-L, L]^3$ with $L = 4$, so the initial mesh size is $h = 2L/n$. The forward Euler with $\Delta t = 0.01$ is used for time discretisation.

Here we plot similar figures as in the 2D case. We mention that the direct computation in 3D is computationally costly so that we cannot afford too many particles and the errors are generally larger than in 2D. Remarkably, even with a small number of particles, up to 60^3, we are still able to observe the second order convergence in L^1 and L^2 norms (L^∞ norm is not very reliable due to the limited number of particles), see Figure 4.3.

4.3.3 Example 3: 2D anisotropic solution with Coulomb potential

Consider the collision kernel

$$A[z] = \frac{1}{16}|z|^{-1}\Pi[z],$$

and the initial condition

$$f(0,v) = \frac{1}{4\pi}\left\{\exp\left(-\frac{(v-u_1)^2}{2}\right) + \exp\left(-\frac{(v-u_2)^2}{2}\right)\right\}, \quad u_1 = (-2,1), \quad u_2 = (0,-1).$$

For this example, we do not have the exact solution to compare with. Therefore, we compare the particle method with the Fourier spectral method in [104]. For the particle method, we choose the

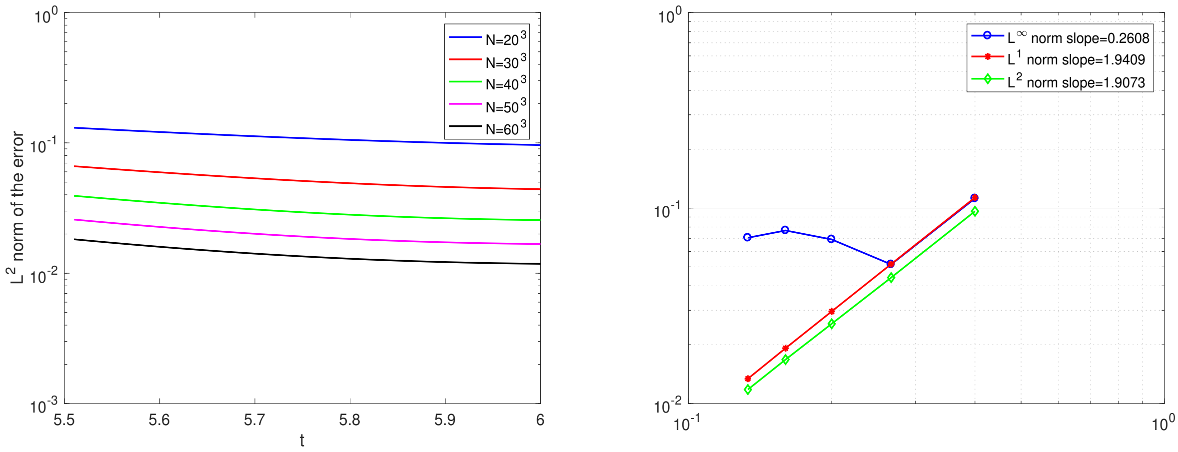


Figure 4.3: Left: Time evolution of $\|f^{\text{num}} - f^{\text{ext}}\|_{L^2}/\|f^{\text{ext}}\|_{L^2}$ with respect to different number of particles. Right: Relative L^∞, L^1, and L^2 norms of the error at time $t = 6.5$ with respect to different h.

following parameters: the number of particles is $N = 120^2$ and the computational domain is $[-10, 10]^2$. The forward Euler with $\Delta t = 0.1$ is used for time discretisation.

For the spectral method, we choose the following parameters: the number of Fourier modes in each velocity dimension is $N_v = 128$; the computational domain is $[-10, 10]^2$. The second order Heun's method with $\Delta t = 0.1$ is used for time discretisation.

The results are shown in Figure 4.4. The results of the two methods match very well.

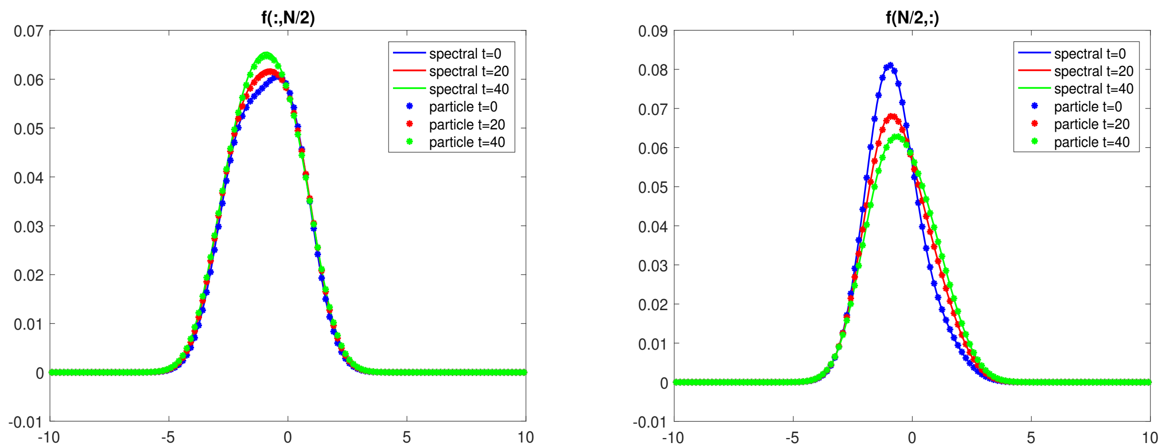


Figure 4.4: Comparison of the particle method (particle number $N = 120^2$) with the spectral method ($N_v = 128^2$). Slices of the solutions at different times.

To better check the convergence of the particle method, we use the spectral method solution with $N_v = 128$ as a reference solution. For the particle method, we test $N = 60^2, 80^2, 100^2, 120^2$, and for each of them reconstruct the solution on the same mesh as the spectral method (so that we can directly compare the error). The results are shown in Figure 4.5 where we can observe better matching as N

increases. We also compute the convergence order similarly to the example in Section 4.3.1. Strikingly, we can still obtain almost second order convergence, see Figure 4.6.

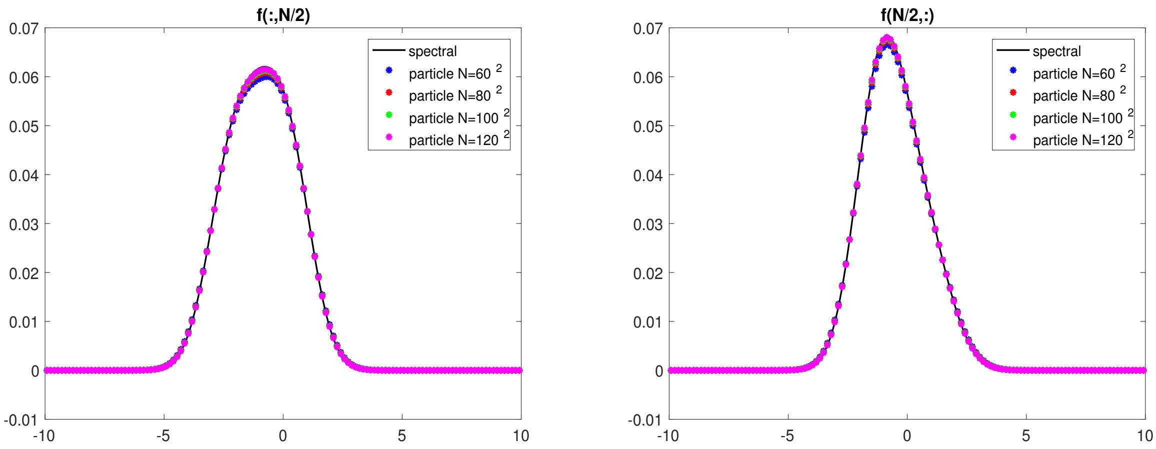


Figure 4.5: Comparison of the particle method (using different particle numbers) with the spectral method ($N_v = 128^2$). Slices of the solutions at time $t = 20$.

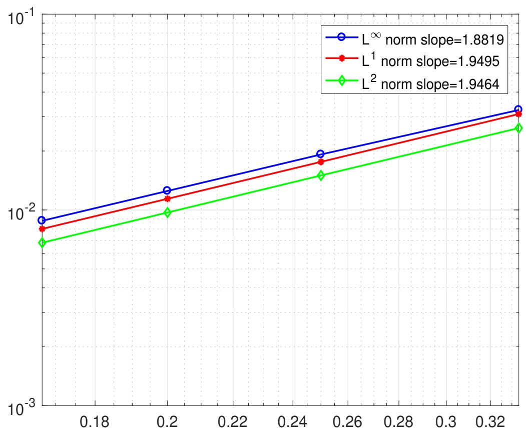


Figure 4.6: Relative L^∞, L^1, and L^2 norms of the error at time $t = 20$ with respect to different h.

4.3.4 Example 4: 3D Rosenbluth problem with Coulomb potential

Consider the collision kernel

$$A[z] = \frac{1}{4\pi}|z|^{-1}\Pi[z],$$

and the initial condition

$$f(0, v) = \frac{1}{S^2}\exp\left(-S\frac{(|v| - \sigma)^2}{\sigma^2}\right), \quad \sigma = 0.3, \quad S = 10.$$

A similar test has been considered in other papers [104]. For the particle method, we choose the following parameters: the number of particles is $N = 50^3$; the computational domain is $[-1,1]^3$. The forward Euler with $\Delta t = 0.2$ is used for time discretisation.

For the spectral method, we choose the following parameters: the number of Fourier modes in each velocity dimension is $N_v = 64$; the computational domain is $[-1,1]^3$. The second order Heun's method with $\Delta t = 0.2$ is used for time discretisation.

The cost of computing the particle method in 3D becomes very heavy if the right-hand side of (4.22) is performed by direct sums. We resort to efficient methods for computing large sums involving convolution kernels. One possible choice is to make use of the treecode strategy as in [14, 94] for instance. Its application to the particle method (4.22) can be found in [33, Appendix B]. In Figure 4.7 left, we show the comparison of the direct sum solver to the treecode solver by plotting their solutions at $t = 20$, $N = 50^3$ or $N = 40^3$. The error committed is negligible. In Figure 4.7 right, we illustrate the speed-up of the treecode solver with respect to the direct sum solver. The efficiency of the treecode solver is significant with larger number of particles N as expected. The results are obtained on Minnesota Supercomputer Institute Mesabi machine with 12 nodes.

The result is shown in Figure 4.8. We observe good agreement between the spectral method and the particle method using the treecode acceleration, especially for short time. For longer time, the discrepancy is due to the limited resolution of the particle method. Note that we do get better convergence when increasing the number of particles from $N = 50^3$ to $N = 60^3$.

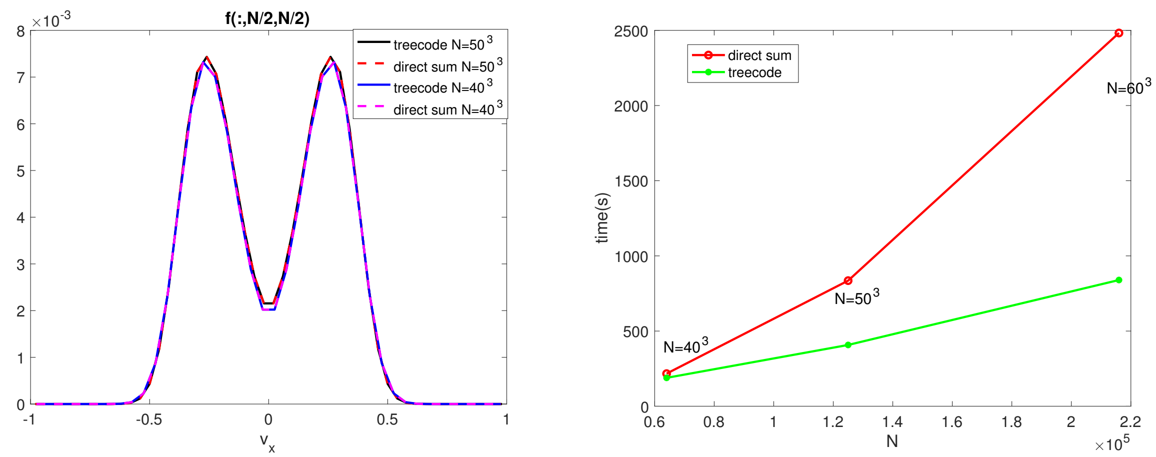


Figure 4.7: Left: comparison a slice of solution with direct sum and treecode at $t = 20$, $N = 50^3$ or $N = 40^3$. Right: comparison of computational time (in seconds) for one step with the treecode solver and with the direct sum solver.

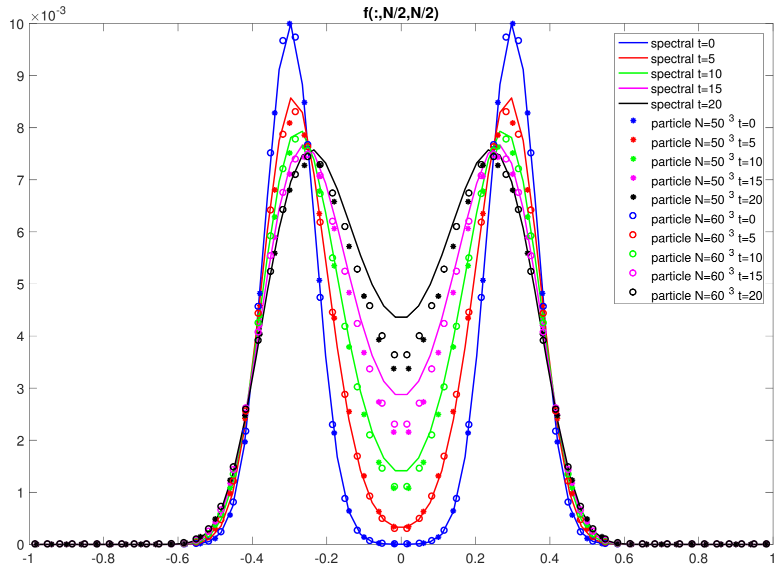


Figure 4.8: Comparison of the particle method using treecode acceleration (using different particle numbers) with the spectral method ($N_v = 64^3$). Slices of the solution at different times.

Chapter 5

Conclusions and Perspectives

With much gratitude, I thank the diligent and persistent reader for making it this far into the thesis. Now, I would like to summarise the results presented here and discuss some of the future avenues of research.

5.1 Gradient Flow structure of the Landau Equation

The fundamental result of this thesis is Theorem 2.10, which gives conditions for when Villani's H-solutions [118] of the Landau equation are equivalent to curves of maximal slope for the entropy functional $\mathcal{H}$ with respect to the Landau metric d_L. There, the only hypothesis which is not already included in the definition of H-solution is the technical assumption (**A2.1**); for $\gamma \in (-3, 0)$, there exists some $0 < \eta \leq \gamma + 3$ such that

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

Our method in Section 2.5 is unable to accommodate $\gamma = -3$ because the requirement for η degenerates in this case. For soft potentials $\gamma \geq -2$, this estimate is propagated globally in time [122, 54] under the appropriate assumptions on the initial data of solutions. Notice that for Maxwellian potentials $\gamma = 0$, this estimate is nothing other than a uniform-in-time bound on the second moment without any L^p requirement for $p > 1$. Hence, gradient flows coincide with H-solutions unconditionally [118, 119]. On the other hand, for very soft potentials $\gamma \in (-3, -2)$, (**A2.1**) is not known to hold for global weak solutions with large initial data. Global solutions with initial data close to equilibrium are known to exist and satisfy (**A2.1**) based on the seminal work of Guo [78]. Conversely, short time solutions for arbitrary initial data are known to exist and satisfy (**A2.1**) (including even the Coulomb case $\gamma = -3$) due to [55, 11].

It would be nice to dispense with the L^p estimate of (**A2.1**) so that at least the equivalence of

H-solutions and gradient flows of (L) is true for global weak solutions with large initial data in the very soft potential case. This could be approached in two different ways. From the gradient flow point of view, one could revisit the proof of Theorem 2.10 in Section 2.5 and investigate sources of improvement. From the classical kinetic theory point of view, one could also try to prove that (**A2.1**) is propagated (missing for $\gamma < -2$) for H-solutions with appropriate initial data.

Moreover, existing literature aside, the technical computations in Section 2.5 gave me an appreciation for the particular difficulty of the Coulomb case $\gamma = -3$. In relation to the previous discussion, I would very much like to extend our equivalence result to this case but, again, this would require technical improvements to the already-very-technical computations of Section 2.5.

While Theorem 2.10 is a result which compares the notions of H-solutions and gradient flows of (L), we have not yet explored global existence for gradient flows of (L) with minimal assumptions on the initial data. One strategy here would be to modify the results in Section 2.4 by replacing the regularised entropy functional $\mathcal{H}_\varepsilon$ with $\mathcal{H}$. More precisely, the natural construction of gradient flows is to use the JKO scheme [90]. Given a fixed time step $\tau > 0$ and initial datum $\mu \in \mathscr{P}_{2,E}$, solve the minimisation problem for $n \in \mathbb{N}$

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

and then concatenate the minimisers into a piecewise constant curve

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

In fact, arguing in the same way as Proposition 2.27, this construction is well-defined (minimisers exist) and up to a subsequence, the piecewise constant curve converges weakly to some limit

$$\mu_t^\tau \overset{\tau \downarrow 0}{\rightharpoonup} \mu_t \in \mathscr{P}_{2,E}, \quad \forall t \in [0, \infty).$$

The problem now is that, while the JKO scheme yields the existence of *some* curve μ_t, it is not clear that μ_t is a curve of maximal slope corresponding to (L). The difficulty here is proving an analogue of Lemmas 2.28 and 2.29 with $\mathcal{H}$ considered in place of $\mathcal{H}_\varepsilon$. In particular, the analogous statement for Lemma 2.29 would assert that $|\partial^- \mathcal{H}|$ is a strong upper gradient for $\mathcal{H}$, but this requires the chain rule Proposition 2.23 at the $\varepsilon = 0$ level. We have already seen, from Section 2.5, the difficulty associated with this result as it ties into the previous discussion for the necessity of Assumption (**A2.1**).

Another option is to start at the $\varepsilon > 0$ level and prove a Γ-convergence type result for the sequence of curves of maximal slope constructed in Theorem 2.9. More precisely, for every $\varepsilon > 0$ and $\gamma \in [-4, 0]$, the curves μ^ε constructed in Theorem 2.9 satisfy the chain rule (Proposition 2.23) and the Energy

Dissipation Inequality (EDI)

$$\mathcal{H}_\varepsilon[\mu^\varepsilon(T)] + \frac{1}{2}\int_0^T D_\varepsilon(\mu^\varepsilon(t))dt + \frac{1}{2}\int_0^T |\dot{\mu}^\varepsilon|_L^2(t)dt \le \mathcal{H}_\varepsilon[\mu^\varepsilon(0)], \quad \forall T > 0. \tag{5.1}$$

To construct a curve of maximal slope for the full Landau equation (L) given initial condition $\mu_0 \in \mathscr{P}_2$, it is natural to investigate sources of compactness for the sequence μ^ε. In particular, assuming compatibility and finite entropy of the initial data

$$\lim_{\varepsilon\downarrow 0} \mathcal{H}_\varepsilon[\mu^\varepsilon(0)] = \mathcal{H}[\mu_0] < +\infty,$$

the EDI (5.1) implies the bounds

$$\sup_{\varepsilon>0} \mathcal{H}_\varepsilon[\mu^\varepsilon(T)] + \sup_{\varepsilon>0}\int_0^T D_\varepsilon(\mu^\varepsilon(t))dt + \sup_{\varepsilon>0}\int_0^T |\dot{\mu}^\varepsilon|_L^2(t)dt < +\infty. \tag{5.2}$$

I would like to understand how these estimates imply compactness of $(\mu^\varepsilon)_\varepsilon$. Supposing there is a convergent subsequence for which $\mu^\varepsilon(t) \rightharpoonup \mu(t)$, the next difficulty would be to show that the limit μ is a curve of maximal slope for $\mathcal{H}$ with respect to d_L. As usual, this involves passing to the limit $\varepsilon \downarrow 0$ for the chain rule

$$\begin{aligned} \mathcal{H}_\varepsilon[\mu^\varepsilon(T)] - \mathcal{H}_\varepsilon[\mu^\varepsilon(0)] &= -\frac{1}{2}\int_0^T \iint \tilde{\nabla} G^\varepsilon * \log[\mu^\varepsilon * G^\varepsilon]\cdot M^\varepsilon \\ \overset{\varepsilon\downarrow 0}{\to} \qquad \mathcal{H}[\mu(T)] - \mathcal{H}[\mu_0] &= -\frac{1}{2}\int_0^T \iint \tilde{\nabla}\log\mu \cdot M, \end{aligned}$$

where $(\mu^\varepsilon, M^\varepsilon)$, $(\mu, M) \in GCE_T$. Here, M^ε is an arbitrary collision rate associated to μ^ε and one would have to prove compactness of M^ε so that it converges to M. Notice also that, unlike the proof of Theorem 2.10, the curves depend on $\varepsilon > 0$ which further complicates any repetition of the strategy in Section 2.5.

The other major extension is the question of uniqueness of gradient flow solutions to (L). To illustrate the promise offered by gradient flows, I would like to return to the finite dimensional example of (1.7); curves $x(t) \in \mathbb{R}^d$ solving $\dot{x}(t) = -\nabla E(x(t))$ for some smooth and convex function $E : \mathbb{R}^d \to \mathbb{R}$. If x^1, $x^2 \in C^1$ are two solutions of (1.7) with respective initial condition x_0^1, $x_0^2 \in \mathbb{R}^d$, then the convexity of E yields

$$\frac{d}{dt}\frac{1}{2}|x^1 - x^2|^2 = -(\nabla E(x^1) - \nabla E(x^2))\cdot(x^1 - x^2) \le 0.$$

If E is λ-convex, then the same computation quantifies not only uniqueness of solution, but also trend to equilibrium for one then obtains

$$|x^1(t) - x^2(t)| \le |x^1(0) - x^2(0)| e^{-\frac{\lambda}{2}t}.$$

Here, we see the importance of quantifying convexity of E which translates similarly to PDEs [101, 35]. As uniqueness of weak solutions for (L) is not known in full generality for $\gamma = -3$, an interesting avenue of research is proving whether $\mathcal{H}$ is convex along generalised geodesics with respect to d_L. This would help close a longstanding open question in the theory of the Landau equation. A first step in this direction, which is interesting in its own right, is to revisit the convexity results of functionals [35] with respect to optimal transport distances from just the *dynamic* formulation [17, 60] of these distances. Another preliminary area of investigation is to consider (L) for *radial* solutions as the isotropy somehow compensates for the degeneracy of Π [73, 72, 75]. My supervisors and I understand how to treat the $\gamma = 0$ case using gradient flow techniques with and without radial symmetry. Concerning the $\gamma < 0$ case even with radial symmetry, we performed the formal computations in this direction which quickly escalated in complexity.

5.2 The Boltzmann and Landau equations

In Chapter 3, we recovered the grazing collision limit between the Boltzmann equation and the Landau equation using the gradient flow framework developed in Chapter 2 (for Landau) and [61] (for Boltzmann). There, the intention was to simplify the grazing collision limit using the Γ-convergence ideas of Sandier and Serfaty [113, 109]. We were able to prove this result using the weaker notion of solutions (H-gradient flows from Definition 3.1) which were based *only* on the Energy Dissipation Inequality (EDI). This is in contrast with the usual notion of curves of maximal slope (Definition 2.5), which are based on the EDI *and* the use of the entropy-dissipation as a strong upper gradient. The gap here lends itself to a direction of future research. There are two independent and interesting problems associated with this endeavour.

One of the obstructions to proving the grazing collision limit for curves of maximal slope from Boltzmann to Landau is the aforementioned technical difficulty of relaxing (**A2.1**) to establish the chain rule for the Landau equation. It is not clear from the compactness and regularity estimates of Chapter 3 how to deduce the integrability requirements of (**A2.1**). Moreover, as we emphasised with the finite second moment and entropy assumption (**A3.1**), the higher order L^p integrability in (**A2.1**) is extraneous for the grazing collision limit [118].

Another reason for considering H-gradient flows is that Erbar's result equating weak solutions and curves of maximal slope for the Boltzmann equation applies only for bounded collision kernels [61]; the Boltzmann analogue of Theorem 2.10 holds only for $\gamma = 0$. It would be nice to extend Erbar's result for the soft potentials $\gamma < 0$. As with us, the major source of technical complexity for Erbar was proving the chain rule. This suggests to me that the strategy in Section 2.5 could be used to reach $\gamma \in (-3, 0)$. One of the key ingredients for us was Desvillettes' Theorem 2.32 which controlled a weighted H^1 norm of $\sqrt{f}$ with the Landau entropy-dissipation. To treat Boltzmann for $\gamma < 0$, my first instinct would be to see if the fractional Sobolev control of $\sqrt{f}$ (c.f. Appendix C and [2, 86, 80]) with respect to the Boltzmann entropy-dissipation can play a similar role here.

5.3 A particle method for the Landau equation

The focus of Chapter 4 was fixed on the numerical approximation of solutions to (L) by studying the regularised problem (L_ε) with right-hand side given by

$$Q_L^\varepsilon(f, f) = \nabla \cdot \left(f \int f_* |v - v_*|^{2+\gamma} \Pi[v - v_*] (\nabla \psi_\varepsilon * \log [\mu * \psi_\varepsilon] - \nabla_* \psi_\varepsilon * \log [\mu_* * \psi_\varepsilon]) dv_* \right).$$

Here, ψ_ε is the Maxwellian mollifer from (ψ_ε), so that (L_ε) is the $\mathcal{H}_{2,\varepsilon}$ Landau equation (in the sense of (2.1)) where

$$\mathcal{H}_{2,\varepsilon}[\mu] = \mathcal{H}[\mu * \psi_\varepsilon], \quad \frac{\delta \mathcal{H}_{2,\varepsilon}}{\delta \mu} = \psi_\varepsilon * \log [\mu * \psi_\varepsilon].$$

This particular regularisation was motivated by the fact that (smooth) solutions, $f \in L^1_+$, to (L_ε) satisfy an H-theorem with the regularised entropy

$$\begin{aligned} \frac{d}{dt} \mathcal{H}_{2,\varepsilon}[f(t)] &= -D_L^\varepsilon(f) \\ &= -\frac{1}{2} \iint f f_* |v - v_*|^{2+\gamma} |\Pi[v - v_*] (\nabla \psi_\varepsilon * \log [f * \psi_\varepsilon] - \nabla_* \psi_\varepsilon * \log [f_* * \psi_\varepsilon])|^2 \leq 0. \end{aligned}$$

In addition, we proved that this regularisation preserves other structural properties of the original Landau equation; the mass, momentum, and energy of solutions are conserved quantities, and equilibrium states are given by Maxwellians. Moreover, while the regularisation destroys the parabolic property of (L), it allows the use of particle methods to approximate (L_ε).

Chapter 4 leaves open the question of using a particle method on (L_ε) to rigorously approximate (L). Although numerical experiments provide heuristic arguments in favour of this approximation, a rigorous justification is missing. The approximation discussed in Chapter 4 has two layers; the first layer is approximating (L) with (L_ε) and the second layer is approximating solutions of (L_ε) with particle solutions. At the time of writing, I am confidently happy to declare that the latter layer has been studied, albeit with the mollifying sequence G^ε, by my Ph.D. supervisors and myself. We are currently preparing a manuscript which more or less states that, for $\gamma \in (-3,0]$ (again, the pesky case of $\gamma = -3$ is not attained yet), solutions of (L_ε) (using G^ε as the mollifier) are well-approximated by particle solutions with convergence as the number of particles tends to infinity. This is also supported by Theorems 2.8 and 2.9, which assert that curves of maximal slope for $\mathcal{H}_\varepsilon$ with respect to d_L exist and are equivalent to weak solutions to the $\mathcal{H}_\varepsilon$ Landau equation.

The numerical experiments in Section 4.3 suggest that convergence of the particle method also holds with ψ_ε as the mollifying sequence. Moreover, I also expect that the statements of Theorems 2.8 and 2.9 are true with $\mathcal{H}_{2,\varepsilon}$ replacing $\mathcal{H}_\varepsilon$. Recall, $\mathcal{H}_{2,\varepsilon}$ uses the mollifier $\psi_\varepsilon(v) \sim \exp\{-|v/\varepsilon|^2\}$ whereas $\mathcal{H}_\varepsilon$ uses $G^\varepsilon(v) \sim \exp\{-\langle v/\varepsilon\rangle\}$. Owing to the different tailed behaviour, the following estimates from Lemma 2.24 read

$$|\log[\mu * G^\varepsilon]| \lesssim_\varepsilon \langle v\rangle, \quad |\log[\mu * \psi_\varepsilon]| \lesssim_\varepsilon \langle v\rangle^2.$$

At the level of the entropy-dissipations induced by either mollifier, we require only $\mu \in \mathscr{P}_2$ with G^ε, whereas $\mu \in \mathscr{P}_4$ seems necessary with ψ_ε. Different assumptions for μ are required when adjusting the tailed behaviour of the mollifying sequence.

Concerning the approximation of (L) by (L_ε), this problem is also present and open for non-linear diffusion models [30, 44, 24]. In fact, this is closely related to the discussion at the end of Section 5.1. The previous paragraph shows that the choice of mollifying sequence adds another element to this problem, although the structure is very general; given ϕ^ε a sequence of (positive) mollifiers and setting

$$\mathcal{H}_{\phi^\varepsilon}[\mu] := \mathcal{H}[\mu * \phi^\varepsilon] = \int (\mu * \phi^\varepsilon) \log(\mu * \phi^\varepsilon),$$

the first variation is always of the form

$$\frac{\delta \mathcal{H}_{\phi^\varepsilon}}{\delta \mu} = \phi^\varepsilon * \left[\frac{\delta \mathcal{H}}{\delta(\mu * \phi^\varepsilon)}\right] = \phi^\varepsilon * \log(\mu * \phi^\varepsilon).$$

Therefore, the $\mathcal{H}_{\phi^\varepsilon}$ Landau equation has right-hand side

$$Q_L^{\phi^\varepsilon}(f^\varepsilon, f^\varepsilon) = \nabla\cdot\left\{f^\varepsilon \int f_*^\varepsilon |v-v_*|^{2+\gamma}\Pi[v-v_*]\left(\nabla\left(\phi^\varepsilon * \left[\frac{\delta\mathcal{H}}{\delta(\mu*\phi^\varepsilon)}\right]\right) - \nabla_*\left(\phi^\varepsilon * \left[\frac{\delta \mathcal{H}_*}{\delta(\mu_* * \phi^\varepsilon)}\right]\right)\right) dv_*\right\},$$

and we ask if solutions, f^ε, converge to some f which solves (L).

Appendix A

Some inequalities and the spherical Laplacian

This appendix gathers various inequalities and identities primarily used in Chapters 2 and 3. As such, it contains material found in the joint works with José A. Carrillo, Matias G. Delgadino, and Laurent Desvillettes [31] and with José A. Carrillo and Matias G. Delgadino [32, Appendix A]. The former is a preprint submitted for publication and the latter has been published in Nonlinear Analysis volume 219, page 112824 in June 2022.

Lemma A.1. *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 $x = 0$ nor $y = 0$ or else the statement holds trivially. Let θ be an oriented angle between x and y. We expand the definition of $\Pi[x]$ to obtain

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

□

Lemma A.2 (Peetre)**.** *For any $p \in \mathbb{R}$ and $x, y \in \mathbb{R}^d$, we have*

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

Proof. Our proof follows [15]. 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 $a = x - y$ and $b = -y$, we obtain the inequality for $p = 2$

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

Taking non-negative 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 non-negative powers here proves the inequality for $p < 0$. □

Lemma A.3 (ALG inequality). *The logarithmic mean separates the arithmetic and geometric means;*

$$\sqrt{ab} \leq \frac{b-a}{\log b - \log a} \leq \frac{a+b}{2}, \quad \forall a,b>0,\ a\neq b.$$

Equality is achieved in any of the inequalities if and only if $a = b$ and equality is achieved in all of them, where the logarithmic mean between $a = b > 0$ is defined as a.

Proof. We follow the elegant proof by Sándor [110]. First we claim the following

$$\frac{4}{(t+1)^2} < \frac{1}{t} < \frac{1}{2\sqrt{t}} + \frac{1}{2t\sqrt{t}}, \quad \forall t>1. \tag{A.1}$$

The left-hand inequality of (A.1) follows from $(t+1)^2 - 4t = (t-1)^2 > 0$. For the right-hand inequality of (A.1), we use Young's (strict) inequality since $t > 1$

$$\frac{1}{t} = \frac{1}{t^{\frac{1}{4}}}\cdot\frac{1}{t^{\frac{3}{4}}} < \frac{1}{2t^{\frac{1}{2}}} + \frac{1}{2t^{\frac{3}{2}}}.$$

Assume without loss of generality that $0 < a < b$. We integrate (A.1) from 1 to $b/a > 1$. This leads to

$$2\frac{b-a}{b+a} < \log b - \log a < \sqrt{\frac{b}{a}} - \sqrt{\frac{a}{b}} = \frac{b-a}{\sqrt{ab}}.$$

Dividing by $b - a$ and inverting yields the result. □

Corollary A.1. *For all $a, b > 0$ with $a \neq b$ we have*

$$(a-b)\log\frac{a}{b} \geq 4|\sqrt{a}-\sqrt{b}|^2.$$

Proof. We write $a-b$ as a difference of squares to deduce

$$\begin{aligned}(a-b)\log\frac{a}{b} &= 2(\sqrt{a}-\sqrt{b})(\sqrt{a}+\sqrt{b})(\log\sqrt{a}-\log\sqrt{b})\\ &= 4|\sqrt{a}-\sqrt{b}|^2\left(\frac{\sqrt{a}+\sqrt{b}}{2}\right)\left(\frac{\log\sqrt{a}-\log\sqrt{b}}{\sqrt{a}-\sqrt{b}}\right)\\ &\geq 4|\sqrt{a}-\sqrt{b}|^2.\end{aligned}$$

In the last line, we used Lemma A.3 after recognising the arithmetic and logarithmic means between $\sqrt{a}$ and $\sqrt{b}$. □

Corollary A.2 ($\Lambda(f)$ bounds)**.** $\Lambda(f) = \frac{f'f'_* - ff_*}{\log f'f'_* - \log ff_*}$ *grows 'quadratically' in f (in the sense of tensor products);*

$$\sqrt{ff_*f'f'_*} < \Lambda(f) < \frac{f'f'_* + ff_*}{2}.$$

Moreover, there holds

$$(f'f'_* - ff_*)\log\frac{f'f'_*}{ff_*} \geq 4|\sqrt{f'f'_*} - \sqrt{ff_*}|^2.$$

Lemma A.4. *Suppose $(k,\, h,\, i)$ is an orthonormal basis of $\mathbb{R}^3$. There holds*

$$\int_0^{2\pi} (\cos\phi\, h + \sin\phi\, i) \otimes (\cos\phi\, h + \sin\phi\, i) d\phi = \pi\Pi[k] = \pi(I - k\otimes k).$$

This is equivalent to (c.f. Section 3.2.2)

$$\int_{\mathbb{S}^1_{k^\perp}} p\otimes p dp = \int_0^{2\pi} p\otimes p d\phi = \pi\Pi[k],$$

where p is orthonormal to k with azimuthal angle ϕ i.e. $p = \cos\phi\, h + \sin\phi\, i$.

Proof. In the basis of (k, h, i), we can represent the matrix in the integral as

$$\int_0^{2\pi} (\cos\phi\, h + \sin\phi\, i) \otimes (\cos\phi\, h + \sin\phi\, i) d\phi = \int_0^{2\pi} \begin{pmatrix} 0 & 0 & 0 \\ 0 & \cos^2\phi & \cos\phi\sin\phi \\ 0 & \cos\phi\sin\phi & \sin^2\phi \end{pmatrix} d\phi$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & \pi & 0 \\ 0 & 0 & \pi \end{pmatrix} = \pi\,(I - k\otimes k) = \pi\Pi[k].$$

□

Next we turn to explicit expressions for the spherical Laplacian/Laplace-Beltrami operator, see [91] for details. Consider a smooth function $f : \mathbb{R}^d \to \mathbb{R}$. Let us write $x \in \mathbb{R}^d$ as $x = r\omega$ for $r = |x|$ and $\omega = \frac{x}{|x|} \in \mathbb{S}^{d-1}$. The spherical Laplacian of f denoted $\Delta_{\mathbb{S}^{d-1}} f$ is obtained by

$$\Delta_{\mathbb{S}^{d-1}} f(x) = \Delta f\left(\frac{x}{|x|}\right)$$

and it satisfies

$$\Delta f = \partial_r^2 f + \frac{d-1}{r}\partial_r f + \frac{1}{r^2}\Delta_{\mathbb{S}^{d-1}} f. \tag{A.2}$$

We write the spherical Laplacian in terms of ∇_ω and $\Pi[\omega]$. Here, for functions $\phi = \phi(x)$ on $\mathbb{R}^d$, ∇_ω is the differential operator applied to the zero-homogeneous extension of ϕ on the sphere;

$$\nabla_\omega \phi = \nabla_x \phi\left(\frac{x}{|x|}\right).$$

$\nabla_\omega\cdot$ will denote the adjoint of ∇_ω.

Lemma A.5. *Under the spherical coordinates $x = r\omega$, and for smooth functions $\phi = \phi(x) = \phi(r\omega)$, the spherical Laplacian of ϕ reads*

$$\Delta_{\mathbb{S}^{d-1}} \phi = \nabla_\omega \cdot \left(\Pi[\omega]\nabla_\omega \phi\right).$$

Proof. We recompute $\Delta\phi$ using spherical coordinates. Then we identify the corresponding $\Delta_{\mathbb{S}^{d-1}}\phi$ term in (A.2). For every index i, the chain rule gives

$$[\nabla\phi]^i = \partial^i \phi = \frac{\partial r}{\partial x^i}\partial_r \phi + \frac{\partial \omega^j}{\partial x^i}\partial_{\omega^j}\phi = \omega^i \partial_r \phi + \frac{1}{r}\Pi^{ij}[\omega]\partial_{\omega^j}\phi.$$

Here, we have recalled the simple computations $\frac{\partial r}{\partial x^i} = \omega^i$ and $\frac{\partial \omega^j}{\partial x^i} = \frac{1}{r}\Pi^{ij}[\omega]$. Writing the Laplacian with repeated indices, we further compute

$$
\begin{aligned}
\Delta\phi = \partial^i\partial^i\phi &= \frac{\partial r}{\partial x^i}\partial_r\left(\omega^i\partial_r\phi + \frac{1}{r}\Pi^{ij}[\omega]\partial_{\omega^j}\phi\right) + \frac{\partial\omega^k}{\partial x^i}\partial_{\omega^k}\left(\omega^i\partial_r\phi + \frac{1}{r}\Pi^{ij}[\omega]\partial_{\omega^j}\phi\right) \\
&= \omega^i\left(\omega^i\partial_r^2\phi + \underbrace{\Pi^{ij}[\omega]\partial_r\left(\frac{1}{r}\partial_{\omega^j}\phi\right)}_{\perp\omega^i}\right) + \frac{1}{r}\Pi^{ik}[\omega]\left(\delta^{ik}\partial_r\phi + \underbrace{\omega^i\partial_{\omega^j}\partial_r\phi}_{\perp\Pi^{ik}[\omega]} + \frac{1}{r}\partial_{\omega^k}\left(\Pi^{ij}[\omega]\partial_{\omega^j}\phi\right)\right).
\end{aligned}
$$

Here, we have expanded the derivatives using the previous computations. In particular, the underbraced terms contribute nothing (as expected since these are the mixed derivatives in the radial and spherical directions). Recall now that

$$
\omega^i\omega^i = 1, \quad \Pi^{ik}[\omega]\delta^{ik} = \operatorname{trace}(\Pi[\omega]) = d-1, \quad \partial_{\omega^k}\Pi^{ij}[\omega] = -(\delta^{ik}\omega^j + \delta^{jk}\omega^i).
$$

Using these identities, we further simplify

$$
\begin{aligned}
\Delta\phi &= \partial_r^2\phi + \frac{(d-1)}{r}\partial_r\phi + \frac{1}{r^2}\Pi^{ik}[\omega]\left(-(\delta^{ik}\omega^j + \underbrace{\delta^{jk}\omega^i}_{\perp\Pi^{ik}[\omega]})\partial_{\omega^j}\phi + \Pi^{ij}[\omega]\partial_{\omega^k}\partial_{\omega^j}\phi\right) \\
&= \partial_r^2\phi + \frac{(d-1)}{r}\partial_r\phi + \frac{1}{r^2}\left(-(d-1)\omega^j\partial_{\omega^j}\phi + \Pi^{jk}[\omega]\partial_{\omega^k}\partial_{\omega^j}\phi\right).
\end{aligned}
$$

We can repackage the spherical Laplacian term in another neat way by noticing that

$$
\nabla_\omega\cdot(\Pi[\omega]\nabla_\omega\phi) = \partial_{\omega^k}\left(\Pi^{jk}[\omega]\partial_{\omega^j}\phi\right) = -(d-1)\omega^j\partial_{\omega^j}\phi + \Pi^{jk}[\omega]\partial_{\omega^k}\partial_{\omega^j}\phi.
$$

Putting this back gives

$$
\Delta\phi = \partial_r^2\phi + \frac{(d-1)}{r}\partial_r\phi + \frac{1}{r^2}\nabla_\omega\cdot(\Pi[\omega]\nabla_\omega\phi).
$$

□

Corollary A.3. *Under the same notations as Lemma A.5, in particular $x = r\omega$, for any smooth vector field V, we have*

$$
\nabla_x\cdot(\Pi[x]V) = \frac{1}{r}\nabla_\omega\cdot(\Pi[\omega]V).
$$

Moreover, for smooth ϕ, we have

$$
\nabla_x\cdot(\Pi[x]\nabla_x\phi) = \frac{1}{r^2}\nabla_\omega\cdot(\Pi[\omega]\nabla_\omega\phi).
$$

Proof. The first identity is a direct computation. For the second identity, repeat the calculations in

Lemma A.5 noticing that $\Pi[x]$ applied to $\nabla_x \phi$ removes the radial derivative contribution. □

Appendix B

An auxiliary PDE for Lemma 2.28

The content of this appendix is based on joint work with José A. Carrillo, Matias G. Delgadino, and Laurent Desvillettes [31, Appendix A]. It is a preprint submitted for publication.

In this section, we fix $\varepsilon > 0$ throughout and study the following PDE where the unknown to be solved for is $\mu \in C([0,T]; \mathscr{P}_2(\mathbb{R}^d))$,

$$\begin{cases} \partial_t \mu &= \nabla \cdot \{\mu \phi_{R_1} \int_{\mathbb{R}^d} \phi_{R_1 *} \psi_{R_2}(v - v_*)|v - v_*|^{\gamma+2} \Pi[v - v_*](J_0^\varepsilon - J_{0*}^\varepsilon) d\mu(v_*)\} \\ \mu(0) &= \mu_0 \in \mathscr{P}_2 \end{cases} . \tag{B.1}$$

For $R_1, R_2 > 0$, the functions $0 \le \phi_{R_1}, \psi_{R_2} \le 1$ are smooth cut-off functions approximating the identity in different ways

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

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

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

The main result of this appendix is

Theorem B.1. *Fix $\varepsilon, R_1, R_2 > 0$, $\gamma \in \mathbb{R}$, 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* (B.1).

By Lemma 2.25, 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 functions ϕ_{R_1} and $\phi_{R_1 *}$ cut off the growth of J_0^ε and J_{0*}^ε for large velocities v, v_*. The $\psi_{R_2}(v - v_*)$ term treats the possible singularities around $v = v_*$ from the weight $|v - v_*|^{2+\gamma}$. In total, these cut-off functions allow us to interpret (B.1) as a continuity equation with a smooth and compactly supported velocity field.

The construction of the solution in Theorem B.1 is given in two steps. Firstly, a local well-posedness theory established to some finite time interval $T > 0$ which depends on ε, γ, 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 (B.1).

We fix $T > 0$ to be determined explicitly later. Our proof strategy closely follows the fixed point argument of [26] in the space $C([0,T];\mathscr{P}_2)$ with the following metric

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

where W_2 is the 2-Wasserstein distance on $\mathscr{P}_2$ [121].

Remark B.1. *The growth of J_0^ε is taken care of by the ϕ_{R_1} and ϕ_{R_1*} terms. Hence the results of this section can be applied when replacing the convolution kernel of J_0^ε with general tailed exponential distributions $G^{s,\varepsilon}(v)$ for $s > 0$.*

For $\mu \in \mathscr{P}_2$, we denote by $U[\mu](v)$ the following function

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

so that the PDE in (B.1) can be written as a nonlinear continuity equation

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

To fix ideas, the weak formulation of (B.1) means that the following equality holds for all test functions $\tau \in C_c^\infty(\mathbb{R}^d)$ and times $t \in [0,T]$

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

Thanks to all the smooth cutoffs and $\mu_0 \in \mathscr{P}_2$, we can enlarge the class of test functions to smooth functions with quadratic growth. In particular, by choosing $\tau(v) = |v|^2$ and symmetrising the right-hand side by swapping $v \leftrightarrow v_*$, we see that the second moment of μ_0 is conserved along the evolution of (B.1).

Our first task is to study the characteristic equation associated to (B.1).

Lemma B.2 (Characteristic equation). *For any $T > 0$, $\mu \in C([0,T]; \mathscr{P}_2)$, 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 following ODE*

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

Furthermore, the growth rate satisfies

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

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 denote by Φ^t_μ the flow map associated to the characteristic ODE, so that

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

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

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

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

$$\int_{\mathbb{R}^d} \tau(v) d(\Phi^t_\nu \# \mu_0)(v) = \int_{\mathbb{R}^d} \tau(\Phi^t_\nu(v)) d\mu_0(v).$$

Clearly, a unique fixed point of $\mu \mapsto \Phi^t_\mu \# \mu_0$ would solve (B.1). 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 B.3 (L^∞ estimate for velocity field). *There exists a constant $C = C(\varepsilon, \gamma, R_1, R_2, \mu_0) > 0$ such that for every $T > 0$ and $\nu \in C([0,T]; \mathscr{P}_2)$, we have*

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

Proof. Estimate for $\gamma \geq -2$:

We have the following three inequalities

$$|v - v_*|^{\gamma+2} \lesssim_\gamma |v|^{\gamma+2} + |v_*|^{\gamma+2}, \quad \|\Pi[v - v_*]\| \leq 1, \quad J^\varepsilon_0 \lesssim_{\varepsilon,\mu_0} 1$$

due to the range of γ, boundedness of Π, and Lemma 2.25, respectively. These three inequalities provide the estimate

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

where we have dropped ψ_{R_2} altogether. Using the compact support of ϕ_{R_1}, we obtain

$$|U[\nu(t)](v)| \lesssim_{\gamma,\varepsilon,\mu_0} \phi_{R_1}(v)(R_1^{2+\gamma} + \langle v\rangle^{2+\gamma}) \lesssim_{R_1} \phi_{R_1}(v)\langle v\rangle^{2+\gamma}.$$

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

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

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

$$|U[\nu(t)](v)| \lesssim_{\gamma,\varepsilon,\mu_0,R_2} \phi_{R_1}(v) \int_{\mathbb{R}^d} \phi_{R_1}(v_*) d\nu_t(v_*) \leq 1,$$

which concludes the proof. □

The next result follows exactly as in [26].

Lemma B.4 (Time continuity of flow map)**.** *Let $C = C(\varepsilon, \gamma, R_1, R_2, \mu_0) > 0$ be the same constant from Lemma B.3. Then for any $T > 0$, and $\nu \in C([0,T]; \mathscr{P}_2)$ we have*

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

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 B.5. *Define*

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

The function F is smooth and compactly supported. In particular, for every $k, l \in \mathbb{N}$, there is a constant

$C = C(\varepsilon, \gamma, R_1, R_2, \mu_0, k, l) > 0$ *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 2.25.

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 $\psi_{R_2}(v - w)$. □

Corollary B.1 (Pointwise and measurewise regularity of U)**.** *Consider the constant $C = C(\varepsilon, \gamma, R_1, R_2, \mu_0, k, l) > 0$ from Lemma B.5 above. We have the following*

1. *Take $C_1 = C(\varepsilon, \gamma, R_1, R_2, \mu_0, 0, 1) > 0$. For every $T > 0; \nu^1, \nu^2 \in C([0,T]; \mathscr{P}_2)$; $t \in [0,T]$; $v \in \mathbb{R}^d$ we have the estimate*
$$|U[\nu^1(t)](v) - U[\nu^2(t)](v)| \leq C_1 W_2(\nu_t^1, \nu_t^2).$$

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

Remark B.2. *By considering the anti-symmetric property of F when swapping variables $v \leftrightarrow w$, one really obtains $C_1 = C_2$. Their distinction in this corollary is artificial.*

Proof. Item 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 $\nu^1(t)$ and $\nu^2(t)$ which exists, see [121]. We estimate the difference with notation from Lemma B.5

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

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.

Item 2:

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

$$
\begin{aligned}
|U[\nu(t)](v_1) - U[\nu(t)](v_2)| &= \left| \int_{\mathbb{R}^d} F(v_1, w) - F(v_2, w) d\nu_t(w) \right| \\
&\leq \int_{\mathbb{R}^d} |F(v_1, w) - F(v_2, w)| d\nu_t(w) \leq C_2 |v_1 - v_2|.
\end{aligned}
$$

Once more, a Mean-Value type estimate is applied (in the first variable of F). □

The next result combines both items of Corollary B.1 to estimate the regularity of the flow map with respect to measures and follows exactly as in [26].

Lemma B.6 (Continuity of flow map with respect to measures)**.** *For $T > 0$ fix any $\nu^1, \nu^2 \in C([0, T]; \mathscr{P}_2)$ and $t \in [0, T]$. With $C := C_1 = C_2$ the same constants in Corollary B.1, we have the estimate*

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

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

The proof of Theorem B.1 is now classical from Corollary B.1 and Lemma B.6, see [26, 29, 68] for instance. The time of existence can be given by any $0 < T < \frac{1}{C} \log 2$ where $C > 0$ is chosen as in Lemma B.6 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 by the evolution of (B.1) and so the maximal time of existence is $+\infty$.

Appendix C

Strong compactness from bounded Boltzmann dissipation

The content of this appendix is based on joint work with José A. Carrillo and Matias G. Delgadino [32, Appendix B]. It has been published in Nonlinear Analysis volume 219, page 112824 in June 2022.

The purpose of this appendix is to derive an estimate guaranteeing strong compactness in the grazing collision limit $\sqrt{f^\epsilon} \to \sqrt{f}$ in L^2_{loc} for Proposition 3.10. We repeat here the main results of [2, 4] which we emphasise are independent of the grazing collision parameter $\epsilon \downarrow 0$ provided the finite angular momentum transfer (3.2) and uniform moments and entropy bounds hold (3.5). Let f_R denote $f\chi_R$ where χ_R is a smooth cut-off function equal to 1 in B_R and vanishing outside of B_{R+1} (we make this precise later). The estimate we wish to show is

$$\int_{\mathbb{R}^3} \left|\mathcal{F}\left[\sqrt{f_R^\epsilon}\right](\xi)\right|^2 \min(|\xi|^2,\, |\xi|^\nu) d\xi \leq C_R(D_B^\epsilon(f^\epsilon) + 1), \quad \forall R > 1, \tag{C.1}$$

where $\mathcal{F}$ stands for the Fourier transform and the constant $C_R > 0$ depending on $R > 1$ is *independent* of $\epsilon > 0$. We recall from Section 3.2 that $\nu > 0$ is the quantity which controls the angular singularity of the collision kernel. Adhering to (**A3.2**), we insist on decoupling $\nu \in (0,2)$, $\gamma \in [-4, 0]$.

As in [2], we first outline the main steps and postpone the details. We show (C.1) for the particular Boltzmann collision kernel

$$B^\epsilon(z,\sigma) = |z|^\gamma_{kin} b^\epsilon\left(\frac{z}{|z|}\cdot\sigma\right), \quad |z|^\gamma_{kin} = \begin{cases} 1 & |z| \leq 1 \\ |z|^\gamma & |z| \geq 1 \end{cases} \leq 1.$$

The dissipation associated to this kernel is certainly less than the dissipation for those kernels without cutting off in the kinetic singularity near $v = v_*$ (such as those we consider from (**A3.2**)).

Proof of (C.1). For ease of notation, we identify $f \equiv f^\epsilon$. Proving (C.1) in this setting of cut-off kinetic singularities clearly implies the full generality of the result since $|z|_{kin}^\gamma \le |z|^\gamma$. Cutting off the angular singularity part of B^ϵ if necessary, and then passing to the limit, we can rewrite the Boltzmann dissipation using the pre-post-collisional change of velocities

$$\begin{aligned} D_B^\epsilon(f) &= -\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} B^\epsilon(f'f'_* - ff_*)\log f d\sigma dv_* dv = \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} B^\epsilon ff_*\log\frac{f}{f'} d\sigma dv_* dv \\ &= \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} B^\epsilon f_*\left(f\log\frac{f}{f'} - f + f'\right) d\sigma dv_* dv + \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} B^\epsilon f_*(f - f')d\sigma dv_* dv. \end{aligned}$$

According to the cancellation lemma (Lemma C.1), we can estimate the second integral with

$$\left|\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} B^\epsilon f_*(f' - f)d\sigma dv_* dv\right| \le C_1,$$

with C_1 being a constant depending only on the moments and entropy. For the first integral, we make the square root appear with the classical inequality

$$x\log\frac{x}{y} - x + y \ge (\sqrt{x} - \sqrt{y})^2, \quad \forall x, y > 0,$$

which can be proven by reducing to the case $y = 1$ and applying the ALG inequality (Lemma A.3). Continuing, we have

$$\begin{aligned} D_B^\epsilon(f) + C_1 &\ge \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} B^\epsilon f_*(\sqrt{f'} - \sqrt{f})^2 d\sigma dv_* dv \\ &= \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} |v - v_*|_{kin}^\gamma b^\epsilon(k\cdot\sigma) f_*(\sqrt{f'} - \sqrt{f})^2 d\sigma dv_* dv, \end{aligned}$$

and we also recall $k = (v - v_*)/|v - v_*|$. We set now $F(v) = \sqrt{f(v)}$ and use F_*, F', F'_* as usual. Having revealed $\sqrt{f}$, we apply a smooth cut-off and pass to Fourier space. For $R > 1$ we take $\chi_R \in C_c^\infty(\mathbb{R}^3)$ a smooth indicator function on B_R such that $0 \le \chi_R \le 1$, $\chi_R|_{B_R} = 1$, and $\mathrm{supp}\chi_R \subset B_{R+1}$. According to the truncation lemma (Lemma C.2), there are constants C_2, $C_3 > 0$ such that C_2 depends only on (**A3.1**) while C_3 depends only on R and γ such that

$$\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} B^\epsilon f_*(\sqrt{f'} - \sqrt{f})^2 d\sigma dv_* dv + C_2 \ge C_3 \iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} b^\epsilon(k\cdot\sigma) f_*\chi_{R*}(F'\chi'_R - F\chi_R)^2 d\sigma dv_* dv.$$

Using Lemma C.3, we are able to pass to Fourier variables so that the last integral can be minorised by

$$\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} b^\epsilon(k\cdot\sigma)f_*\chi_{R*}(F'\chi_R' - F\chi_R)^2 d\sigma dv_* dv$$
$$\geq \frac{1}{2(2\pi)^3}\int_{\mathbb{R}^3}|\mathcal{F}[F\chi_R](\xi)|^2\left\{\int_{\mathbb{S}^2} b^\epsilon\left(\frac{\xi}{|\xi|}\cdot\sigma\right)(\mathcal{F}[f\chi_R](0) - |\mathcal{F}[f\chi_R](\xi^-)|)d\sigma\right\}d\xi,$$

where $\xi^- = (\xi - |\xi|\sigma)/2$. Finally, the integral in curly brackets can be estimated using Lemma C.4 so that there is a constant $C_4 > 0$ depending on the uniform bounds of moments and entropy and finite angular momentum transfer (3.2) giving

$$\int_{\mathbb{S}^2} b^\epsilon\left(\frac{\xi}{|\xi|}\cdot\sigma\right)(\mathcal{F}[f\chi_R](0) - |\mathcal{F}[f\chi_R](\xi^-)|)d\sigma \geq C_4\min(|\xi|^2, |\xi|^\nu).$$

Putting these considerations together, we have

$$\frac{C_3C_4}{2(2\pi)^3}\int_{\mathbb{R}^3}|\mathcal{F}[F\chi_R](\xi)|^2\min(|\xi|^2,|\xi|^\nu)d\xi \leq D_B^\epsilon(f) + C_1 + C_2.$$

□

The rest of this section is devoted to (re)proving the lemmas that were invoked in the previous proof. In particular, we repeat the proofs involving estimates pertaining to the collision kernel B^ϵ since we want to make certain that our constants are independent of $\epsilon > 0$.

Lemma C.1 (Cancellation lemma)**.** *For almost every $v_* \in \mathbb{R}^3$ and $\epsilon > 0$ sufficiently small, we have*

$$\int_{\mathbb{R}^3}\int_{\mathbb{S}^2} B^\epsilon(v - v_*, \sigma)(f' - f)d\sigma dv = [f * S^\epsilon](v_*),$$

where S^ϵ is given by

$$S^\epsilon(z) = 2\pi|z|_{kin}^\gamma \int_0^{\epsilon/2}\left[\cos^{-3}\left(\frac{\theta}{2}\right) - 1\right]\beta^\epsilon(\theta)d\theta.$$

Moreover, we have the trivial estimate $|S^\epsilon(z)| \leq 12$. Finally, the previous estimates lead to

$$\left|\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} B^\epsilon f_*(f' - f)\right| = \left|\iint_{\mathbb{R}^6} f(v)f(v_*)S^\epsilon(v - v_*)\right| \leq 12.$$

Proof. As in the beginning of the proof of (C.1), we split the 'gain' and 'loss' part of the integral by an approximation argument, cutting off the angular singularity as necessary. Focusing on the gain term, for fixed $\sigma \in \mathbb{S}^2$ and $v_* \in \mathbb{R}^3$, we consider the change of coordinates $v \mapsto v'$.

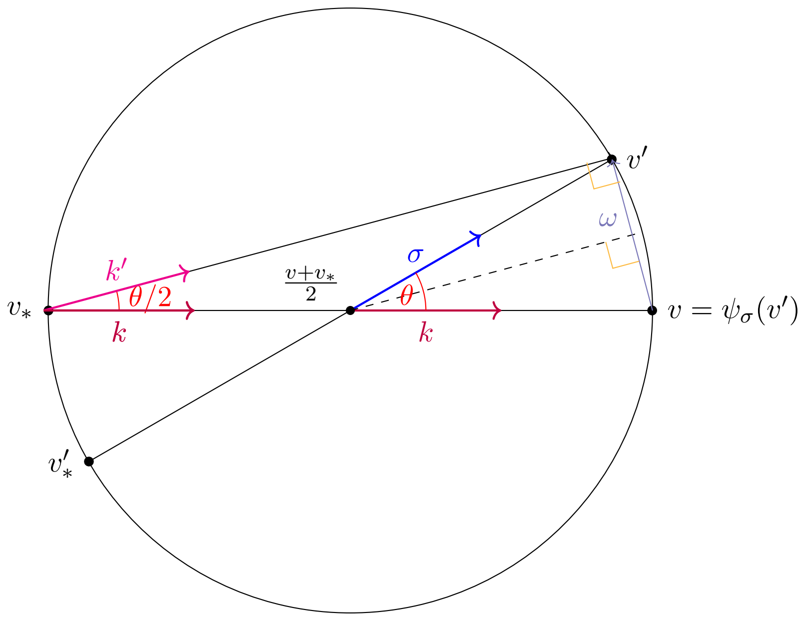


Figure C.1: Geometry of elastic binary collisions with additional angles.

Recalling

$$v' = \frac{v+v_*}{2} + \frac{|v-v_*|}{2}\sigma = v - \frac{|v-v_*|}{2}(\sigma - k) = v_* + \frac{|v-v_*|}{2}(k+\sigma),$$

the first of these identities implies the following equality for the Jacobian

$$\left|\frac{\partial v'}{\partial v}\right| = \left|\frac{1}{2}I + \frac{1}{2}k \otimes \sigma\right| = \frac{1}{8}(1 + k\cdot\sigma).$$

Graphically, see Figure C.1, we can switch from $k = \frac{v-v_*}{|v-v_*|}$ to $k' = \frac{v'-v_*}{|v'-v_*|}$ using the standard half-angle trigonometric identity $1 + k\cdot\sigma = 1 + \cos\theta = 2\cos^2\frac{\theta}{2} = 2(k'\cdot\sigma)^2$, where the last equality can be seen pictorially or by employing the same trigonometric identity from the definition of k' using

$$k'\cdot\sigma = \frac{1+k\cdot\sigma}{|k+\sigma|}, \quad |k+\sigma| = 2\cos\frac{\theta}{2}.$$

This leads to another form of the Jacobian determinant

$$\left|\frac{\partial v'}{\partial v}\right| = \frac{1}{8}(1+k\cdot\sigma) = \frac{(k'\cdot\sigma)^2}{4}.$$

Now, since $\theta \in [0,\pi/2]$ (see Section 3.2.2), we therefore have $k'\cdot\sigma = \cos\frac{\theta}{2} \geq \frac{1}{\sqrt{2}}$. This shows that the transformation is invertible and we define the inverse transfromation $v' \mapsto \psi_\sigma(v') = v$. Employing similar trigonometric identities as before, some computations lead to

$$|v-v_*| = \frac{|v'-v_*|}{k'\cdot\sigma} \iff |\psi_\sigma(v) - v_*| = \frac{|v-v_*|}{k\cdot\sigma},$$

since the collision map is involutive. Returning to the change of variable, we have

$$\begin{aligned}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2} B^\epsilon(|v-v_*|, k\cdot\sigma) f(v') dv d\sigma &= \int_{\mathbb{R}^3}\int_{\mathbb{S}^2} B^\epsilon(|v-v_*|, k\cdot\sigma) f(v') \left|\frac{\partial v}{\partial v'}\right| dv' d\sigma \\ &= \int_{\mathbb{R}^3}\int_{\mathbb{S}^2} B^\epsilon(|v-v_*|, 2(k'\cdot\sigma)^2 - 1) f(v') \frac{4}{(k'\cdot\sigma)^2} dv' d\sigma \\ &= \int_{k\cdot\sigma\geq 1/\sqrt{2}} B^\epsilon(|\psi_\sigma(v) - v_*|, 2(k\cdot\sigma)^2 - 1) f(v) \frac{4}{(k\cdot\sigma)^2} dv d\sigma,\end{aligned}$$

where we just relabel $v \leftrightarrow v'$ in the last line. Inserting this back into the difference, we obtain

$$\begin{aligned}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2} B^\epsilon(|v-v_*|, k\cdot\sigma)(f'-f) d\sigma dv = \int_{\mathbb{R}^3} f(v) \Bigg[&\int_{k\cdot\sigma\geq 1/\sqrt{2}} B^\epsilon\left(\frac{|v-v_*|}{k\cdot\sigma}, 2(k\cdot\sigma)^2 - 1\right)\frac{4}{(k\cdot\sigma)^2} d\sigma \\ &- \int_{k\cdot\sigma\geq 0} B^\epsilon(v-v_*, k\cdot\sigma) d\sigma\Bigg] dv.\end{aligned}$$

Thus, we identify the term in square brackets as $S^\epsilon(|v-v_*|)$. Focusing again on the gain part, we change to spherical coordinates (see Figure 3.2), remembering now that $\cos\theta = k\cdot\sigma \geq 1/\sqrt{2}$, so $\theta \in [0, \pi/4]$ and therefore, we have

$$\begin{aligned}\int_{k\cdot\sigma\geq 1/\sqrt{2}} B^\epsilon\left(\frac{|v-v_*|}{k\cdot\sigma}, 2(k\cdot\sigma)^2 - 1\right)\frac{4}{(k\cdot\sigma)^2} d\sigma &= \int_{\mathbb{S}^1_{k^\perp}}\int_0^{\pi/4} \frac{4\sin\theta}{\cos^2\theta} B^\epsilon\left(\frac{|v-v_*|}{\cos\theta}, \cos 2\theta\right) d\theta dp \\ &= 2\pi\int_0^{\pi/4} \frac{2\sin 2\theta}{\cos^3\theta} B^\epsilon\left(\frac{|v-v_*|}{\cos\theta}, \cos 2\theta\right) d\theta \\ &= 2\pi |v-v_*|^\gamma_{kin} \int_0^{\epsilon/2} \cos^{-3}\left(\frac{\theta}{2}\right)\beta^\epsilon(\theta) d\theta.\end{aligned}$$

In the last line, we doubled the integration region while also decomposing B^ϵ with respect to β^ϵ with $\epsilon > 0$ sufficiently small so that $\theta \in [0, \epsilon/2] \implies \cos^\gamma_{kin}(\theta) = 1$. This completes the identification of S^ϵ.

Turning to the L^∞ bound, we note that the fundamental theorem of calculus gives the estimate

$$\cos^{-3}\frac{\theta}{2} - 1 = \int_0^1 \frac{d}{dt}\cos^{-3}\frac{t\theta}{2} dt = \frac{3}{2}\theta\int_0^1 \cos^{-4}\left(\frac{t\theta}{2}\right)\sin\left(\frac{t\theta}{2}\right) dt \leq \frac{3}{2}\theta\sin\left(\frac{\theta}{2}\right) \sim \frac{3}{4}\theta^2.$$

Thus, we obtain

$$|S^\epsilon(z)| \leq \frac{3\pi}{2}\int_0^{\epsilon/2} \theta^2\beta^\epsilon(\theta) d\theta \leq 12.$$

The final estimate of the lemma is now easy because

$$\left|\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} B^\epsilon f_*(f'-f) d\sigma dv_* dv\right| = \left|\iint_{\mathbb{R}^6} f_* f S^\epsilon(v-v_*) dv dv_*\right| \leq 12.$$

□

Lemma C.2 (Truncation lemma)**.** *One can take the constant*

$$C_2 = 150\pi \left(\iint_{\mathbb{R}^6} (|v|^2 + |v_*|^2) f_* f dv_* dv \right) \left(\int_{\theta=0}^{\epsilon/2} \theta^2 \beta^\epsilon(\theta) d\theta \right) < +\infty$$

such that for all $R > 1$, we have

$$\iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} B^\epsilon f_* (F' - F)^2 d\sigma dv_* dv + C_2 \geq \frac{(2\sqrt{2}(R+1))^\gamma}{2} \iint_{\mathbb{R}^6} \int_{\mathbb{S}^2} b^\epsilon(k \cdot \sigma) f_* \chi_{R*} (F' \chi_R' - F \chi_R)^2 d\sigma dv_* dv,$$

where we recall the notation $F = \sqrt{f}$ and $\chi_R \in C_c^\infty(\mathbb{R}^3)$ is a smooth indicator function such that

$$0 \leq \chi_R \leq 1, \quad \chi_R|_{B_R} = 1, \quad supp \chi_R \subset B_{R+1}.$$

Proof. Firstly, it is clear that $f_*(F' - F)^2 \geq f_* \chi_{R*} (F' - F)^2 \chi_R^2$. We wish to pair the indicators with F in the right velocity variables so we estimate

$$(F' \chi_R' - F \chi_R)^2 = (F'(\chi_R' - \chi_R) + (F' - F)\chi_R)^2 \leq 2F'^2 (\chi_R' - \chi_R)^2 + 2(F' - F)^2 \chi_R^2.$$

Including B^ϵ, we have

$$\begin{aligned} B^\epsilon(|v - v_*|, k \cdot \sigma)(F' - F)^2 &\geq |v - v_*|_{kin}^\gamma b^\epsilon(k \cdot \sigma) \chi_{R*} (F' - F)^2 \chi_R^2 \\ &\geq |v - v_*|_{kin}^\gamma b^\epsilon(k \cdot \sigma) \left[\frac{1}{2} \chi_{R*} (F' \chi_R' - F \chi_R)^2 - \chi_{R*} F'^2 (\chi_R' - \chi_R)^2 \right]. \end{aligned} \tag{C.2}$$

Now, for the second term with the minus sign, we use similar Mean-Value estimates as in Lemma 3.4 to deduce $(\chi_R' - \chi_R)^2 \leq \text{Lip}(\chi_R)^2 |v - v_*|^2 |\sigma - k|^2$. We choose χ_R in such a way that $\text{Lip}(\chi_R) \leq 2$ (i.e. its height changes by 1 over a horizontal distance of 0.95). Before proceeding with the estimate of the second term, we write down the following relation which can be obtained as in the proof of the cancellation lemma (Lemma C.1)

$$|v' - v_*| = \frac{1}{2} |\sigma + k| |v - v_*|. \tag{C.3}$$

Recalling $dv = \frac{4}{(k' \cdot \sigma)^2} dv'$ and similar pre-post-collision velocity relations from Lemma C.1, the integral of the second term of (C.2) can be estimated by

$$\begin{aligned}&\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|v-v_*|^\gamma_{kin}b^\epsilon(k\cdot\sigma)f_*\chi_{R*}f'(\chi_R'-\chi_R)^2d\sigma dv_*dv\\ &\le 16\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|v-v_*|^\gamma_{kin}|v-v_*|^2b^\epsilon(2(k'\cdot\sigma)^2-1)f_*\chi_{R*}f'\frac{|\sigma-k|^2}{(k'\cdot\sigma)^2}dv'dv_*d\sigma.\end{aligned}$$

By expanding the square, one obtains

$$|\sigma+k|^2=4(k'\cdot\sigma)^2,\quad |\sigma-k|^2=4(1-(k'\cdot\sigma)^2),\tag{C.4}$$

which allows the estimate to continue as

$$\begin{aligned}&\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|v-v_*|^\gamma_{kin}b^\epsilon(k\cdot\sigma)f_*\chi_{R*}f'(\chi_R'-\chi_R)^2d\sigma dv_*dv\\ &\le 64\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}\left(\frac{|v'-v_*|}{(k'\cdot\sigma)}\right)^\gamma_{kin}|v'-v_*|^2b^\epsilon(2(k'\cdot\sigma)^2-1)f_*\chi_{R*}f'\frac{1-(k'\cdot\sigma)^2}{(k'\cdot\sigma)^4}dv'dv_*d\sigma.\end{aligned}$$

Relabelling v' as v and moving to polar coordinates, we finally have

$$\begin{aligned}&\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}|v-v_*|^\gamma_{kin}b^\epsilon(k\cdot\sigma)f_*\chi_{R*}f'(\chi_R'-\chi_R)^2d\sigma dv_*dv\\ &\le 64\iint_{\mathbb{R}^6}\int_{\mathbb{S}^1_{k^\perp}}\int_{\theta=0}^{\epsilon/2}\left(\frac{|v-v_*|}{\cos\theta}\right)^\gamma_{kin}|v-v_*|^2\beta^\epsilon(2\theta)f_*\chi_{R*}f\cos^{-4}\theta(1-\cos^2\theta)d\theta dp dv_*dv\\ &\le 150\pi\left(\iint_{\mathbb{R}^6}(|v|^2+|v_*|^2)f_*f dv_*dv\right)\left(\int_{\theta=0}^{\epsilon/2}\theta^2\beta^\epsilon(\theta)d\theta\right)=:C_2<+\infty.\end{aligned}$$

In the last inequality, we bluntly estimated the negative powers of $\cos\theta\sim 1$ since $\theta\le\epsilon/2$.

Turning to the first term of (C.2), we combine (C.3) and (C.4) together with the identification $k'\cdot\sigma=\cos\theta/2$ (see the proof of Lemma C.1) to deduce for $\theta\in[0,\pi/2]$

$$|v'-v_*|=\cos\frac{\theta}{2}|v-v_*|\implies |v'-v_*|\le|v-v_*|\le\sqrt{2}|v'-v_*|.$$

This implies that whenever $|v_*|\le R+1$ and at least one of $|v|\le R+1$ or $|v'|\le R+1$ hold, we immediately obtain $|v-v_*|^2\le 8(R+1)^2$. In this case, we can estimate the kinetic contribution

$$|v-v_*|^\gamma_{kin}\ge(2\sqrt{2}(R+1))^\gamma.$$

Adding C_2 to both sides of (C.2) and integrating, we obtain

$$2\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}B^\epsilon f_*(F'-F)^2+2C_2\ge(2\sqrt{2}(R+1))^\gamma\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2}b^\epsilon(k\cdot\sigma)f_*\chi_{R*}(F'\chi_R'-F\chi_R)^2.$$

$\square$

Lemma C.3 (Fourier representation)**.** *For $f \in L^1(\mathbb{R}^3)$ and $f \geq 0$, we have*

$$\iint_{\mathbb{R}^6}\int_{\mathbb{S}^2} b^\epsilon(k\cdot\sigma) f_*(F'-F)^2 d\sigma dv_* dv$$
$$\geq \frac{1}{2(2\pi)^3}\int_{\mathbb{R}^3}|\mathcal{F}[F](\xi)|^2\left\{\int_{\mathbb{S}^2} b^\epsilon\left(\frac{\xi}{|\xi|}\cdot\sigma\right)(\mathcal{F}[f](0)-|\mathcal{F}[f](\xi^-)|)d\sigma\right\}d\xi,$$

with the notation $\xi^- = \frac{\xi-|\xi|\sigma}{2}$ recalling $F=\sqrt{f}$.

Furthermore, there is a constant C_f depending only on bounds for the entropy, mass, and energy of f such that for every $\xi \in \mathbb{R}^3$ we have

$$\mathcal{F}[f](0)-|\mathcal{F}[f](\xi)| \geq C_f \min(|\xi|^2, 1).$$

For the first part of the lemma, we direct the reader to [2, Section 5, Corollary 3]. We only show the second estimate of the result to verify that the constant C_f can be taken independently of $\epsilon > 0$.

Proof. Recall that for real numbers a, b one can take $\theta = \tan^{-1}(a/b) \in \mathbb{R}$ such that $\sqrt{a^2+b^2} = a\cos\theta - b\sin\theta$. Applying this to the real and imaginary parts of the Fourier transform of f, there is some $\theta \in \mathbb{R}$ such that

$$\mathcal{F}[f](0)-|\mathcal{F}[f](\xi)| = \int_{\mathbb{R}^3} f(v)(1-\cos(v\cdot\xi+\theta))dv = 2\int_{\mathbb{R}^3} f(v)\sin^2\left(\frac{v\cdot\xi+\theta}{2}\right)dv$$
$$\geq 2\sin^2\delta\int_{B_r\cap A_\delta} f(v)dv.$$

Here, $r>0$ is some (large) radius to be specified later. For $\delta>0$, we consider the set $A_\delta := \{v\in\mathbb{R}^3 : \forall p\in\mathbb{Z},\ |v\cdot\xi+\theta-2\pi p| \geq 2\delta\}$. The partition $\mathbb{R}^d = (\mathbb{R}^d\setminus B_r)\cup(B_r\cap A_\delta)\cup(B_r\cap(\mathbb{R}^d\setminus A_\delta))$ leads to the estimate

$$\sin^2\delta\int_{B_r\cap A_\delta} f(v)dv = \sin^2\delta\left(\int_{\mathbb{R}^d} - \int_{\mathbb{R}^d\cap B_r} - \int_{B_r\cap(\mathbb{R}^d\setminus A_\delta)}\right) f(v)dv$$
$$\geq \sin^2\delta\left(\|f\|_{L^1} - \frac{\|f\|_{L^1_2}}{r^2} - \int_{B_r\cap(\mathbb{R}^d\setminus A_\delta)} f(v)dv\right).$$

We now further investigate the set $B_r\cap(\mathbb{R}^d\setminus A_\delta) = \{v\in\mathbb{R}^3 : |v|\leq r, \exists p\in\mathbb{Z} \text{ s.t. } |v\cdot\xi+\theta-2\pi p|\leq 2\delta\}$. By considering (rotate and translate v as appropriate) $\xi = ke_1$, $\theta = 0$, with $k>0$ and $e_1 = (1,0,0)$, one can show

$$|B_r\cap(\mathbb{R}^d\setminus A_\delta)| \leq \frac{4\delta}{|\xi|}(2r)^{d-1}\left(3+\frac{r|\xi|}{\pi}\right).$$

More precisely, one should think of $B_r\cap(\mathbb{R}^d\setminus A_{\delta=0})$ as the set of integer lattice points in B_r lying along the axial direction of $\xi/|\xi|$. So the inequality above estimates the measure of a δ neighbourhood version of this set. Continuing the estimate, we thus obtain

$$\mathcal{F}[f](0) - |\mathcal{F}[f](\xi)| \geq 2\sin^2\delta \left(\|f\|_{L^1} - \frac{\|f\|_{L^1_2}}{r^2} - \sup_{|A| \leq \frac{4\delta}{|\xi|}(2r)^{d-1}\left(3+\frac{r|\xi|}{\pi}\right)} \int_A f(v)dv \right). \tag{C.5}$$

In the case $|\xi| \geq 1$, notice that

$$\frac{4\delta}{|\xi|}(2r)^{d-1}\left(3 + \frac{r|\xi|}{\pi}\right) \leq 12\delta(2r)^{d-1} + \frac{6\delta}{\pi}(2r)^d.$$

Therefore, choose large $r > 0$ and small $\delta > 0$ such that the bracketed quantity is strictly positive (appealing to equi-integrability of f). In the case $|\xi| \leq 1$, one again chooses large $r > 0$ but small $\delta \sim |\xi|$ so that $\sin^2\delta \geq |\xi|^2$. □

Lemma C.4 (Fourier average estimate)**.** *For every $\xi \in \mathbb{R}^3$ and $\epsilon \leq 1$ we have*

$$\int_{\mathbb{S}^2} b^\epsilon\left(\frac{\xi}{|\xi|}\cdot\sigma\right)\min(|\xi^-|^2, 1)d\sigma \geq \frac{2c_1}{\pi}\left(\int_0^{\pi/2}\phi^{1-\nu}d\phi\right)\min(|\xi|^2, |\xi|^\nu),$$

*recalling the notations c_1, ν from (**A3.2**) and $\xi^- = \frac{\xi - |\xi|\sigma}{2}$.*

Proof. From the definition of ξ^-, we have

$$|\xi^-|^2 = \frac{|\xi|^2}{2}\left(1 - \frac{\xi}{|\xi|}\cdot\sigma\right).$$

Using spherical coordinates with radial direction given by $\xi/|\xi|$ (see Section 3.2.2), we use the lower bound of (3.12) and directly integrate over $\mathbb{S}^1_{\xi^\perp}$ to obtain

$$\begin{aligned}\int_{\mathbb{S}^2} b^\epsilon\left(\frac{\xi}{|\xi|}\cdot\sigma\right)\min(|\xi^-|^2, 1)d\sigma &= \int_{\mathbb{S}^1_{\xi^\perp}}\int_{\theta=0}^{\epsilon/2}\beta^\epsilon(\theta)\min\left(\frac{|\xi|^2}{2}(1-\cos\theta), 1\right)d\theta d\xi^\perp \\ &\geq \frac{4}{\pi}\int_{\theta=0}^{\epsilon/2}\beta^\epsilon(\theta)\min\left(\frac{|\xi|^2\theta^2}{2}, 1\right)d\theta.\end{aligned}$$

We introduce the change of variables $\theta = \epsilon\chi/\pi$ and the lower bound of (**A3.2**) giving

$$\int_{\mathbb{S}^2} b^\epsilon\left(\frac{\xi}{|\xi|}\cdot\sigma\right)\min(|\xi^-|^2, 1)d\sigma \geq \frac{4}{\pi}\int_0^{\pi/2}\beta(\chi)\min\left(\frac{|\xi|^2\chi^2}{2}, \frac{\pi^2}{\epsilon^2}\right)d\chi \geq \frac{4c_1}{\pi}\int_0^{\pi/2}\min\left(\frac{|\xi|^2\chi^2}{2}, \frac{\pi^2}{\epsilon^2}\right)\frac{1}{\chi^{1+\nu}}d\chi.$$

We use one more change of variable $\phi = |\xi|\chi$. In the case $|\xi| \geq 1$ we can further minorise by

$$\frac{4c_1}{\pi}\left(\int_0^{\pi/2}\min\left(\frac{\phi^2}{2}, \frac{\pi^2}{\epsilon^2}\right)\frac{1}{\phi^{1+\nu}}d\phi\right)|\xi|^\nu.$$

In the case $|\xi| \leq 1$, we explicitly obtain

$$\frac{2c_1}{\pi}\left(\int_0^{|\xi|\pi/2} \phi^{1-\nu} d\phi\right)|\xi|^\nu = C|\xi|^2.$$

$\square$

Bibliography

[1] D. Alexander, I. Kim, and Y. Yao. Quasi-static evolution and congested crowd transport. *Nonlinearity*, 27(4):823–858, 2014. 15

[2] R. Alexandre, L. Desvillettes, C. Villani, and B. Wennberg. Entropy dissipation and long-range interactions. *Arch. Ration. Mech. Anal.*, 152(4):327–355, 2000. 2, 20, 71, 75, 76, 80, 139, 153, 160

[3] R. Alexandre, J. Liao, and C. Lin. Some a priori estimates for the homogeneous Landau equation with soft potentials. *Kinet. Relat. Models*, 8(4):617–650, 2015. 16, 18

[4] R. Alexandre and C. Villani. On the Boltzmann equation for long-range interactions. *Comm. Pure Appl. Math.*, 55(1):30–70, 2002. 20, 70, 71, 72, 75, 76, 91, 153

[5] R. Alexandre and C. Villani. On the Landau approximation in plasma physics. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 21(1):61–95, 2004. 17, 19, 20, 72

[6] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008. 4, 12, 22, 26, 27, 31, 32, 38, 39, 42, 54, 56, 70, 87, 91, 117

[7] L. Arkeryd. On the Boltzmann equation. I. Existence. *Arch. Rational Mech. Anal.*, 45:1–16, 1972. 20

[8] L. Arkeryd. On the Boltzmann equation. II. The full initial value problem. *Arch. Rational Mech. Anal.*, 45:17–34, 1972. 20

[9] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré, and M. Veneroni. Passing to the limit in a Wasserstein gradient flow: from diffusion to reaction. *Calc. Var. Partial Differential Equations*, 44(3-4):419–454, 2012. 15

[10] A. A. Arsen'ev and O. E. Buryak. On a connection between the solution of the Boltzmann equation and the solution of the Landau-Fokker-Planck equation. *Mat. Sb.*, 181(4):435–446, 1990. 20

[11] A. A. Arsen'ev and N. V. Peskov. The existence of a generalized solution of Landau's equation. *Ž. Vyčisl. Mat i Mat. Fiz.*, 17(4):1063–1068, 1096, 1977. 29, 135

[12] R. Bailo, J. A. Carrillo, and J. Hu. Fully discrete positivity-preserving and energy-dissipating schemes for aggregation-diffusion equations with a gradient-flow structure. *Commun. Math. Sci.*, 18(5):1259–1303, 2020. 12, 22

[13] R. Bailo, J. A. Carrillo, H. Murakawa, and M. Schmidtchen. Convergence of a fully discrete and energy-dissipating finite-volume scheme for aggregation-diffusion equations. *Math. Models Methods Appl. Sci.*, 30(13):2487–2522, 2020. 12, 22

[14] J. Barnes and P. Hut. A hierarchical $o(n \log n)$ force-calculation algorithm. *Nature*, 324(6096):446, 1986. 133

[15] J. Barros-Neto. *An introduction to the theory of distributions.* Marcel Dekker, Inc. New York, 1973. Pure and Applied Mathematics, 14. 141

[16] G. Bellettini, L. Bertini, M. Mariani, and M. Novaga. Convergence of the one-dimensional Cahn-Hilliard equation. *SIAM J. Math. Anal.*, 44(5):3458–3480, 2012. 15

[17] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000. vi, 4, 5, 12, 30, 138

[18] C. K. Birdsall and A. B. Langdon. *Plasma Physics via Computer Simulation.* McGraw-Hill, 1985. 22

[19] T. Blass, I. Fonseca, G. Leoni, and M. Morandotti. Dynamics for systems of screw dislocations. *SIAM J. Appl. Math.*, 75(2):393–419, 2015. 15

[20] A. V. Boblylev, M. Pulvirenti, and C. Saffirio. From particle systems to the Landau equation: a consistency result. *Comm. Math. Phys.*, 319(3):683–702, 2013. 16

[21] A. V. Bobylëv. The theory of the nonlinear spatially uniform Boltzmann equation for Maxwell molecules. In *Mathematical physics reviews, Vol. 7*, volume 7 of *Soviet Sci. Rev. Sect. C Math. Phys. Rev.*, pages 111–233. Harwood Academic Publ., Chur, 1988. 82

[22] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Universitext. Springer, New York, 2011. 32

[23] C. Buet, S. Cordier, P. Degond, and M. Lemou. Fast algorithms for numerical, conservative, and entropy approximations of the Fokker-Planck-Landau equation. *J. Comput. Phys.*, 133(2):310–322, 1997. 22

[24] M. Burger and A. Esposito. Porous medium equation as limit of nonlocal interaction. *arXiv preprint arXiv:2202.05030*, 2022. 140

[25] G. Buttazzo. *Semicontinuity, relaxation and integral representation in the calculus of variations*, volume 207 of *Pitman Research Notes in Mathematics Series.* Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989. 13, 36

[26] 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):515–539, 2011. 148, 150, 152

[27] 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):575–608, 1992. 47

[28] 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):381–420, 2017. 16, 18

[29] J. A. Carrillo, Y.-P. Choi, and M. Hauray. The derivation of swarming models: mean-field limit and Wasserstein distances. In *Collective dynamics from bacteria to crowds*, volume 553 of *CISM Courses and Lect.*, pages 1–46. Springer, Vienna, 2014. 152

[30] J. A. Carrillo, K. Craig, and F. S. Patacchini. A blob method for diffusion. *Calc. Var. Partial Differential Equations*, 58(2):Paper No. 53, 53, 2019. 12, 15, 21, 23, 114, 122, 127, 128, 140

[31] J. A. Carrillo, M. G. Delgadino, L. Desvillettes, and J. Wu. The Landau equation as a Gradient Flow. *arXiv preprint arXiv:2007.08591*, 2020. ii, 4, 24, 141, 147

[32] J. A. Carrillo, M. G. Delgadino, and J. Wu. Boltzmann to Landau from the Gradient Flow Perspective. *Nonlinear Analysis*, 219:112824, 2022. ii, 69, 141, 153

[33] J. A. Carrillo, J. Hu, L. Wang, and J. Wu. A particle method for the homogeneous Landau equation. *J. Comput. Phys. X*, 7:100066, 24, 2020. ii, 113, 133

[34] J. A. Carrillo, S. Lisini, G. Savaré, and D. Slepčev. Nonlinear mobility continuity equations and generalized displacement convexity. *J. Funct. Anal.*, 258(4):1273–1309, 2010. 12

[35] J. A. Carrillo, R. J. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006. 4, 12, 13, 22, 138

[36] C. Cercignani. *The Boltzmann equation and its applications*, volume 67 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1988. 1

[37] C. Cercignani, R. Illner, and M. Pulvirenti. *The mathematical theory of dilute gases*, volume 106 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1994. 1, 20

[38] A. Chertock. A practical guide to deterministic particle methods. In *Handbook of numerical methods for hyperbolic problems*, volume 18 of *Handb. Numer. Anal.*, pages 177–202. Elsevier/North-Holland, Amsterdam, 2017. 21

[39] R. Choksi and M. A. Peletier. Small volume-fraction limit of the diblock copolymer problem: II. Diffuse-interface functional. *SIAM J. Math. Anal.*, 43(2):739–763, 2011. 15

[40] A. Cohen and B. Perthame. Optimal approximations of transport equations by particle and pseudoparticle methods. *SIAM J. Math. Anal.*, 32(3):616–636, 2000. 21, 128

[41] G.-H. Cottet and P.-A. Raviart. Particle methods for the one-dimensional Vlasov-Poisson equations. *SIAM J. Numer. Anal.*, 21(1):52–76, 1984. 21

[42] T. M. Cover and J. A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006. 3

[43] K. Craig and A. L. Bertozzi. A blob method for the aggregation equation. *Math. Comp.*, 85(300):1681–1717, 2016. 21, 128

[44] K. Craig, K. Elamvazhuthi, M. Haberland, and O. Turanova. A blob method method for inhomogeneous diffusion with applications to multi-agent control and sampling. *arXiv preprint arXiv:2202.12927*, 2022. 23, 140

[45] 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):167–182, 1992. 3, 9, 16, 20

[46] P. Degond and B. Lucquin-Desreux. An entropy scheme for the Fokker-Planck collision operator of plasma kinetic theory. *Numer. Math.*, 68(2):239–262, 1994. 22

[47] P. Degond and S. Mas-Gallic. The weighted particle method for convection-diffusion equations. I. The case of an isotropic viscosity. *Math. Comp.*, 53(188):485–507, 1989. 21

[48] P. Degond and S. Mas-Gallic. The weighted particle method for convection-diffusion equations. II. The anisotropic case. *Math. Comp.*, 53(188):509–525, 1989. 21

[49] P. Degond and F.-J. Mustieles. A deterministic approximation of diffusion equations using particles. *SIAM J. Sci. Statist. Comput.*, 11(2):293–310, 1990. 21

[50] M. G. Delgadino. Convergence of a one-dimensional Cahn-Hilliard equation with degenerate mobility. *SIAM J. Math. Anal.*, 50(4):4457–4482, 2018. 15

[51] L. Desvillettes. On asymptotics of the Boltzmann equation when the collisions become grazing. *Transport Theory and Statistical Physics*, 21(3):259–276, 1992. 3, 9, 16, 20

[52] L. Desvillettes. Entropy dissipation estimates for the Landau equation in the Coulomb case and applications. *J. Funct. Anal.*, 269(5):1359–1403, 2015. 16, 19, 61

[53] L. Desvillettes. Entropy dissipation estimates for the Landau equation: general cross sections. In *From particle systems to partial differential equations. III*, volume 162 of *Springer Proc. Math. Stat.*, pages 121–143. Springer, [Cham], 2016. 19, 30, 61

[54] L. Desvillettes. Autour du thèorème H de Boltzmann. *To appear at Séminaire Laurent Schwartz, EDP et Applications*, 2020. 19, 29, 61, 135

[55] 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. *arXiv preprint arXiv:2011.00386*, 2020. 17, 29, 135

[56] 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):179–259, 2000. 16, 18, 30

[57] 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):261–298, 2000. 16, 30

[58] R. J. DiPerna and P.-L. Lions. On the Cauchy problem for Boltzmann equations: global existence and weak stability. *Ann. of Math. (2)*, 130(2):321–366, 1989. 19, 20, 91

[59] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989. 19, 20

[60] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009. vi, 12, 13, 30, 37, 38, 40, 42, 138

[61] M. Erbar. A gradient flow approach to the Boltzmann equation. *arXiv preprint arXiv:1603.00540*, 2019. iv, vi, 4, 10, 20, 22, 23, 24, 30, 36, 37, 38, 41, 54, 80, 87, 88, 91, 108, 117, 118, 138, 139

[62] I. Fonseca and G. Leoni. *Modern methods in the calculus of variations: L^p spaces.* Springer Monographs in Mathematics. Springer, New York, 2007. 13, 36

[63] N. Fournier. Uniqueness of bounded solutions for the homogeneous Landau equation with a Coulomb potential. *Comm. Math. Phys.*, 299(3):765–782, 2010. 16

[64] N. Fournier and H. Guérin. On the uniqueness for the spatially homogeneous Boltzmann equation with a strong angular singularity. *J. Stat. Phys.*, 131(4):749–781, 2008. 20, 76

[65] N. Fournier and H. Guérin. Well-posedness of the spatially homogeneous Landau equation for soft potentials. *J. Funct. Anal.*, 256(8):2542–2560, 2009. 4, 16

[66] N. Fournier and C. Mouhot. On the well-posedness of the spatially homogeneous Boltzmann equation with a moderate angular singularity. *Comm. Math. Phys.*, 289(3):803–824, 2009. 20, 76

[67] D. Godinho. Asymptotic of grazing collisions for the spatially homogeneous Boltzmann equation for soft and Coulomb potentials. *Stochastic Process. Appl.*, 123(11):3987–4039, 2013. 20, 22, 75

[68] F. Golse. On the dynamics of large particle systems in the mean field limit. In *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*, volume 3 of *Lect. Notes Appl. Math. Mech.*, pages 1–144. Springer, [Cham], 2016. 152

[69] F. Golse, M. P. Gualdani, C. Imbert, and A. Vasseur. Partial regularity in time for the space homogeneous Landau equation with Coulomb potential. *arXiv preprint arXiv:1906.02841*, 2019. 4, 17

[70] 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):253–295, 2019. 4, 19

[71] F. Golse, P.-L. Lions, B. Perthame, and R. Sentis. Regularity of the moments of the solution of a transport equation. *J. Funct. Anal.*, 76(1):110–125, 1988. 20, 91

[72] M. Gualdani and N. Zamponi. A review for an isotropic Landau model. In *PDE models for multi-agent phenomena*, volume 28 of *Springer INdAM Ser.*, pages 115–144. Springer, Cham, 2018. 16, 138

[73] M. P. Gualdani and N. Guillen. Estimates for radial solutions of the homogeneous Landau equation with Coulomb potential. *Anal. PDE*, 9(8):1772–1809, 2016. 4, 16, 138

[74] M. P. Gualdani and N. Zamponi. Spectral gap and exponential convergence to equilibrium for a multi-species Landau system. *Bull. Sci. Math.*, 141(6):509–538, 2017. 11, 119

[75] M. P. Gualdani and N. Zamponi. Global existence of weak even solutions for an isotropic Landau equation with Coulomb potential. *SIAM J. Math. Anal.*, 50(4):3676–3714, 2018. 16, 138

[76] J. Guerand and C. Imbert. Log-transform and the weak Harnack inequality for kinetic Fokker-Planck equations. *arXiv preprint arXiv:2102.04105*, 2021. 19

[77] J. Guerand and C. Mouhot. Quantitative De Giorgi methods in kinetic theory. *arXiv preprint arXiv:2103.09646*, 2021. 19

[78] Y. Guo. The Landau equation in a periodic box. *Comm. Math. Phys.*, 231(3):391–434, 2002. 4, 16, 29, 135

[79] L. He. Asymptotic analysis of the spatially homogeneous Boltzmann equation: grazing collisions limit. *J. Stat. Phys.*, 155(1):151–210, 2014. 20

[80] L.-B. He. Sharp bounds for Boltzmann and Landau collision operators. *Ann. Sci. Éc. Norm. Supér. (4)*, 51(5):1253–1341, 2018. 139

[81] E. Hebey. *Nonlinear analysis on manifolds: Sobolev spaces and inequalities*, volume 5 of *Courant Lecture Notes in Mathematics.* New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999. 104

[82] A. M. Helmenstine. What is plasma used for, and what is it made of? 2019. Last accessed 11 March 2022, `https://www.thoughtco.com/what-is-plasma-608345`. 1

[83] C. Henderson and S. Snelson. C^∞ smoothing for weak solutions of the inhomogeneous Landau equation. *Arch. Ration. Mech. Anal.*, 236(1):113–143, 2020. 19

[84] D. Hilbert. Mathematical problems. *Bull. Amer. Math. Soc. (N.S.)*, 37(4):407–436, 2000. Reprinted from Bull. Amer. Math. Soc. **8** (1902), 437–479. 2

[85] R. W. Hockney and J. W. Eastwood. *Computer Simulation Using Particles.* CRC Press, 1988. 22

[86] C. Imbert and L. Silvestre. The Schauder estimate for kinetic integral equations. *Anal. PDE*, 14(1):171–204, 2021. 19, 139

[87] ITER organisation. Making it work. 2022. Last accessed 11 March 2022, `https://www.iter.org/sci/makingitwork`. 1

[88] ITER organisation. What is ITER? 2022. Last accessed 11 March 2022, `https://www.iter.org/proj/inafewlines`. 1

[89] P.-E. Jabin and S. Junca. A continuous model for ratings. *SIAM J. Appl. Math.*, 75(2):420–442, 2015. 2

[90] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998. 4, 5, 12, 15, 22, 24, 53, 136

[91] J. Jost. *Riemannian geometry and geometric analysis.* Universitext. Springer, Cham, seventh edition, 2017. 144

[92] G. Lacombe and S. Mas-Gallic. Presentation and analysis of a diffusion-velocity method. In *Flows and related numerical methods (Toulouse, 1998)*, volume 7 of *ESAIM Proc.*, pages 225–233. Soc. Math. Appl. Indust., Paris, 1999. 21

[93] M. Lemou. Multipole expansions for the Fokker-Planck-Landau operator. *Numer. Math.*, 78(4):597–618, 1998. 22

[94] P. Li, H. Johnston, and R. Krasny. A Cartesian treecode for screened Coulomb interactions. *J. Comput. Phys.*, 228(10):3858–3868, 2009. 133

[95] E. M. Lifshitz and L. P. Pitaevskiĭ. *Course of theoretical physics ["Landau-Lifshits"]. Vol. 10.* Pergamon International Library of Science, Technology, Engineering and Social Studies. Pergamon Press, Oxford-Elmsford, N.Y., 1981. Translated from the Russian by J. B. Sykes and R. N. Franklin. 1, 3, 16

[96] P.-L. Lions. Régularité et compacité pour des noyaux de collision de Boltzmann sans troncature angulaire. *C. R. Acad. Sci. Paris Sér. I Math.*, 326(1):37–41, 1998. 91

[97] P.-L. Lions and S. Mas-Gallic. Une méthode particulaire déterministe pour des équations diffusives non linéaires. *C. R. Acad. Sci. Paris Sér. I Math.*, 332(4):369–376, 2001. 21

[98] N. F. G. Martin and J. W. England. *Mathematical theory of entropy*, volume 12 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley Publishing Co., Reading, Mass., 1981. With a foreword by James K. Brooks. 3

[99] S. Mas-Gallic. The diffusion velocity method: a deterministic way of moving the nodes for solving diffusion equations. volume 31, pages 595–605. 2002. Special issue comprised of papers presented at the Conference on Asymptotic and Numerical Methods for Kinetic Equations (Oberwolfach, 2001). 21

[100] D. Matthes and H. Osberger. A convergent Lagrangian discretization for a nonlinear fourth-order equation. *Found. Comput. Math.*, 17(1):73–126, 2017. 15

[101] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997. 138

[102] K. Oelschläger. Large systems of interacting particles and the porous medium equation. *J. Differential Equations*, 88(2):294–346, 1990. 21

[103] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001. 12, 41, 73

[104] L. Pareschi, G. Russo, and G. Toscani. Fast spectral methods for the Fokker-Planck-Landau collision operator. *J. Comput. Phys.*, 165(1):216–236, 2000. 21, 130, 133

[105] L. Pareschi and G. Toscani. *Interacting multiagent systems. Kinetic equations and Monte Carlo methods.* Oxford University Press, 2013. 2

[106] H. L. Royden. *Real analysis.* Macmillan Publishing Company, New York, third edition, 1988. 58

[107] G. Russo. Deterministic diffusion of particles. *Comm. Pure Appl. Math.*, 43(6):697–733, 1990. 21

[108] G. Russo. A particle method for collisional kinetic equations. I. Basic theory and one-dimensional results. *J. Comput. Phys.*, 87(2):270–300, 1990. 21

[109] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004. 15, 21, 74, 138

[110] J. Sándor. A note on the logarithmic mean. *Amer. Math. Monthly*, 123(1):112, 2016. 142

[111] F. Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications.* Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling. 4

[112] F. Santambrogio. {Euclidean, metric, and Wasserstein} gradient flows: an overview. *Bull. Math. Sci.*, 7(1):87–154, 2017. 13

[113] S. Serfaty. Gamma-convergence of gradient flows on Hilbert and metric spaces and applications. *Discrete Contin. Dyn. Syst.*, 31(4):1427–1451, 2011. vi, 15, 21, 74, 138

[114] G. Toscani. Kinetic models of opinion formation. *Commun. Math. Sci.*, 4(3):481–496, 2006. 2

[115] C. Truesdell and R. G. Muncaster. *Fundamentals of Maxwell's kinetic theory of a simple monatomic gas*, volume 83 of *Pure and Applied Mathematics*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980. Treated as a branch of rational mechanics. 1

[116] D. Tskhakaya, K. Matyash, R. Schneider, and F. Taccogna. The particle-in-cell method. *Contributions to Plasma Physics*, 47(8-9):563–594, 2007. 22

[117] C. Villani. On the Cauchy problem for Landau equation: sequential stability, global existence. *Adv. Differential Equations*, 1(5):793–816, 1996. 119

[118] C. Villani. On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Rational Mech. Anal.*, 143(3):273–307, 1998. 2, 3, 7, 16, 17, 18, 20, 26, 71, 72, 73, 75, 82, 83, 118, 135, 138

[119] C. Villani. On the spatially homogeneous Landau equation for Maxwellian molecules. *Math. Models Methods Appl. Sci.*, 8(6):957–983, 1998. 16, 135

[120] C. Villani. A review of mathematical topics in collisional kinetic theory. In *Handbook of mathematical fluid dynamics, Vol. I*, pages 71–305. North-Holland, Amsterdam, 2002. 2

[121] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new. 91, 148, 151

[122] K.-C. Wu. Global in time estimates for the spatially homogeneous Landau equation with soft potentials. *J. Funct. Anal.*, 266(5):3134–3155, 2014. 16, 18, 29, 135