

The Landau equation does not blow up

by

NESTOR GUILLEN

*Texas State University
 San Marcos, TX, U.S.A.*

LUIS SILVESTRE

*The University of Chicago
 Chicago, IL, U.S.A.*

1. Introduction

The Landau equation was derived in 1936 by Lev Landau [42] to model the evolution of densities of particles performing Coulomb collisions in a plasma. It can be derived as a limit case of the Boltzmann equation when the grazing collisions dominate the evolution. It is one of the central equations in kinetic theory. In the space-homogeneous case, it is given by a very simple formula

$$f_t = \bar{a}_{ij} \partial_{ij} f + f^2, \quad \text{where} \quad \bar{a}_{ij} = -\partial_{ij}(-\Delta)^{-2} f.$$

A more general family of equations is usually studied. It has the form

$$f_t = q(f), \tag{1.1}$$

where the operator $q(f)$ is given by the formula

$$q(f)(v) = \partial_{v_i} \int_{\mathbb{R}^3} \alpha(|v-w|) a_{ij}(v-w) (\partial_{v_j} - \partial_{w_j}) [f(v)f(w)] dw. \tag{1.2}$$

Here $\alpha(0, +\infty) \rightarrow [0, +\infty)$ is an arbitrary non-negative function and

$$a_{ij}(z) = |z|^2 \delta_{ij} - z_i z_j.$$

We will refer to α as the interaction potential. It is common to study the case $\alpha(r) = r^\gamma$ with $\gamma \in [-3, 1]$. The most important case is $\alpha(r) = r^{-3}$ that corresponds to the original Landau equation written above, for charged particles interacting with Coulomb potentials.

Luis Silvestre was supported by NSF grants DMS-2054888 and DMS-2350263. Nestor Guillen was supported by NSF grant DMS-2144232.

For a positive function $f: \mathbb{R}^3 \rightarrow (0, \infty)$, its Fisher information is defined by the following expression

$$i(f) := \int_{\mathbb{R}^3} \frac{|\nabla f|^2}{f} dv. \quad (1.3)$$

The formula (1.3) is extended to the case that f has vacuum regions by setting

$$\frac{|\nabla f|^2}{f} = 0$$

at those points.

Our main result is the following.

THEOREM 1.1. *Assume that $f: [0, T] \times \mathbb{R}^3 \rightarrow [0, \infty)$ is a classical solution to the space-homogeneous Landau equation (1.1). Assume that the interaction potential α satisfies, for all $r > 0$,*

$$\frac{r|\alpha'(r)|}{\alpha(r)} \leq \sqrt{19},$$

then the Fisher information $i(f)$ is monotone decreasing as a function of time.

We present Theorem 1.1 as an a-priori estimate for classical solutions. By that, we mean smooth functions f that decay sufficiently fast as $|v| \rightarrow \infty$. Our second main theorem tells us that the equation has global smooth solutions for every reasonably nice initial data.

The assumption of Theorem 1.1 holds for any power-law interaction $\alpha(r) = r^\gamma$, with γ in the usual range $\gamma \in [-3, 1]$. It even goes beyond that range if the Landau collision operator is understood appropriately taking into account the cancellations in the integrand of (1.2) for $|v-w|$ small. Our parameter $\sqrt{19}$ is not optimal. After we posted online a first version of the this article, Sehyun Ji improved the computation of the threshold of applicability of Theorem 1.1 to

$$\frac{r|\alpha'(r)|}{\alpha(r)} \leq \sqrt{22}$$

in [38]. There is no reason to expect this condition to be optimal either. We are currently not aware of any example, for any interaction potential, of a solution for which the Fisher information is not monotone decreasing.

As a consequence of the monotonicity of the Fisher information, we deduce the global existence of smooth solutions in the very-soft-potential range.

THEOREM 1.2. *Assume that $\alpha(r)=r^\gamma$, for $\gamma\in[-3,1]$. Let $f_0:\mathbb{R}^3\rightarrow[0,\infty)$ be an initial data that is bounded by a Maxwellian in the sense that*

$$f_0(v)\leq C_0 e^{-\beta|v|^2},$$

for some positive parameters C_0 and β .

Then, there is a unique global classical solution $f:[0,\infty)\times\mathbb{R}^3\rightarrow[0,\infty)$ to the Landau equation (1.1), with initial data $f(0,v)=f_0(v)$. For any positive time, this function f is strictly positive, in the Schwartz space, and bounded above by a Maxwellian. The Fisher information $i(f)$ is non-increasing.

The remarkable new feature of Theorem 1.2 is that it applies to the very-soft potential range $\gamma\leq -2$. The most important case of Theorem 1.2 is $\alpha(r)=r^{-3}$, which is the original equation by Landau and corresponds to the evolution of the velocity density of charged particles that interact by Coulomb potentials. Regularity estimates for this equation have remained an elusive well-known open problem for several years. The difficulty comes from the fact that the reaction term is too singular to be bested by the diffusion term when we only use the previously known controlled coercive quantities of mass, energy and entropy. The boundedness of the Fisher information, which is provided by Theorem 1.1, overcomes this difficulty altogether.

Remark 1.3. Theorem 1.2 is a relatively direct consequence of Theorem 1.1 combined with well-known techniques. There are various short-time well-posedness results in the literature that one can apply, as well as conditional a-priori estimates. A slightly different version of Theorem 1.2 would result from different choices between them. For example, it is possible to replace the Maxwellian upper bound for f_0 with the condition $f_0\in L_k^\infty$ for k sufficiently large, and it would lead to correspondingly weaker decay conditions for the solution f .

Using our Theorem 1.1, minimalist assumptions on the initial data for which the Cauchy problem is solvable are investigated in the recent preprints [17], [39].

Remark 1.4. A version of Theorem 1.1 for the homogeneous Boltzmann equation is obtained in [36]. In that recent preprint, Cyril Imbert, Cedric Villani and the second author reproduce much of the analysis of this paper in the context of the Boltzmann equation. However, a new difficulty emerges that requires the development of a highly non-trivial integro-differential version of the inequality of §9.

1.1. Historical remarks and references

Given that the Landau equation is one of the main equations in statistical mechanics, especially in the Coulombic case $\alpha(r)=r^{-3}$, it is no surprise that it has received a significant amount of attention from the mathematical community through the years. It is impossible to review all the literature in the subject. We discuss some of the most relevant results below.

The analysis of the Boltzmann equation, as well as the Landau equation, involves significantly simpler formulas in the case of the Maxwell-molecules model (which is $\alpha\equiv 1$) than in the case of general interaction potentials. One of the reasons is that there is a manageable expression for the Fourier transform of the operator due to Alexander Bobylev (see [6] and [7]). It is also well known that the evolution of the equation is contractive with respect to the quadratic Wasserstein distance (see [49] and [10]), and that the Fisher information is non-increasing along the flow (see [52] and [55]). We naturally wonder if these properties fail to be true in the case of other interaction potentials, or if they hold in more generality but proofs are only well understood in the case of Maxwell molecules due to the simpler arithmetic structure of the operator. Is the Fisher information monotone decreasing for the Boltzmann and Landau equation for some interaction potential other than Maxwell molecules? What about the entropy dissipation? What about the contractivity of the Wasserstein distance? We answer the first of these questions for the Landau equation in Theorem 1.1.

It is rather unusual to find a new explicit Lyapunov functional for a very well studied equation in mathematical physics. Moreover, most basic conserved and monotone quantities in PDEs are verified by a relatively simple computation. It is also unusual to find a simple Lyapunov functional whose proof is non-trivial.

The first reference to the monotonicity of the Fisher information in a kinetic equation is in an interesting paper by McKean in 1966 [44]. He considers Kac's 1-dimensional caricature of the Boltzmann equation with Maxwell molecules. He proves that the Fisher information is monotone in this context. The paper includes a number of opinions and conjectures including the following.

- The author writes that the Fisher information “probably fails to decrease” in the 3-dimensional problem.
- He conjectures (in the context of the Kac equation) that the functionals that result as higher-order derivatives of the entropy by heat flow are all monotone. He reports that he tried to prove it but could not do it.
- He conjectures that the entropy dissipation for the Boltzmann equation should be monotone decreasing. He even suggests that its derivative may be monotone as well.

In 1992, Giuseppe Toscani proved the monotonicity of the Fisher information for the 2-dimensional Boltzmann equation in the Maxwell molecules case [50]. Toscani conjectures that the Fisher information can probably be proved to be monotone “at least in the case of inverse power interaction potentials”, but he states that it is still an open question.

In 1998, Villani proved the monotonicity of the Fisher information for the Boltzmann equation with Maxwell molecules in arbitrary dimension [52]. The proof, which one may say uses similar ideas as in Toscani’s paper, involves more complicated geometry and formulas. In 2012, Matthes and Toscani provide a very short alternative proof based on Fourier analysis [43]. This proof is very specific to $\alpha \equiv 1$ and provides no intuition about other interaction potentials. In the introduction of [52], Villani writes “We also investigate briefly the case of arbitrary potentials, and show precisely why the Maxwellian case seems to depart from the other”, which reflects the limitations of methods available then when it came to general power law potentials.

The monotonicity of the Fisher information for the Boltzmann equation implies the monotonicity for the Landau equation as a limit case. In 2000, Villani wrote a direct proof in the case of the Landau equation with Maxwell molecules [55]. While the paper [55] is fairly short, the proof is non-trivial. The intuition is arguably harder to grasp than in the Boltzmann case. In §5, we provide an alternative proof for the Landau equation in the Maxwell molecules case, which is in some ways the starting point for our general method for Theorem 1.1.

Other recent publications studying the evolution of the Fisher information are [2] for the hard-potential case and [45] for moderately soft potentials. They present upper bounds (not monotonicity) for the Fisher information that are uniform in time. These are scenarios in which there are well-known global-in-time regularity estimates.

The great majority of the regularity estimates for the Landau equation in the past use a decomposition for the Landau collision operator (1.2) as a sum of a diffusion term plus a lower-order term. In non-divergence form, it reads

$$q(f) = \bar{a}_{ij} \partial_{ij} f + \bar{c} f,$$

where

$$\bar{a}_{ij} = \int_{\mathbb{R}^3} \alpha(|v-w|) a_{ij}(v-w) f(w),$$

and $\bar{c} = -\partial_{ij} \bar{a}_{ij}$. Some ellipticity bounds can be deduced for the coefficients \bar{a}_{ij} based only on the mass, energy and entropy of f . The reaction term $\bar{c} f$ is more singular when γ is more negative. The majority of the estimates in the literature are obtained using parabolic estimates for the diffusion term and using them to control the other term. We

do not use this decomposition in our proof of Theorem 1.1. When $\gamma \geq 0$, the reaction term is very simple because \bar{c} is bounded pointwise in terms of the mass and energy of f . When $\gamma \in [-2, 0]$, the reaction term can still be controlled with the help of ellipticity estimates from the diffusion term. For the very-soft potential range $\gamma < -2$, the reaction term is too singular to be controlled from the diffusion term. This is vaguely the reason of the main difficulty in establishing unconditional bounds in the very-soft potential range.

The early results on classical well-posedness for the Landau equation focused on the Maxwell-molecules case ($\gamma = 0$) or the hard-potentials case ($\gamma > 0$). Villani investigated the Maxwell-molecules case first [54], and later the case of hard potentials in collaboration with Laurent Desvillettes [19], [20]. The regularity in the hard-potentials case was revisited by El Safadi in [21]. Our understanding of the global well-posedness and smoothness for $\gamma \geq 0$ is very satisfactory. For moderately soft potentials, there are regularity estimates in [58], [46], [31], from which one can construct global smooth solutions, as in Theorem 1.2.

For very soft potentials, which is the range $\gamma \in [-3, -2)$, the global classical well-posedness of the equation has been an elusive and well-known open question for several years. This is the most interesting range, because it includes the original Landau equation for Coulomb potentials which is $\gamma = -3$. The development of our understanding before this work was comparable with our current understanding of the Navier–Stokes equation (in terms of what the known results are). The results in the current literature can be roughly classified in the following groups.

Global-in-time weak solutions. In 1998, Villani defined a notion of generalized solution for the space-homogeneous Boltzmann and Landau equation [53], which he called H -solution. He was able to prove the existence of global solutions of this kind, but not the uniqueness. In some way, Villani’s result for the Landau equation plays the same role as Leray’s global weak solutions do for the Navier–Stokes equation. See also [3] for an earlier notion of generalized solution.

From an entropy dissipation estimate, Desvillettes proved that the H -solutions constructed by Villani belong to $L^1([0, T], L^3_{-3}(\mathbb{R}^3))$ and are in fact weak solutions in a more classical sense [16] (see also [37]).

François Golse, Maria Gualdani, Cyril Imbert and Alexis Vasseur showed that the global weak solutions constructed by Villani are smooth outside of a potential set of times of dimension at most $\frac{1}{2}$ (see [27] and [26]).

In a recent preprint [28], Golse, Imbert and Vasseur prove that the set of potential singularities in (t, v) has parabolic Hausdorff dimension at most $\frac{7}{2}$ in the case of Coulombic potential. This result is more or less comparable with the Caffarelli–Kohn–Nirenberg partial regularity theorem for Navier–Stokes, except that the dimension they obtain is

larger in the case of the Landau equation. In [28], they also prove that axially symmetric solutions of the Landau equation with very soft potentials are smooth away from the axis of symmetry.

Short-time existence. Short time existence results are obtained by Nicolas Fournier and Hélène Guérin [23], by Chris Henderson, Stanley Snelson and Andrei Tarfulea [34], by Hyung Ju Hwang and Jin Woo Jang [35], and by William Golding and Amélie Loher [25].

Perturbative theory around the Maxwellian. When the initial data is sufficiently close to a fixed Maxwellian in an appropriate sense, Yan Guo constructed a global smooth solution [33] (even for the space-inhomogeneous case). See also [24] for a more recent result with weaker assumptions on the initial data but restricted to the space-homogeneous case.

Partial progress aiming at the regularity of solutions. Conditional regularity results are obtained by the second author in [46], and by Gualdani and the first author in [31]. These works imply in particular that if the solution to (1.1) stays bounded in certain L^p space, for an appropriately large exponent p , then the solution will be smooth.

Recent results by Ricardo Alonso, Véronique Bagland, Laurent Desvillettes and Bertrand Lods [1] and by William Golding and Amélie Loher [25], provide Prodi–Serrin-type conditional regularity estimates depending on $f \in L^p([0, T], L^q(\mathbb{R}^3))$ for suitable pairs of exponents p and q .

Other partial results regarding the regularity of the space-homogeneous Landau equation with Coulomb potentials are given in [9], [5], [13], [47], [18].

Modified equations. In [41], Joachim Krieger and Robert Strain proposed an isotropic toy model that retains some of the features of the original Landau equation for the case of Coulomb potentials, and where blow up could be ruled out (for radial solutions). This model and other variations is further analyzed in [29], [30], [32], [48]. Other simplified radially symmetric models are analyzed by Alexander Bobylev in [8].

The difficulty of obtaining unconditional regularity estimates for the Landau equation with Coulomb potentials (resolved in Theorem 1.2) is mentioned as an outstanding open problem in the majority of these papers.

The problem of establishing the regularity of the Landau equation with Coulomb potentials is also described in the open problems section of Villani’s book [56, §5.1.3]. Villani first argues that finite-time blow-up may be expected from an analogy between the Landau equation and the non-linear heat equation. However, he reports that after seeing some numerical simulations by Francis Filbet, he changed his mind and became convinced that blow-up should not occur. See also the related numerical computations in [12]. That a blow up does not take place is finally confirmed by our Theorem 1.2.

In [57], Villani provides further details about the history of these problems, as well as a description of the ideas leading to the proof of the monotonicity of the Fisher information for the Boltzmann equation in [36].

1.2. Outline of the paper and main ideas in the proof.

In this section we attempt to explain a rough outline of the ideas involved in the proof of Theorem 1.1 and the general organization of the paper.

In §2, we review some of the results in the literature that we apply in this paper. None of the results presented in this preliminary section is new.

The first idea leading to Theorem 1.1 is described in §3. We observe that we can write the Landau operator as a composition of a linear elliptic operator in \mathbb{R}^6 , and the projection of a distribution function F in \mathbb{R}^6 over its marginal. More precisely, for a function $F: \mathbb{R}^6 \rightarrow [0, \infty)$ with unit integral, we write

$$\pi F(v) := \int_{\mathbb{R}^3} F(v, w) dw.$$

We also define the degenerate elliptic operator Q acting on functions $F=F(v, w)$ as

$$Q(F)(v, w) = (\partial_{v_i} - \partial_{w_i})(\alpha(|v-w|)a_{ij}(v-w)(\partial_{v_j} - \partial_{w_j})F).$$

With this notation, we observe that

$$q(f) = \pi Q(f \otimes f), \quad \text{where} \quad (f \otimes f)(v, w) := f(v)f(w).$$

It turns out that the projection over marginals reduces the value of the Fisher information, in the sense that, for any symmetric probability distribution F on \mathbb{R}^6 ,

$$i(f) = \frac{1}{2} I(f \otimes f) \leq I(F), \quad \text{when} \quad f = \pi F.$$

Here, we write $I(F)$ for the Fisher information applied to a function in \mathbb{R}^6 . From this inequality we argue that the Fisher information will decrease along the direction of $q(f)$ if $I(F)$ decreases along the direction of $Q(F)$ for any function $F=F(v, w)$ defined on \mathbb{R}^6 such that $F(v, w)=F(w, v)$.

This is a substantial simplification. The Landau operator, q , which is non-linear and non-local, is now replaced by a linear elliptic operator with explicit coefficients, Q . Within this framework the known monotonicity of the Fisher information for Maxwell molecules easily follows (see below). For all other potentials, estimating the derivative of I along Q is still non-trivial and requires further ideas.

In §4 we observe that, for certain vector fields \tilde{b}_1, \tilde{b}_2 and \tilde{b}_3 , we have

$$Q(F) = \sum_{k=1}^3 \sqrt{\alpha} \tilde{b}_k \cdot \nabla (\sqrt{\alpha} \tilde{b}_k \cdot \nabla F).$$

This observation, while elementary, will help us organize our formulas efficiently. We use these vectors and their properties extensively in the rest of the paper.

The transport of each of these vector fields $\tilde{b}_k \cdot \nabla$ individually goes by isometries in \mathbb{R}^6 . Thus, $I(F)$ is constant along the flow of $\tilde{b}_k \cdot \nabla$. The fact that $I(F)$ is monotone decreasing in the direction of $Q(F)$ in the case of Maxwell molecules follows immediately from this and the convexity of the Fisher information I . This is explained in §5.

Whenever α is not constant, the flow of $\sqrt{\alpha} \tilde{b}_k \cdot \nabla$ is not an isometry in \mathbb{R}^6 . However, it is an isometry when we restrict it to the level sets of α . With this idea in mind, in §6, we define a modified Fisher information I_{\tan} that only takes into account the gradient of F along the directions parallel to the level sets of $|v-w|$. This functional I_{\tan} is constant along the flow of $\sqrt{\alpha} \tilde{b}_k \cdot \nabla$. Using the same convexity argument as in the case of Maxwell molecules, we conclude that

$$\langle I'_{\tan}(F), Q(F) \rangle \leq 0,$$

with a precise estimate of its dissipation in terms of second derivatives of F .

The difference between the full Fisher information $I(F)$ and the modified one $I_{\tan}(F)$ is only one direction, which we call n , and is the normalized gradient of $|v-w|$ in \mathbb{R}^6 . We write

$$J(F) := I(F) - I_{\tan}(F) = \iint_{\mathbb{R}^6} \frac{|n \cdot \nabla F|^2}{F} dw dv.$$

In §7, we compute $\langle J'(F), Q(F) \rangle$ to express it in a form that resembles as closely as possible the structure of the dissipation of I_{\tan} . We do it through a direct computation using the vector fields \tilde{b}_k and Lie brackets. We end up with a negative diffusion term involving second derivatives of F and a positive error term of the form

$$R := \iint_{\mathbb{R}^6} \frac{|\alpha'|^2}{\alpha} (\tilde{b}_k \cdot \nabla \log F)^2 F dw dv.$$

In §8, we break the favorable diffusion terms obtained in the previous sections into three groups, depending on the directions of their second derivatives. We isolate one term, which we call $D_{\text{spherical}}$, that only takes into account second derivatives in the directions of the vector fields \tilde{b}_k 's. We claim that this diffusion term can be used to control the error term R . After using polar coordinates and performing some manipulations, we reduce

the problem to an inequality for functions on the sphere $f: S^2 \rightarrow (0, \infty)$ of the following form:

$$\int_{S^2} |\nabla_\sigma^2 \log f|^2 f \, d\sigma \geq \lambda \int_{S^2} |\nabla_\sigma \log f|^2 f \, d\sigma.$$

It is not difficult to prove that the inequality holds for some $\lambda > 0$. However, it is important for us to obtain a nearly sharp constant λ . The maximum value that we are able to handle for $r|\alpha'(r)|/\alpha(r)$ in Theorem 1.1 (which is $\sqrt{19}$) depends on the value of λ from this inequality. We prove the inequality in §9 (written slightly differently in terms of the vector fields b_k). It appears to be new as far as we are aware.

Finally, in §11, we explain how Theorem 1.1, combined with results from the literature, can be used to derive Theorem 1.2. There is no difficulty in §11.

Acknowledgements

We would like to thank David Bowman and Sehyun Ji for reading an earlier version of this manuscript and pointing out some typos. We also thank Maria Gualdani and Cyril Imbert for many useful discussions about this problem through several years. We would also like to thank the anonymous reviewer for many helpful comments and suggestions.

2. Preliminaries

In this section we collect the results from the literature that we use in this paper. The results we state here are well known. They either appear explicitly or are straightforward minor modifications of the results in the literature. We provide references in every case.

As it is customary, we write $\langle v \rangle = (1 + |v|^2)^{1/2}$ and, for any $p \in [1, \infty]$ and any $k \in \mathbb{R}$,

$$\|f\|_{L_k^p(\mathbb{R}^3)} := \|\langle v \rangle^k f\|_{L^p(\mathbb{R}^3)}.$$

The mass, momentum and energy, that is

$$\begin{aligned} \text{(mass)} &:= \int_{\mathbb{R}^3} f(v) \, dv, \\ \text{(momentum)} &:= \int_{\mathbb{R}^3} f(v)v \, dv, \\ \text{(energy)} &:= \int_{\mathbb{R}^3} f(v)|v|^2 \, dv. \end{aligned}$$

are constant along the flow of the equation.

To fix ideas, it is a comfortable choice to study solutions of the equation (1.1) with unit mass and zero momentum. This will be preserved along the flow of the equation.

We recall the following formula for the entropy:

$$h(t) := \int_{\mathbb{R}^3} f(v) \log f(v) dv.$$

The entropy is non-increasing for solutions of (1.1). Before this work, it was the only natural quantity associated to the equation known to be monotone decreasing—outside the Maxwell-molecules case $\alpha \equiv 1$.

To begin the analysis of the equation, we start from a short-time existence result. There exist a few in the literature, both in the space-homogeneous and space-inhomogeneous regime. The most complete that we know of is given in [34]. Restricted to the space-homogeneous regime, it becomes the following result.

THEOREM 2.1. (from [34]) *Assume that $\alpha(r) = r^\gamma$ for some $\gamma \in [-3, 0)$. Assume that $f_0: \mathbb{R}^3 \rightarrow [0, \infty)$ is an initial value such that $f_0 \in L_{k_0}^\infty$, with $k_0 = \max\{5, 15/(5+\gamma)\}$. There exists a positive time $T > 0$ and a classical solution $f: [0, T] \times \mathbb{R}^3 \rightarrow [0, \infty)$ to the equation (1.1) with initial data f_0 . This solution satisfies the following properties:*

- $f(t, v) > 0$ for all $t > 0$ and $v \in \mathbb{R}^3$.
- $f \in C^\infty((0, T) \times \mathbb{R}^3)$ with upper bounds for $t > 0$ for all derivatives.
- If f_0 is continuous, then f matches the initial data continuously. Otherwise, it is weakly continuous at $t=0$.
- The solution can be extended for as long as $\|f(t)\|_{L_{k_0}^\infty} < +\infty$.
- Let T_E be the maximum time of existence and $\tau > 0$. For all $m \in \mathbb{N}$, there exist k_m and ℓ_m such that, if $f_0 \in L_{k_m}^\infty$, then $D^{(m)} f(t) \in L^\infty([0, T_E - \tau], L_{k_m - \ell_m}^\infty(\mathbb{R}^3))$.

The main result in [34] is for the space-inhomogeneous equation. There is an extra assumption for the initial data f_0 , which they call *well distributed*. This assumption is only relevant for the space-inhomogeneous case. In the space-homogeneous case, it would be automatically satisfied if f_0 is continuous and non-zero. Even if f_0 is not continuous, it is not difficult to see that $f(t, v)$ will be well distributed for any $t > 0$ small.

There is also a more recent short-time existence result [25], which is specific to the space-homogeneous Landau equation with Coulomb potentials, for initial data that is merely in $L_k^1 \cap L^p$ for some $p > \frac{3}{2}$. The solution constructed in [25] is also classical and smooth. The continuation criteria presented in [25] is a simpler Prodi–Serrin-type condition without weights.

The following result can be found in [46, Theorem 3.7]. We use it to improve the continuation criteria from Theorem 2.1 in the space-homogeneous case.

THEOREM 2.2. (from [46]) *Assume that $\alpha(r)=r^\gamma$ for some $\gamma\in[-3, 2]$. Assume that $f: [0, T]\times\mathbb{R}^3\rightarrow[0, \infty)$ is a classical solution to (1.1). Let $p>3/(5+\gamma)$, $p\geq 1$, and k be sufficiently large (depending on γ). Assume that, for all $t\in[0, T]$,*

$$\|f(t)\|_{L_k^p}\leq C_0.$$

Then, f is bounded for positive time with an upper bound

$$\|f(t)\|_{L^\infty}\leq C_1(1+t^{-3/(2p)})$$

for a constant C_1 that depends on C_0 , the mass of f_0 , and δ .

While the conditional upper bound in Theorem 2.2 does not readily provide any decay for large velocities, they can be deduced using the technique in [14]. The results in [14] are only stated in the moderately soft potential range $\gamma\in[-2, 0]$. However, they apply in the full range of soft potentials when we know in addition that the solution f is bounded.

THEOREM 2.3. (Essentially from [14]) *Assume that $\alpha(r)=r^\gamma$ for some $\gamma\in[-3, 0]$ and let $f: [0, T]\times\mathbb{R}^3\rightarrow[0, \infty)$ be a classical solution to (1.1). Assume that, for k sufficiently large and all $t\in[0, T]$,*

$$\|f(t)\|_{L^\infty}\leq C_0 \quad \text{and} \quad \|f(t)\|_{L_k^1}\leq C_0.$$

Assume further that $f_0\leq C_1 e^{-\beta|v|^2}$ for some $C_1>0$ and $\beta>0$ sufficiently small. Then,

$$f(t, v)\leq C_2 e^{-\beta|v|^2}.$$

for a constant C_2 that depends on C_0 , C_1 and the mass of f_0 .

Sketch of proof. Theorem 2.3 is proved in [14, §5] for the moderately soft potential range $\gamma\in(-2, 0]$. The fact $\gamma>-2$ is only used for the initial L^∞ bounds and the upper bounds on the coefficients given in [14, Lemma 2.1]. We included the extra assumptions that $\|f\|_{L^\infty}\leq C_0$ and $\|f(t)\|_{L_k^1}\leq C_0$ to take care of those requirements. We are only left to check that the upper bounds on the coefficients can be verified with our extra assumptions.

For $\gamma\leq-2$, following the proof of the computation for upper bounds for the coefficients \bar{a}_{ij} in [14, Lemma 2.1], we get

$$\begin{aligned} \bar{a}_{ij}e_i e_j &\lesssim \int |v-w|^{\gamma+2} f(w) dw \\ &\leq \int_{B_r(v)} |v-w|^{\gamma+2} f(w) dw + \int_{B_{|v|/2}\setminus B_r(v)} |v-w|^{\gamma+2} f(w) dw \\ &\quad + \int_{\mathbb{R}^3\setminus(B_{|v|/2}\cup B_r(v))} |v-w|^{\gamma+2} f(w) dw \\ &\lesssim r^{\gamma+5} \|f\|_{L^\infty} + \langle v \rangle^{\gamma+2} \|f\|_{L^1} + r^{\gamma+2} \langle v \rangle^{-k} \|f\|_{L_k^1}. \end{aligned}$$

We choose $r = \langle v \rangle^{(\gamma+2)/(\gamma+5)}$, and then k large enough so that

$$r^{\gamma+2} \langle v \rangle^{-k} \leq \langle v \rangle^{\gamma+2}.$$

Thus, we obtain the same upper bounds for the coefficients \bar{a}_{ij} as in [14, Lemma 2.1] for the case $\gamma \in [-3, -2]$ when we assume in addition that $\|f\|_{L^\infty}$ and $\|f\|_{L_k^1}$ are bounded.

The upper bound in [14, Lemma 2.1] improves when e is parallel to v , which is proved along the same lines. Once these upper bounds are established, the rest of §5 in [14] follows without any major change. \square

We also recall a result on the propagation of bounded moments. The following theorem is a simplified version of a result in [16].

THEOREM 2.4. *If f is a solution of the Landau equation (1.1), where $\alpha(r) = r^\gamma$ with $\gamma \in [-3, 1]$. Assume that the initial data f_0 belongs to L_k^1 for some exponent $k \geq 0$. Then, for all $t \in [0, T]$, we have $\|f(t)\|_{L_k^1} \leq C(T)$, for some upper bound $C(T)$ depending on T , $\|f_0\|_{L_k^1}$, and the mass, energy and entropy of f_0 .*

Theorem 2.4 is stated in [16] for the Coulomb case $\gamma = -3$ only. The proof easily applies to any soft potential in the range $\gamma \in [-3, 0]$ (and even slightly smaller than -3). The result in [16] is also stated and proved for H -solutions, which makes it technically more complicated. The case $\gamma = 0$ had been covered earlier in [54]. For $\gamma > 0$, it is proved in [19] that there is even a gain of moments in the sense that $\|f(t)\|_{L_k^1}$ is bounded for $t > 0$ even if $\|f_0\|_{L_k^1}$ is not. Note that, for $\gamma < 0$, even though $C(T)$ may a-priori grow for large time, it does not blow up in finite time.

For uniqueness of solutions, we cite [23] for $\gamma > -3$ and [22] for $\gamma = -3$. The following theorem summarizes both cases.

THEOREM 2.5. *When $\alpha(r) = r^\gamma$ with $\gamma \in [-3, 2]$, there exists at most one weak solution of (1.1) that belongs to the space*

$$f \in L^\infty([0, T], L_2^1(\mathbb{R}^3)) \cap L^1([0, T], L^p(\mathbb{R}^3)).$$

Here, $p = \infty$ for $\gamma = -3$, and $p = 3/(3+\gamma)$ otherwise.

The solutions that we obtain from Theorem 2.1 are C^∞ smooth for $t > 0$. If f_0 is not continuous, the initial data is only achieved in the sense of weak continuity. Because of that, it is not irrelevant that Theorem 2.5 applies to weak solutions. It tells us in particular that there cannot be two different solutions bifurcating from the same initial data at time zero. The precise definition of weak solution given in [22], [23] is not necessarily important here, but only the fact that it is compatible with a smooth solution that achieves a potentially rough initial data by weak continuity.

As we mentioned in the introduction, the Fisher information of a non-negative function $f: \mathbb{R}^3 \rightarrow [0, \infty)$ is given by the formulas

$$i(f) := \int_{\mathbb{R}^3} \frac{|\nabla f|^2}{f} dv = \int_{\mathbb{R}^3} |\nabla \log f|^2 f dv = 4 \int_{\mathbb{R}^3} |\nabla \sqrt{f}|^2 dv.$$

Something must be clarified about the case when f has vacuum regions. In that case, the last expression is the only one that makes immediate sense and suggests that we should set the integrand as zero wherever $f=0$. This apparent difficulty with vacuum regions is actually not relevant for our analysis. In [34], the authors prove that the solution to the Landau equation becomes immediately strictly positive (even in the space inhomogeneous setting).

The Fisher information $i(f)$ is well defined for non-negative functions f so that $\sqrt{f} \in \dot{H}^1(\mathbb{R}^3)$. This will always be the case with the class of solutions that we work with (see Appendix B). Moreover, the following lemma from [51] indicates that the Fisher information is finite for all non-negative functions that are sufficiently smooth and fast decaying at infinity.

LEMMA 2.6. *For all $\varepsilon > 0$, there exists $C_\varepsilon > 0$ (only depending on dimension and ε) such that, for all $f \in H_{3/2+\varepsilon}^2(\mathbb{R}^3)$, one has*

$$I(f) \leq C_\varepsilon \|f\|_{H_{3/2+\varepsilon}^2}.$$

We use the notation $\langle i'(f), g \rangle$ to denote the Gateaux derivative of i in the direction g . More precisely, if $f(t)$ is a curve of non-negative functions such that

$$\sqrt{f(t)} \in \dot{H}^1 \quad \text{and} \quad \partial_t f(0) = g,$$

then, by definition,

$$\langle i'(f), g \rangle = \partial_t i(f(t))|_{t=0}.$$

Likewise, $i''(f)$ denotes the quadratic form such that, if $f(t)$ is second differentiable with respect to t , then

$$\langle i''(f)g, g \rangle + \langle i'(f), \partial_{tt} f|_{t=0} \rangle = \partial_{tt} i(f(t))|_{t=0}.$$

It is a standard fact that the values of these formulas do not depend on the curve $f(t)$ other than through the values of $g = \partial_t f(0)$. Moreover, the convexity of i corresponds to $i''(f)$ being a positive quadratic form.

We use Theorem 2.1 as the starting point to construct our global smooth solution in Theorem 1.2. Combining it with the propagation of moments in Theorem 2.4 and the Maxwellian upper bounds of Theorem 2.3, the solutions we work with are very-well-behaved functions. They are C^∞ smooth functions, they have rapid decay at infinity

together with all their derivatives, and they are strictly positive. This is the class of functions that we work with in this paper, which minimizes the technical inconveniences in our computations. *A posteriori*, Theorem 1.1, and most of the lemmas in this paper, apply to a much wider class of functions by a density argument (see Appendix B). However, we will make no attempt to classify the most general class of functions for which our results make sense.

Even when f and $|\nabla f|$ decay faster than any algebraic rate as $|v| \rightarrow \infty$, it is not completely obvious that $|\nabla f|^2/f$ is integrable in \mathbb{R}^3 . Moreover, there are several integrals throughout this paper that involve first, second and even third derivatives of $\log f$, and these quantities may grow for large velocities. In the same vein as Lemma 2.6, it can be argued that if sufficiently many derivatives of f decay sufficiently fast as $|v| \rightarrow \infty$, then all the integrals in this article make sense. However, we can also avoid this technical annoyance altogether by approximating the function f with one that is bounded from below and above by a multiple of the same Maxwellian. We do not want to distract the reader with this very standard technical point, so we explain it in Appendix B. For the rest of this paper, we manipulate integrals involving the function f (or later F in the lifted variables) multiplying various derivatives of its logarithm without further comments.

3. Lifting and projection

We have found that the question of monotonicity of $i(f)$ for f solving (1.1) is essentially the same as the question of monotonicity of the Fisher information for a specific linear second-order equation in a space with double the original number of variables. This is a significant simplification, it means we can understand our question by studying a simpler and concrete evolution equation, and one that is linear and local—in contrast with the original equation. The starting point is realizing that $q(f)$ is defined as the integral of a function in a space with twice the number of variables, and therefore it makes sense to “lift” the flow and any quantity of interest into this space with double the variables. The purpose of this section is to describe the usage of these lifting and projection operations.

Given a function $f: \mathbb{R}^d \rightarrow \mathbb{R}$, we can build a function $F: \mathbb{R}^{d+d} \rightarrow \mathbb{R}$ by taking the tensor product of f with itself, as follows:

$$F(v, w) = (f \otimes f)(v, w) := f(v)f(w).$$

This defines a map from scalar functions in \mathbb{R}^d to scalar functions in \mathbb{R}^{d+d} . Conversely, we define the following *projection* operator

$$\pi(F)(v) := \int_{\mathbb{R}^d} F(v, w) dw.$$

This is an operator that maps a function $F(v, w)$ defined in \mathbb{R}^{d+d} into a function $\pi(F): \mathbb{R}^d \rightarrow \mathbb{R}$. The map π is a bounded linear map from $L^1(\mathbb{R}^{d+d})$ to $L^1(\mathbb{R}^d)$. Also observe that, if $F=F(v, w)$ is a probability density in \mathbb{R}^{d+d} , then $\pi(F)$ is simply the marginal probability distribution in the $v \in \mathbb{R}^d$ variable.

We will loosely use the terminology that the variables $(v, w) \in \mathbb{R}^{d+d}$ are the “lifted variables”, functions in \mathbb{R}^{d+d} are lifted functions, and so on. The next ingredient is a second-order differential operator for lifted functions. Concretely, we consider the following degenerate-elliptic operator that applies to functions $F=F(v, w)$:

$$Q(F)(v, w) := (\partial_{v_i} - \partial_{w_i})(\alpha(|v-w|)a_{ij}(v-w)(\partial_{v_j} - \partial_{w_j})F(v, w)). \quad (3.1)$$

For any function $f=f(v)$, we may study the initial value problem associated to the linear operator Q with initial data $f \otimes f$:

$$\begin{aligned} F(0, v, w) &= f(v)f(w), \\ F_t &= Q(F). \end{aligned} \quad (3.2)$$

Equation (3.2) is a linear degenerate parabolic equation with explicit coefficients. Its classical well-posedness is straightforward from classical PDE theory.⁽¹⁾ The integral of F on \mathbb{R}^6 is constant in time. From the symmetry of the initial data, and the symmetry of the equation, we deduce that, for all $t > 0$, the solution must be symmetric in the sense that $F(t, v, w) = F(t, w, v)$.

LEMMA 3.1. *For any twice-differentiable function f which is non-negative and has mass 1, the Landau operator q (given in (1.2)) coincides with*

$$q(f) = \pi(Q(F)) = \partial_t[\pi F]|_{t=0},$$

where F is the solution of (3.2).

Proof. The proof follows by a straightforward substitution. We see that, by the linearity of π ,

$$\begin{aligned} \partial_t[\pi(F)]|_{t=0} &= \pi(Q(F))|_{t=0} \\ &= \int_{\mathbb{R}^3} ((\partial_{v_i} - \partial_{w_i})(\alpha(|v-w|)a_{ij}(v-w)(\partial_{v_j} - \partial_{w_j})F(0, v, w))) dw. \end{aligned}$$

The derivatives with respect to w_i integrate to zero respect to dw . Thus,

$$\partial_t[\pi(F)]|_{t=0} = \partial_{v_i} \int_{\mathbb{R}^3} \alpha(|v-w|)a_{ij}(v-w)(\partial_{v_j} - \partial_{w_j})F(0, v, w) dw = q(f). \quad \square$$

⁽¹⁾ In fact, one can see that equation (3.2) is literally the heat equation on each of the spheres that we describe in (4.6) with diffusion coefficient α . See Remark 10.3.

We use I for the Fisher information in the *lifted* variables. For $F: \mathbb{R}^6 \rightarrow (0, \infty)$,

$$I(F) := \iint_{\mathbb{R}^6} \frac{|\nabla F|^2}{F} dw dv = \iint_{\mathbb{R}^6} |\nabla \log F|^2 F dw dv = 4 \iint_{\mathbb{R}^6} |\nabla \sqrt{F}|^2 dw dv \quad (3.3)$$

When F has vacuum regions, only the last expression makes sense literally. For the first two equalities to make sense, and agree with the third one, we must set

$$\frac{|\nabla F|^2}{F} = |\nabla \log F|^2 F = 0$$

when $F=0$.

The next lemmas relate $i(f)$ with $I(F)$. The first one is very standard.

LEMMA 3.2. *For any non-negative C^1 function f with $\int f dv = 1$ and $i(f) < \infty$, we have*

$$i(f) = \frac{1}{2} I(f \otimes f).$$

Proof. We are going to show a bit more: if f is a C^1 function with $f \geq 0$ and f is not identically zero, then

$$I(f \otimes f) = 2 \left(\int_{\mathbb{R}^3} f dv \right) i(f).$$

First, let us prove this when $f(v) > 0$ for all v . Since $(f \otimes f)(v, w) = f(v)f(w)$,

$$\nabla_{\mathbb{R}^6} (f \otimes f)(v, w) = (f(w) \nabla_{\mathbb{R}^3} f(v), f(v) \nabla_{\mathbb{R}^3} f(w)),$$

so

$$|\nabla(f \otimes f)|^2 = f(w)^2 |\nabla f(v)|^2 + f(v)^2 |\nabla f(w)|^2.$$

In particular,

$$\begin{aligned} \frac{|\nabla(f \otimes f)|^2}{(f \otimes f)}(v, w) &= \frac{f(w)^2 |\nabla f(v)|^2 + f(v)^2 |\nabla f(w)|^2}{f(v)f(w)} \\ &= f(w) \frac{|\nabla f(v)|^2}{f(v)} + f(v) \frac{|\nabla f(w)|^2}{f(w)}. \end{aligned}$$

Then, we integrate over \mathbb{R}^6 , getting

$$\begin{aligned} I(f \otimes f) &= \iint_{\mathbb{R}^6} \frac{|\nabla(f \otimes f)|^2}{(f \otimes f)}(v, w) dw dv \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \left(f(w) \frac{|\nabla f(v)|^2}{f(v)} + f(v) \frac{|\nabla f(w)|^2}{f(w)} \right) dw dv \\ &= 2 \left(\int_{\mathbb{R}^3} f(w) dw \right) \left(\int_{\mathbb{R}^3} \frac{|\nabla f(v)|^2}{f(v)} dv \right). \end{aligned}$$

For a general non-negative f of class C^1 , we write $f_\varepsilon = f + \varepsilon\phi$, where $\phi(x) = e^{-|v|^2}$. Then, $f_\varepsilon > 0$ for every $\varepsilon > 0$ and

$$I(f_\varepsilon \otimes f_\varepsilon) = 2 \left(\int_{\mathbb{R}^3} f_\varepsilon dv \right) i(f_\varepsilon).$$

Now, we note that

$$\begin{aligned} i(f_\varepsilon) &= \int_{\mathbb{R}^3} \frac{|\nabla f(v)|^2 + 2\varepsilon \nabla f(v) \cdot \nabla \phi(v) + \varepsilon^2 |\nabla \phi(v)|^2}{f_\varepsilon(v)} dv \\ &= \int_{\mathbb{R}^3} \frac{|\nabla f(v)|^2}{f_\varepsilon(v)} dv + \int_{\mathbb{R}^3} \frac{2\varepsilon \nabla f(v) \cdot \nabla \phi(v) + \varepsilon^2 |\nabla \phi(v)|^2}{f_\varepsilon(v)} dv. \end{aligned}$$

The first integral converges to $i(f)$ as $\varepsilon \rightarrow 0$, due to $i(f) < \infty$ and the dominated convergence theorem. The absolute value of the second one is bounded from above by

$$\varepsilon \int_{\mathbb{R}^3} \frac{|\nabla f(v)|^2}{f(v)} dv + \varepsilon(\varepsilon+1) \int_{\mathbb{R}^3} \frac{|\nabla \phi(v)|^2}{\phi(v)} dv.$$

It follows that $i(f_\varepsilon) \rightarrow i(f)$. It is obvious that

$$\int_{\mathbb{R}^3} f_\varepsilon(v) dv \rightarrow \int_{\mathbb{R}^3} f(v) dv.$$

Lastly,

$$f_\varepsilon \otimes f_\varepsilon = f \otimes f + \varepsilon(f \otimes \phi + \phi \otimes f) + \varepsilon^2 \phi \otimes \phi.$$

From here, one can see that $I(f_\varepsilon \otimes f_\varepsilon) \rightarrow I(f \otimes f)$ as $\varepsilon \rightarrow 0$ in a manner entirely analogous to $i(f_\varepsilon) \rightarrow i(f)$, by noting that $f_\varepsilon \otimes f_\varepsilon \geq \max\{f \otimes f, \phi \otimes f\}$ for every $\varepsilon > 0$ and

$$\nabla_{\mathbb{R}^6}(f_\varepsilon \otimes f_\varepsilon) \rightarrow \nabla_{\mathbb{R}^3}(f \otimes f)$$

pointwise as $\varepsilon \rightarrow 0$. □

The following is a crucial property of the Fisher information.

LEMMA 3.3. *For any non-negative C^1 function $F: \mathbb{R}^6 \rightarrow [0, \infty)$ with $F(v, w) = F(w, v)$, we have*

$$i(\pi F) \leq \frac{1}{2} I(F).$$

The inequality in Lemma 3.3 was first observed by Eric Carlen in [15], without the symmetry assumption. There, he showed that some classical results related to the Fisher information follow as a consequence of this elementary inequality.

Proof. First we show inequality under the additional assumption that $F > 0$ everywhere. We note that

$$\nabla(\pi F)(v) = \nabla_v \left\{ \int_{\mathbb{R}^3} F(v, w) dw \right\} = \int_{\mathbb{R}^3} \nabla_v F(v, w) dw,$$

so

$$|\nabla(\pi F)(v)| \leq \int_{\mathbb{R}^3} |\nabla_v F(v, w)| dw.$$

Since $F > 0$ everywhere, we may apply the Cauchy–Schwarz inequality to get

$$\left(\int_{\mathbb{R}^3} |\nabla_v F(v, w)| dw \right)^2 \leq \left(\int_{\mathbb{R}^3} \frac{|\nabla_v F(v, w)|^2}{F(v, w)} dw \right) \left(\int_{\mathbb{R}^3} F(v, w) dw \right).$$

The last two inequalities combine to

$$|\nabla(\pi F)(v)|^2 \leq \left(\int_{\mathbb{R}^3} \frac{|\nabla_v F(v, w)|^2}{F(v, w)} dw \right) (\pi F)(v) \quad \text{for all } v.$$

Dividing by $(\pi F)(v)$ (which is > 0 for all v , since $F > 0$ everywhere) and integrating in v ,

$$i(\pi F) \leq \iint_{\mathbb{R}^3 \times \mathbb{R}^3} \frac{|\nabla_v F(v, w)|^2}{F(v, w)} dw dv = \frac{1}{2} I(F).$$

To obtain the last equality, we have used that $F(v, w) = F(w, v)$.

Now, let F be just as in the statement of the lemma, and for every $\varepsilon > 0$ define

$$F_\varepsilon(v, w) = F(v, w) + \varepsilon e^{-|v|^2 - |w|^2}.$$

Then, F_ε also satisfies all the assumptions of the lemma, but in addition $F_\varepsilon > 0$ everywhere, and thus

$$i(\pi F_\varepsilon) \leq \frac{1}{2} I(F_\varepsilon) \quad \text{for all } \varepsilon > 0.$$

Observe that

$$(\pi F_\varepsilon)(v) = (\pi F)(v) + c\varepsilon e^{-|v|^2}$$

for some dimensional constant c . By repeating the argument in the proof of the previous lemma, we conclude that $i(\pi F_\varepsilon) \rightarrow i(\pi F)$ and $I(F_\varepsilon) \rightarrow I(F)$ as $\varepsilon \rightarrow 0^+$, and the lemma follows. \square

Lemma 3.3 is very important to make this program work. It would also hold if we replace the Fisher information with the usual entropy, but it would not hold for most usual quantities related to f that involve derivatives.

LEMMA 3.4. *Let f be a solution of (1.1) and F as in (3.2). Then*

$$\partial_t i(f) \leq \frac{1}{2} \partial_t I(F)|_{t=0}.$$

Proof. Using Lemma 3.1, we get

$$\begin{aligned} \partial_t i(f) &= \langle i'(f), q(f) \rangle \\ &= \langle i'(f), \partial_t(\pi F)|_{t=0} \rangle \\ &= \partial_t i(\pi F)|_{t=0} \\ &= \partial_t \left(i(\pi F) - \frac{1}{2} I(F) \right) \Big|_{t=0} + \frac{1}{2} \partial_t I(F) \Big|_{t=0}. \end{aligned}$$

Because of Lemma 3.3, the first term is the time derivative of a non-positive function that achieves its maximum at $t=0$. Therefore, it cannot be positive. We end up with

$$\partial_t i(f) \leq \frac{1}{2} \partial_t I(F) \Big|_{t=0}. \quad \square$$

It is convenient to rewrite Lemma 3.4 in terms of the Gateaux derivatives of i and I . It says that, for $F=f \otimes f$, we have

$$\langle i'(f), q(f) \rangle \leq \frac{1}{2} \langle I'(F), Q(F) \rangle. \quad (3.4)$$

In order to prove that the left-hand side is negative, we will show that the right-hand side is. The remarkable advantage of our analysis is that the question of monotonicity of the Fisher information for the Landau equation, which is non-linear and non-local, is now reduced to its monotonicity for a linear equation with explicit coefficients, in the doubled-variables. Precisely, we will prove Theorem 1.1 as soon as we prove the following inequality for any function $F: \mathbb{R}^6 \rightarrow [0, \infty)$ such that $F(v, w) = F(w, v)$:

$$\langle I'(F), Q(F) \rangle \leq 0. \quad (3.5)$$

Proving inequality (3.5) is the objective of the coming sections. We want to apply it to solutions of the Landau equation. From Theorem 2.1, these solutions are strictly positive, smooth and rapidly decaying functions. We must show that (3.5) holds for any smooth function $F: \mathbb{R}^6 \rightarrow (0, \infty)$ with sufficient decay at infinity. There is no time evolution explicitly involved in the inequality (3.5).

Remark 3.5. Inequality (3.4) is in fact an equality. There is a very elementary justification that we explain here. Consider the function $F(t, v, w)$ given by

$$F(t, v, w) = f(v)f(w) + tQ(f \otimes f).$$

Suppose that $F(t, v, w) \geq 0$ at least in some small interval of time $t \in (-\delta, \delta)$. Otherwise, we may approximate it with a non-negative function appropriately.

Let $\iota(t)$ be the function

$$\iota(t) := i(\pi F(t, \cdot)) - \frac{1}{2} I(F(t, \cdot)).$$

Since

$$F(0, v, w) = f(v)f(w),$$

Lemma 3.2 tells us that $\iota(0) = 0$. Lemma 3.3 tells us that $\iota(t) \leq 0$ for all $t \in (-\delta, \delta)$. Thus, $\iota(t)$ achieves its maximum at $t = 0$ and we deduce that $\iota'(0) = 0$.

If is also possible, although it looks rather magical, to justify the equality by writing $\langle i'(f), q(f) \rangle$ as a double integral and symmetrizing the expression in the correct way.

Since (3.4) is in fact an equality, the inequality (3.5) turns out to be equivalent to the statement of Theorem 1.1. Our analysis so far is sharp.

4. Flowing along vector fields

In this section we decompose the lifted Landau operator (3.1) as a sum of second-order differential operators each acting along one direction in \mathbb{R}^6 . The vector fields \tilde{b}_k giving these directions correspond to the generators of rotations along three perpendicular axes in \mathbb{R}^3 (which are then “lifted” to \mathbb{R}^6). We write the majority of the computations in this paper using the vectors b_k . Because of this, it is important to get used to their basic arithmetic properties which we describe in this section.

First, for each value of $v - w$, we define the following three vectors

$$b_1(v - w) = \begin{pmatrix} 0 \\ w_3 - v_3 \\ v_2 - w_2 \end{pmatrix}, \quad b_2(v - w) = \begin{pmatrix} v_3 - w_3 \\ 0 \\ w_1 - v_1 \end{pmatrix}, \quad b_3(v - w) = \begin{pmatrix} w_2 - v_2 \\ v_1 - w_1 \\ 0 \end{pmatrix}. \quad (4.1)$$

One can check that the vector fields b_1 , b_2 and b_3 are always perpendicular to $v - w$. In fact, for each fixed value of $v - w$, with $v \neq w$ these three vectors span the plane in \mathbb{R}^3 perpendicular to $v - w$. The following identity, that is readily verified by a direct computation, links these vectors with the Landau equation:

$$b_1 \otimes b_1 + b_2 \otimes b_2 + b_3 \otimes b_3 = a_{ij}(v - w), \quad (4.2)$$

where

$$a_{ij}(z) = |z|^2 \delta_{ij} - z_i z_j$$

is the matrix used in the definition of the Landau operator (1.2). As a consequence of these two observations, the lifted operator Q (3.1) can be rewritten as

$$\begin{aligned} Q(F) &= \sum_{k=1}^3 (\operatorname{div}_v - \operatorname{div}_w)(\alpha(|v-w|)b_k \otimes b_k (\nabla_v - \nabla_w)F) \\ &= \sum_{k=1}^3 (\partial_{v_i} - \partial_{w_i})(\alpha(|v-w|)(b_k)_i (b_k)_j \cdot (\partial_{v_j} - \partial_{w_j})F). \end{aligned}$$

Remark 4.1. Denote by $\Pi(z)$ ($z \neq 0$) the orthogonal projection from \mathbb{R}^3 to the plane perpendicular to z . It is elementary that

$$\Pi_{ij}(z) = \delta_{ij} - |z|^{-2} z_i z_j$$

and so we see that

$$a_{ij}(z) = |z|^2 \Pi(z).$$

Accordingly the formula relating a_{ij} and the vector fields b_k 's gives a corresponding one for $\Pi(z)$, which we record here:

$$\Pi_{ij}(z) = |z|^{-2} \sum_{k=1}^3 (b_k(z))_i (b_k(z))_j.$$

Since the differential operators $b_k \cdot (\nabla_v - \nabla_w)$ act on functions in \mathbb{R}^6 , it is both convenient and natural to write it in terms of vectors in \mathbb{R}^6 . For $k=1, 2, 3$, we define the following corresponding vector fields \tilde{b}_k in \mathbb{R}^6 :

$$\tilde{b}_k(v-w) = \begin{pmatrix} b_k \\ -b_k \end{pmatrix}. \quad (4.3)$$

We abuse notation by writing $\tilde{b}_k: \mathbb{R}^6 \rightarrow \mathbb{R}^6$ to denote $\tilde{b}_k = \tilde{b}_k(v-w)$. The vector fields \tilde{b}_1 , \tilde{b}_2 and \tilde{b}_3 are divergence free in \mathbb{R}^6 . They satisfy the following matrix identity:

$$\tilde{b}_1 \otimes \tilde{b}_1 + \tilde{b}_2 \otimes \tilde{b}_2 + \tilde{b}_3 \otimes \tilde{b}_3 = \begin{pmatrix} \{a_{ij}(v-w)\} & -\{a_{ij}(v-w)\} \\ -\{a_{ij}(v-w)\} & \{a_{ij}(v-w)\} \end{pmatrix}. \quad (4.4)$$

Moreover, $\tilde{b}_k(v-w)$ is always perpendicular to

$$\begin{pmatrix} v-w \\ w-v \end{pmatrix}.$$

In particular, for any function $\beta = \beta(|v-w|)$, we always have (for the gradient and the divergence in \mathbb{R}^6)

$$\tilde{b}_k \cdot \nabla \beta = 0 \quad \text{and} \quad \operatorname{div}(\beta(|v-w|) \tilde{b}_k) = 0.$$

In light of all the above, we can write Q in terms of these vector fields \tilde{b}_k :

$$\begin{aligned} Q(F) &= \sum_{k=1}^3 \operatorname{div}(\alpha(|v-w|)(\tilde{b}_k \cdot \nabla F) \tilde{b}_k) \\ &= \sum_{k=1}^3 \alpha(|v-w|)(\tilde{b}_k \cdot \nabla(\tilde{b}_k \cdot \nabla F)) \\ &= \sum_{k=1}^3 \sqrt{\alpha} \tilde{b}_k \cdot \nabla(\sqrt{\alpha} \tilde{b}_k \cdot \nabla F). \end{aligned}$$

In this formula, ∇F stands for the gradient of F with respect to the variable $(v, w) \in \mathbb{R}^6$. We think of the operator $Q(F)$ as the sum of the second derivatives of F along the directions \tilde{b}_k , with a weight $\alpha(|v-w|)$. To make this point clearer, while also introducing additional notation we will use later, let us define, for a given vector field b in \mathbb{R}^6 , the operator

$$L_b(F) := b \cdot \nabla F.$$

Then, $L_b(F)$ is simply the derivative of F along b . One sees that the above decomposition for Q can be restated as a sum of squares of operators of the form L_b , concretely

$$Q = \sum_{k=1}^3 L_{\sqrt{\alpha} \tilde{b}_k} \circ L_{\sqrt{\alpha} \tilde{b}_k} = \sum_{k=1}^3 \alpha L_{\tilde{b}_k} \circ L_{\tilde{b}_k}. \quad (4.5)$$

It is interesting to notice that, if we consider the flow of each vector field \tilde{b}_k ,

$$\begin{pmatrix} v'(t) \\ w'(t) \end{pmatrix} = \tilde{b}_k(v(t) - w(t)),$$

then

$$v(t) + w(t), \quad |v(t)|^2 + |w(t)|^2 \quad \text{and} \quad |v(t) - w(t)|$$

are all constant in t . This is simply the fact that \tilde{b} is everywhere perpendicular to the gradient of $|v|^2 + |w|^2$, to the gradient of $|v-w|$, and to vectors of the form $(e, e) \in \mathbb{R}^6$ with $e \in \mathbb{R}^3$.

Given any point $(v, w) \in \mathbb{R}^6$, the set of points in \mathbb{R}^6 that we may reach by flowing along the three vector fields b_k 's, is exactly the sphere used in the usual formulas for the Boltzmann equation:

$$\operatorname{sphere}(v, w) := \{(v', w') \in \mathbb{R}^6 : v' + w' = v + w \text{ and } |v'|^2 + |w'|^2 = |v|^2 + |w|^2\}. \quad (4.6)$$

Based on this intuition, there are some quantities that we define later on taking into account derivatives along the directions \tilde{b}_k for which we use names referring to this sphere. The set $\text{sphere}(v, w)$ is indeed a 2-dimensional sphere of radius $|v-w|/\sqrt{2}$ isometrically embedded in \mathbb{R}^6 . It is the same sphere used in the usual parametrization of the Boltzmann equation:

$$\text{sphere}(v, w) = \left\{ (v', w') : v' = \frac{v+w}{2} + \frac{|v-w|}{2} \sigma \text{ and } w' = \frac{v+w}{2} - \frac{|v-w|}{2} \sigma \text{ for } \sigma \in S^2 \right\}.$$

At any given point $(v, w) \in \mathbb{R}^6$, the vectors $\tilde{b}_1(v-w)$, $\tilde{b}_2(v-w)$ and $\tilde{b}_3(v-w)$ are tangent to $\text{sphere}(v, w)$. These vectors are necessarily linearly dependent. We can easily verify this fact by hand, since

$$(v_1 - w_1)\tilde{b}_1 + (v_2 - w_2)\tilde{b}_2 + (v_3 - w_3)\tilde{b}_3 = 0.$$

Using the vector fields \tilde{b}'_i s, we will eventually verify that the operator Q is exactly the Laplace–Beltrami operator on $\text{sphere}(v, w)$ times the function $|v-w|^2\alpha(|v-w|)$. See Remark 10.3.

There is one last vector field in \mathbb{R}^6 that will play an important role in our analysis. Let $n: \mathbb{R}^6 \rightarrow \mathbb{R}^6$ be the unit normal vector to the level sets of $|v-w|$. That is

$$n = \frac{1}{\sqrt{2}|v-w|} \begin{pmatrix} v-w \\ w-v \end{pmatrix}. \quad (4.7)$$

We can see that, for $k=1, 2, 3$, $[n, \tilde{b}_k] = 0$. For the reader's convenience, let us recall the definition of the Lie bracket $[a, b]$, this is the vector whose components are given by

$$[a, b]_i = a_j \partial_j b_i - b_j \partial_j a_i.$$

The Lie bracket computes the error of commuting the differentiation with respect to a and b :

$$a \cdot \nabla(b \cdot \nabla u) - b \cdot \nabla(a \cdot \nabla u) = [a, b] \cdot \nabla u \quad (4.8)$$

In contrast to the vectors \tilde{b}_k , in general we have

$$n \cdot \nabla \alpha \neq 0 \quad \text{whenever} \quad \alpha = \alpha(|v-w|).$$

Moreover, we have

$$[n, \alpha \tilde{b}_k] = (n \cdot \nabla \alpha) \tilde{b}_k = \sqrt{2} \alpha'(|v-w|) \tilde{b}_k.$$

For the last identity, note that

$$\begin{aligned}
n \cdot \nabla \alpha &= n \cdot \frac{\alpha'(|v-w|)}{|v-w|} \binom{v-w}{w-v} \\
&= \frac{1}{\sqrt{2}|v-w|} \binom{v-w}{w-v} \cdot \frac{\alpha'(|v-w|)}{|v-w|} \binom{v-w}{w-v} \\
&= \frac{2|v-w|^2}{\sqrt{2}|v-w|^2} \alpha'(|v-w|) \\
&= \sqrt{2} \alpha'(|v-w|).
\end{aligned}$$

We will also be using vector fields to decompose the Fisher information, and we record here a definition and two preparatory lemmas that will be used later. Given a vector field $e: \mathbb{R}^6 \rightarrow \mathbb{R}^6$, we define the functional

$$I_e(F) := \iint_{\mathbb{R}^6} \frac{(e \cdot \nabla F)^2}{F} dw dv.$$

One can see $I_e(F)$ as a kind of Fisher information⁽²⁾ of F that only uses ∇F in the direction e . The first of the preparatory lemmas provides a formula for the Gateaux derivative of I_e at F (for some vector field e) in the direction given by $L_b(F)$ (for some divergence-free vector field b).

LEMMA 4.2. *Let e and b be vector fields in \mathbb{R}^6 and assume that b satisfies $\operatorname{div}(b)=0$. Then, the following identity holds for any smooth positive function $F: \mathbb{R}^6 \rightarrow (0, \infty)$ with rapid decay at infinity:*

$$\langle I'_e(F), L_b(F) \rangle = 2 \iint_{\mathbb{R}^6} (e \cdot \nabla \log F) ([e, b] \cdot \nabla \log F) F dw dv.$$

Proof. We directly compute

$$\langle I'_e(F), L_b(F) \rangle = \iint_{\mathbb{R}^6} \left(2 \frac{(e \cdot \nabla F)(e \cdot \nabla(b \cdot \nabla F))}{F} - \frac{(e \cdot \nabla F)^2}{F^2} (b \cdot \nabla F) \right) dw dv.$$

Observe that

$$e \cdot \nabla(b \cdot \nabla F) = b \cdot \nabla(e \cdot \nabla F) + [e, b] \cdot \nabla F.$$

This results in

$$\begin{aligned}
&\langle I'_e(F), L_b(F) \rangle \\
&= \iint_{\mathbb{R}^6} \left(2 \frac{(e \cdot \nabla F)(b \cdot \nabla(e \cdot \nabla F))}{F} - \frac{(e \cdot \nabla F)^2}{F^2} (b \cdot \nabla F) + 2 \frac{(e \cdot \nabla F)([e, b] \cdot \nabla F)}{F} \right) dw dv.
\end{aligned}$$

⁽²⁾ In fact, $I_e(F)$ is the standard Fisher information (as understood in statistics) at $\theta=0$ for the 1-parameter distribution F_θ obtained by transporting the distribution F along the flow given by e .

The first two terms in the integrand add up to a divergence, since

$$2\frac{(e \cdot \nabla F)(b \cdot \nabla(e \cdot \nabla F))}{F} - \frac{(e \cdot \nabla F)^2}{F^2}(b \cdot \nabla F) = b \cdot \nabla \frac{(e \cdot \nabla F)^2}{F} = \operatorname{div} \left(\frac{(e \cdot \nabla F)^2}{F} b \right).$$

It follows that

$$\langle I'_e(F), L_b(F) \rangle = 2 \iint_{\mathbb{R}^6} \frac{(e \cdot \nabla F)([e, b] \cdot \nabla F)}{F} dw dv.$$

and the lemma now follows using that

$$\nabla F = F \nabla \log F.$$

□

The second preparatory lemma is the corresponding pure second derivative for $I_e(F)$.

LEMMA 4.3. *Let e and b be vector fields in \mathbb{R}^6 . Then, the following identity holds for any smooth positive function $F: \mathbb{R}^6 \rightarrow (0, \infty)$ with rapid decay at infinity:*

$$\langle I''_e(F)L_b(F), L_b(F) \rangle = 2 \iint_{\mathbb{R}^6} (e \cdot \nabla(b \cdot \nabla \log F))^2 F dw dv.$$

Proof. For any function $G: \mathbb{R}^6 \rightarrow \mathbb{R}$, we see that

$$\langle I'_e(F), G \rangle = \iint_{\mathbb{R}^6} \left(2 \frac{(e \cdot \nabla F)(e \cdot \nabla G)}{F} - \frac{(e \cdot \nabla F)^2}{F^2} G \right) dw dv.$$

Differentiating again, we get

$$\begin{aligned} \langle I''_e(F)G, G \rangle &= \iint_{\mathbb{R}^6} \left(2 \frac{(e \cdot \nabla G)^2}{F} - 4 \frac{(e \cdot \nabla F)(e \cdot \nabla G)}{F^2} G + 2 \frac{(e \cdot \nabla F)^2}{F^3} G^2 \right) dw dv \\ &= \iint_{\mathbb{R}^6} 2F \left(\frac{e \cdot \nabla G}{F} - G \frac{(e \cdot \nabla F)}{F^2} \right)^2 dw dv. \end{aligned}$$

In particular, if $G = L_b(F) = b \cdot \nabla F$, it reduces to

$$\begin{aligned} \langle I''_e(F)L_b(F), L_b(F) \rangle &= \iint_{\mathbb{R}^6} 2F \left(\frac{e \cdot \nabla(b \cdot \nabla F)}{F} - \frac{(b \cdot \nabla F)}{F} \frac{(e \cdot \nabla F)}{F} \right)^2 dw dv \\ &= \iint_{\mathbb{R}^6} 2F \left| e \cdot \nabla \left(b \cdot \frac{\nabla F}{F} \right) \right|^2 dw dv. \end{aligned}$$

□

5. The case of Maxwell molecules

In case $\alpha(|v-w|) \equiv 1$, the monotonicity of the Fisher information was obtained in [55], and also follows from the earlier work for the Boltzmann equation in [50], [52]. Here, we provide a quick proof by verifying (3.5) in that case. It also serves to explain some of the ideas that will be used later for more general interaction potentials α . Let \tilde{b}_1, \tilde{b}_2 and \tilde{b}_3 be the vector fields as in (4.3). We use formula (4.5) for $\alpha \equiv 1$:

$$Q(F) = \sum_{k=1}^3 L_{\tilde{b}_k} \circ L_{\tilde{b}_k}(F).$$

We see that, when $\alpha \equiv 1$, the operator $Q(F)$ is a sum of squares of differential operators given by vector fields which are generators of certain rotations in \mathbb{R}^6 . Since the Fisher information is invariant under rotations, it follows that $I(F)$ is unchanged if one flows F by $L_{\tilde{b}_k} F$.

LEMMA 5.1. *For any smooth positive function $F: \mathbb{R}^6 \rightarrow (0, \infty)$ with rapid decay at infinity, we have*

$$\langle I'(F), L_{\tilde{b}_k}(F) \rangle = 0.$$

Proof. The result already follows by the rotational invariance, but we also verify directly by differentiating the integral. We proceed by direct computation:

$$\begin{aligned} \langle I'(F), L_{\tilde{b}_k}(F) \rangle &= \iint_{\mathbb{R}^6} \left(2 \frac{\nabla F \cdot \nabla L_{\tilde{b}_k}(F)}{F} - \frac{|\nabla F|^2}{F^2} L_{\tilde{b}_k}(F) \right) dw dv \\ &= \iint_{\mathbb{R}^6} \left(\frac{\tilde{b}_k \cdot \nabla |\nabla F|^2 + 2 \langle (D\tilde{b}_k) \nabla F, \nabla F \rangle}{F} - \frac{|\nabla F|^2}{F^2} L_{\tilde{b}_k}(F) \right) dw dv \end{aligned}$$

Observe that $D\tilde{b}_k$ is antisymmetric for $k=1, 2, 3$, and thus

$$\langle (D\tilde{b}_k) \nabla F, \nabla F \rangle \equiv 0.$$

Then,

$$\begin{aligned} \langle I'(F), L_{\tilde{b}_k}(F) \rangle &= \iint_{\mathbb{R}^6} \left(\frac{\tilde{b}_k \cdot \nabla |\nabla F|^2}{F} - \frac{|\nabla F|^2}{F^2} \tilde{b}_k \cdot \nabla F \right) dw dv \\ &= \iint_{\mathbb{R}^6} \tilde{b}_k \cdot \nabla \frac{|\nabla F|^2}{F} dw dv = 0. \end{aligned} \quad \square$$

Motivated by Lemma 5.1, we study the first-order transport equation, where we flow F by $L_{\tilde{b}_k}$. For any function $F_0 = F_0(v, w)$, let us consider the initial value problem

$$\begin{aligned} F^k(0, v, w) &= F_0(v, w), \\ F_t^k(t, v, w) &= L_{\tilde{b}_k}(F). \end{aligned} \quad (5.1)$$

LEMMA 5.2. *Let F^k be as in (5.1), then $I(F)$ is constant in t .*

Proof. It follows from Lemma 5.1. \square

LEMMA 5.3. *For any smooth positive function $F_0: \mathbb{R}^6 \rightarrow (0, \infty)$ with rapid decay at infinity, if F^k solve the problems (5.1) for $k=1, 2, 3$, then we have*

$$Q(F_0)(v, w) = \sum_{k=1}^3 \partial_{tt} F^k(0, v, w).$$

Proof. This is simply the fact that

$$Q = \sum_{k=1}^3 L_{\tilde{b}_k} \circ L_{\tilde{b}_k}$$

as in (4.5). \square

The following result was proved by Villani in [55]. Using our current framework, we are able to provide a short proof.

PROPOSITION 5.4. *The Fisher information is monotone decreasing along the flow of the Landau equation (1.1) when $\alpha \equiv 1$.*

Proof. We have to verify that (3.5) holds in this case. Then, the result follows using Lemma 3.4.

Let $F^k(t, v, w)$ solve (5.1) with initial data $F_0 = F(v, w)$. We differentiate $I(F^k)$ twice using Lemma 5.2, to get that

$$0 = \partial_{tt} I(F^k) = \langle I'(F^k), \partial_{tt} F^k \rangle + \langle I''(F^k) \partial_t F^k, \partial_t F^k \rangle.$$

Since the Fisher information is convex, then

$$\langle I''(F^k) \partial_t F^k, \partial_t F^k \rangle \geq 0.$$

Thus, using Lemma 5.3, we conclude

$$\langle I'(F), Q(F) \rangle = - \sum_{k=1}^3 \langle I''(F) \partial_t F^k, \partial_t F^k \rangle|_{t=0} \leq 0. \quad \square$$

Lemma 5.3 still holds if we use $\sqrt{\alpha} \tilde{b}_k$ instead of \tilde{b}_k for problem (5.1). However, Lemma 5.2 would not hold for any non-constant function α . In order to verify inequality (3.5), we need another idea whenever α is not constant.

6. The Fisher information along layers

As we explained in §4 and recorded in (4.5), the operator Q can be expressed in terms of a composition of transport operators $L_{\sqrt{\alpha}\tilde{b}_k}$. In §5 we saw how this decomposition can be used to show the monotonicity of the Fisher information in the case of Maxwell molecules $\alpha \equiv 1$. In this section we begin the analysis for non-constant α .

Since we will be writing $L_{\sqrt{\alpha}\tilde{b}_k}$ repeatedly, from now on we will use the simpler notation

$$L_k(F) := \sqrt{\alpha} \tilde{b}_k \cdot \nabla F. \quad (6.1)$$

As noted in §5, when $\alpha \equiv 1$ the operator L_k is the generator of a flow by isometries in \mathbb{R}^6 , and therefore it preserves the Fisher information $I(F)$. This is not true any more when α is not constant. We cannot use the analysis in §5 to immediately deduce (3.5).

The main observation for this section is that, since α depends only on $|v-w|$ and the vectors \tilde{b}_k are tangent to these level sets, the flow defining L_k is still an isometry layer-by-layer when we restrict our analysis to the level sets of $|v-w|$. Thus, if we modify the Fisher information to only take into account the components of ∇F that are tangent to these level sets, we obtain a quantity that is indeed preserved by the flow of L_k .

The unit normal to the level set of $|v-w|$ is precisely the vector field n defined in (4.7). The *modified* Fisher information that we want to study is the following:

$$\begin{aligned} I_{\tan}(F) &:= \iint_{\mathbb{R}^6} \left(\frac{|\nabla F|^2}{F} - \frac{(n \cdot \nabla F)^2}{F} \right) dw dv \\ &= \iint_{\mathbb{R}^6} \left(\frac{|\partial_{v_i} F|^2 + |\partial_{w_i} F|^2}{F} - \frac{(v_i - w_i)(v_j - w_j)}{2|v-w|^2} \frac{(\partial_{v_i} - \partial_{w_i})F(\partial_{v_j} - \partial_{w_j})F}{F} \right) dw dv. \end{aligned} \quad (6.2)$$

It is convenient to write I_{\tan} as $I_{\text{sph}} + I_{\text{par}}$ where

$$\begin{aligned} I_{\text{sph}}(F) &:= \iint_{\mathbb{R}^6} \frac{a_{ij}(v-w)}{2|v-w|^2} \cdot \frac{(\partial_{v_i} - \partial_{w_i})F(\partial_{v_j} - \partial_{w_j})F}{F} dw dv, \\ I_{\text{par}}(F) &:= \iint_{\mathbb{R}^6} \frac{|(\partial_{v_i} + \partial_{w_i})F|^2}{2F} dw dv. \end{aligned}$$

We used the notation

$$a_{ij}(z) = |z|^2 \delta_{ij} - z_i z_j.$$

Remark 6.1. The subindices in I_{\tan} , I_{sph} and I_{par} are meant to convey which directions are being included. We already noted that $I_{\tan}(F)$ is the Fisher information of F in the tangential directions to the level sets of $(v, w) \mapsto \alpha(|v-w|)$, and so here we call it the *tangential* (with respect to the level sets of α) *Fisher information*. Meanwhile,

$I_{\text{sph}}(F)$ involves the directions tangent to sphere(v, w): we call it the *spherical Fisher information*. Last but not least, $I_{\text{par}}(F)$ involves the directions in the 3-dimensional linear subspace of \mathbb{R}^6 of “parallel” pairs of velocities $\{(e, e) \in \mathbb{R}^6 : e \in \mathbb{R}^3\}$: we call it the *parallel Fisher Information*.

PROPOSITION 6.2. *Let \tilde{b}_k be the vector fields from (4.3), let $\alpha = \alpha(|v-w|)$ be an arbitrary non-negative function and let L_k be as in (6.1). For any smooth positive function $F: \mathbb{R}^6 \rightarrow (0, \infty)$ with rapid decay at infinity and any $k=1, 2, 3$, we have*

$$\begin{aligned}\langle I'_{\text{par}}(F), L_k(F) \rangle &= 0 \\ \langle I'_{\text{sph}}(F), L_k(F) \rangle &= 0 \\ \langle I'_{\text{tan}}(F), L_k(F) \rangle &= 0.\end{aligned}$$

Moreover,

$$\langle I'_{\text{tan}}(F), Q(F) \rangle = - \sum_{k=1}^3 \langle I''_{\text{tan}}(F) L_k(F), L_k(F) \rangle \leq 0,$$

and similar identities hold for I_{sph} and I_{par} .

Proof. Since $I_{\text{tan}} = I_{\text{par}} + I_{\text{sph}}$, then the result follows for I_{tan} after we prove it for I_{sph} and I_{par} . Let us start with the case of I_{par} . Let us define the unit vectors p_1, p_2 and p_3 in \mathbb{R}^6 as

$$p_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad p_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad p_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

We write I_{par} in terms of these vectors. We have

$$I_{\text{par}}(F) = \sum_{i=1}^3 \iint_{\mathbb{R}^6} \frac{|p_i \cdot \nabla F|^2}{F} dw dv.$$

We observe that $[p_i, \sqrt{\alpha} \tilde{b}_k] = 0$ for $i, k = 1, 2, 3$. Then, using Lemma 4.2, we get that

$$\langle I'_{\text{par}}(F), L_k(F) \rangle = 0.$$

We now move on to I_{sph} . Because of (4.4), we can write I_{sph} as

$$I_{\text{sph}} = \sum_{i=1}^3 \iint_{\mathbb{R}^6} \frac{1}{2|v-w|^2} \frac{|\tilde{b}_i \cdot \nabla F|^2}{F} dw dv.$$

Differentiation in the directions \tilde{b}_i and \tilde{b}_k does not commute when $i \neq k$. Their commutators are easy to compute:

$$[\tilde{b}_1, \tilde{b}_2] = -2\tilde{b}_3, \quad [\tilde{b}_2, \tilde{b}_3] = -2\tilde{b}_1 \quad \text{and} \quad [\tilde{b}_3, \tilde{b}_1] = -2\tilde{b}_2.$$

Taking this into account, we compute $\langle I'_{\text{sph}}(F), L_k(F) \rangle$ using Lemma 4.2:

$$\langle I'_{\text{sph}}(F), L_k(F) \rangle = \sum_{i=1}^3 \iint_{\mathbb{R}^6} \frac{\sqrt{\alpha}}{|v-w|^2} (\tilde{b}_i \cdot \nabla \log F) \cdot ([\tilde{b}_i, \tilde{b}_k] \cdot \nabla \log F) F dw dv = 0.$$

The last identity holds because the two terms where $k \neq i$ cancel out. For example, for $k=1$, we have $[\tilde{b}_2, \tilde{b}_1] = 2\tilde{b}_3$ and $[\tilde{b}_3, \tilde{b}_1] = -2\tilde{b}_2$, which makes the integrand

$$\frac{2\sqrt{\alpha}}{|v-w|^2} F((\tilde{b}_2 \cdot \nabla \log F) \cdot (\tilde{b}_3 \cdot \nabla \log F) - (\tilde{b}_3 \cdot \nabla \log F) \cdot (\tilde{b}_2 \cdot \nabla \log F)) = 0.$$

Therefore, we obtain both

$$\langle I'_{\text{sph}}(F), L_k(F) \rangle = 0 \quad \text{and} \quad \langle I'_{\text{par}}(F), L_k(F) \rangle = 0$$

Adding them, we also obtain

$$\langle I'_{\text{tan}}(F), L_k(F) \rangle = 0.$$

The final identity follows mimicking the proof of Proposition 5.4. \square

7. Using commutators to estimate the remaining directional derivative

In §6 we found a quantity I_{tan} that is monotone decreasing in the direction of Q . The difference between this quantity I_{tan} and the full Fisher information I depends only on the component of ∇F in the single direction n perpendicular to the level sets of $|v-w|$. In this section, we analyze the value of $J := I - I_{\text{tan}}$ and compute its derivative explicitly.

LEMMA 7.1. *Let n be as in (4.7) and $b = \tilde{b}_k$ for one of the vector fields (4.3) with $k=1, 2, 3$. Let J be the functional*

$$J(F) := \iint_{\mathbb{R}^6} \frac{|n \cdot \nabla F|^2}{F} dw dv,$$

and let

$$Q_b(F) := \alpha b \cdot \nabla(b \cdot \nabla F),$$

for some scalar function $\alpha = \alpha(|v-w|)$. Then, following identity holds:

$$\langle J'(F), Q_b(F) \rangle = \int_{\mathbb{R}^6} \left(-2F(n \cdot \nabla(\sqrt{\alpha} b \cdot \nabla \log F))^2 + \frac{(\alpha')^2}{\alpha} (b \cdot \nabla \log F)^2 F \right) dw dv.$$

Proof. The proof is a direct computation. We write it in detail.

Recall that $\operatorname{div}(ab)=0$ and also $\operatorname{div}((n \cdot \nabla \alpha)b)=0$. Moreover,

$$[n, ab] = (n \cdot \nabla \alpha)b.$$

It is good to remember that, since $[n, b]=0$, then

$$n \cdot \nabla(b \cdot \nabla G) = b \cdot \nabla(n \cdot \nabla G)$$

for any function G .

We differentiate and follow the computation:

$$\begin{aligned} \langle J'(F), Q_b(F) \rangle &= \iint_{\mathbb{R}^6} [2(n \cdot \nabla \log F)(n \cdot \nabla Q_b(F)) - (n \cdot \nabla \log F)^2 Q_b(F)] dw dv \\ &= \iint_{\mathbb{R}^6} [2(n \cdot \nabla \log F)(n \cdot \nabla(\alpha b \cdot \nabla(b \cdot \nabla F))) - (n \cdot \nabla \log F)^2 (\alpha b \cdot \nabla(b \cdot \nabla F))] dw dv. \end{aligned}$$

We commute differentiation with respect to n and αb introducing an error term:

$$\begin{aligned} \langle J'(F), Q_b(F) \rangle &= \iint_{\mathbb{R}^6} [2(n \cdot \nabla \log F)(\alpha b \cdot \nabla(n \cdot \nabla(b \cdot \nabla F))) \\ &\quad + 2(n \cdot \nabla \log F)(n \cdot \nabla \alpha)b \cdot \nabla(b \cdot \nabla F)) \\ &\quad - (n \cdot \nabla \log F)^2(\alpha b \cdot \nabla(b \cdot \nabla F))] dw dv. \end{aligned}$$

Integrating by parts, and using that $\operatorname{div} b=0$ and $\operatorname{div}(ab)=0$, we get

$$\begin{aligned} \langle J'(F), Q_b(F) \rangle &= \iint_{\mathbb{R}^6} [-2\alpha b \cdot \nabla(n \cdot \nabla \log F)(n \cdot \nabla(b \cdot \nabla F)) \\ &\quad - 2(n \cdot \nabla \alpha)b \cdot \nabla(n \cdot \nabla \log F)(b \cdot \nabla F) \\ &\quad + \alpha b \cdot \nabla(n \cdot \nabla \log F)^2(b \cdot \nabla F)] dw dv. \end{aligned}$$

We expand the derivative in the last term and observe that

$$n \cdot \nabla(b \cdot \nabla \log F)F = (n \cdot \nabla(b \cdot \nabla F)) - (n \cdot \nabla \log F)(b \cdot \nabla F),$$

from where we get

$$\begin{aligned} \langle J'(F), Q_b(F) \rangle &= \iint_{\mathbb{R}^6} [-2\alpha F(n \cdot \nabla(b \cdot \nabla \log F))^2 - 2(n \cdot \nabla \alpha)b \cdot \nabla \log F(n \cdot \nabla(b \cdot \nabla \log F))F] dw dv. \end{aligned}$$

We complete squares and conclude this is equal to

$$\begin{aligned} &\iint_{\mathbb{R}^6} \left[-2F \left(\sqrt{\alpha} n \cdot \nabla(b \cdot \nabla \log F) + \frac{(n \cdot \nabla \alpha)}{2\sqrt{\alpha}} b \cdot \nabla \log F \right)^2 + \frac{1}{2} \frac{(n \cdot \nabla \alpha)^2}{\alpha} (b \cdot \nabla \log F)^2 F \right] dw dv \\ &= \iint_{\mathbb{R}^6} \left[-2F(n \cdot \nabla(\sqrt{\alpha} b \cdot \nabla \log F))^2 + \frac{(\alpha')^2}{\alpha} (b \cdot \nabla \log F)^2 F \right] dw dv. \end{aligned} \quad \square$$

Remark 7.2. It is possible to compute $\langle J'(F), Q(F) \rangle$ using a method that follows more closely the ideas of §5. For that, we would have to compute $\langle J', L_k \rangle$ (which is no longer zero) in terms of commutators of n and b_k . Then, following the ideas in §5, we may proceed with computing the second derivative in time of $I(F)$ when $F_t=L_k(F)$, and derive the formula in Lemma 7.1.

It is also true that we can rewrite the proofs of Propositions 5.4 and 6.2 using a more direct computation with commutators, like the one presented here for Lemma 7.1.

We believe that the method of §5 is more intuitive than the computation presented in this section. Yet, there might be some value in presenting alternative approaches to the computation.

8. The three distinct diffusion terms

In §6, we analyzed the derivative of I_{\tan} in the direction $Q(F)$ and obtained a negative value in terms of I''_{\tan} . In §7, we analyzed the derivative of J in the direction $Q(F)$ and obtained a negative value in terms of J'' , and a positive error term. In this section, we analyze the derivative of the full Fisher information I in the direction $Q(F)$, which results from adding the estimates we computed in previous sections.

We want to analyze the second derivative of the Fisher information.

LEMMA 8.1. *Let I be the Fisher information of a smooth positive function*

$$F: \mathbb{R}^6 \longrightarrow (0, \infty),$$

as in (3.3). Let $b: \mathbb{R}^6 \rightarrow \mathbb{R}^6$ be any smooth vector field. Then

$$\langle I''(F)(b \cdot \nabla F), (b \cdot \nabla F) \rangle = 2 \iint_{\mathbb{R}^6} F |\nabla(b \cdot \nabla \log F)|^2 dw dv.$$

Proof. We apply Lemma 4.3 with the canonical basis of \mathbb{R}^6 : e_1, e_2, \dots, e_6 . We add the six corresponding identities and obtain the desired result. \square

We apply Lemma 8.1 with $b = \sqrt{\alpha} \tilde{b}_k$ for each $k=1, 2, 3$ and $\alpha = \alpha(|v-w|)$. These vector fields are divergence free for any scalar function α . We obtain

$$\langle I''(F)(\sqrt{\alpha} \tilde{b}_k \cdot \nabla F), (\sqrt{\alpha} \tilde{b}_k \cdot \nabla F) \rangle = 2 \iint_{\mathbb{R}^6} F |\nabla(\sqrt{\alpha} \tilde{b}_k \cdot \nabla \log F)|^2 dw dv.$$

LEMMA 8.2. *Let Q be the linear operator associated to the Landau equation as in (3.1). Let I be the usual Fisher information. Then,*

$$\begin{aligned} \frac{1}{2}\langle I'(F), Q(F) \rangle &= - \sum_{k=1}^3 \iint_{\mathbb{R}^6} |\nabla[\sqrt{\alpha} \tilde{b}_k \cdot \nabla \log F]|^2 F \, dw \, dv \\ &\quad + \sum_{k=1}^3 \iint_{\mathbb{R}^6} \frac{\alpha'(|v-w|)^2}{2\alpha(|v-w|)} (\tilde{b}_k \cdot \nabla \log F)^2 F \, dw \, dv. \end{aligned}$$

Here, \tilde{b}_k are the vector fields defined in (4.3).

Both terms are homogeneous of degree 1 in F . The second term vanishes when α is constant, which corresponds to the monotonicity of the Fisher information in the Maxwell-molecules case.

Proof. We start by recalling that I is almost the same as I_{\tan} defined in §6, except for one extra term

$$I(F) = I_{\tan}(F) + \iint_{\mathbb{R}^6} F |n \cdot \nabla \log F|^2 \, dw \, dv.$$

Here, n be the vector field defined in (4.7).

Recall that, with the vector fields \tilde{b}_k defined in (4.3),

$$Q(F) = \sum_{k=1}^3 \sqrt{\alpha} \tilde{b}_k \cdot \nabla (\sqrt{\alpha} \tilde{b}_k \cdot \nabla F).$$

We established the monotonicity of I_{\tan} along the flow of Q in Proposition 6.2. Moreover, for $L_k(F) = \sqrt{\alpha} \tilde{b}_k \cdot \nabla F$, we computed

$$\langle I'_{\tan}(F), Q(F) \rangle = \sum_{k=1}^3 -\langle I''_{\tan}(F) L_k(F), L_k(F) \rangle. \quad (8.1)$$

Let us write $J := I - I_{\tan}$. That is,

$$J(F) = \iint_{\mathbb{R}^6} F |n \cdot \nabla \log F|^2 \, dw \, dv.$$

We now apply Lemma 7.1 to with each vector field \tilde{b}_k , and add the three identities. We get the following equality:

$$\begin{aligned} \langle J'(F), Q(F) \rangle &= \sum_{k=1}^3 \left(\langle J''(F) L_k(F), L_k(F) \rangle + \iint_{\mathbb{R}^6} \frac{\alpha'(|v-w|)^2}{\alpha(|v-w|)} F (\tilde{b}_k \cdot \nabla \log F)^2 \, dw \, dv \right). \end{aligned} \quad (8.2)$$

Adding (8.1) and (8.2), we get

$$\begin{aligned}
& \langle I'(F), Q(F) \rangle \\
&= \sum_{k=1}^3 \left(\langle I''(F) L_k(F), L_k(F) \rangle + \iint_{\mathbb{R}^6} \frac{\alpha'(|v-w|)^2}{\alpha(|v-w|)} F \left(\tilde{b}_k \cdot \nabla \log F \right)^2 dw dv \right) \\
&= \sum_{k=1}^3 \left(-2 \iint_{\mathbb{R}^6} F |\nabla[\sqrt{\alpha} \tilde{b}_k \cdot \nabla \log F]|^2 dw dv \right. \\
&\quad \left. + \iint_{\mathbb{R}^6} \frac{\alpha'(|v-w|)^2}{\alpha(|v-w|)} F (\tilde{b}_k \cdot \nabla \log F)^2 dw dv \right). \quad \square
\end{aligned}$$

LEMMA 8.3. *Given any differentiable function $G: \mathbb{R}^6 \rightarrow \mathbb{R}$, the following identity holds at every point $(v, w) \in \mathbb{R}^6$:*

$$|\nabla G|^2 = \frac{1}{2} \sum_{i=1}^3 |(\partial_{v_i} + \partial_{w_i})G|^2 + |n \cdot \nabla G|^2 + \frac{1}{2|v-w|^2} \sum_{k=1}^3 |\tilde{b}_k \cdot \nabla G|^2.$$

Here, n is as in (4.7) and \tilde{b}_k are as in (4.3).

Proof. We start from the elementary identities

$$((\partial_{v_i} \pm \partial_{w_i})G)^2 = (\partial_{v_i} G)^2 \pm 2(\partial_{v_i} G)(\partial_{w_i} G) + (\partial_{w_i} G)^2.$$

These can be added together, resulting in

$$|\nabla G|^2 = \sum_{i=1}^3 [(\partial_{v_i} G)^2 + (\partial_{w_i} G)^2] = \frac{1}{2} \sum_{i=1}^3 [((\partial_{v_i} + \partial_{w_i})G)^2 + ((\partial_{v_i} - \partial_{w_i})G)^2].$$

The sum of the squares $((\partial_{v_i} - \partial_{w_i})G)^2$ is simply $|\nabla_v G - \nabla_w G|^2$, and we have

$$|\nabla_v G - \nabla_w G|^2 = |\Pi(v-w)(\nabla_v G - \nabla_w G)|^2 + \left(\frac{v-w}{|v-w|} \cdot (\nabla_v G - \nabla_w G) \right)^2,$$

where $\Pi(v-w)$ denotes the orthogonal projector from \mathbb{R}^3 onto the space perpendicular to $v-w$, which we already introduced in Remark 4.1. The formula for Π in Remark 4.1 says that

$$|\Pi(v-w)(\nabla_v G - \nabla_w G)|^2 = \frac{1}{|v-w|^2} \sum_{k=1}^3 (b_k \cdot (\nabla_v G - \nabla_w G))^2 = \frac{1}{|v-w|^2} \sum_{k=1}^3 (\tilde{b}_k \cdot \nabla G)^2.$$

From the definition of n in (4.7), it follows that

$$\frac{1}{2|v-w|^2} ((\nabla_v G - \nabla_w G) \cdot (v-w))^2 = |n \cdot \nabla G|^2,$$

and the lemma is proved. \square

Based on Lemma 8.3, we may decompose the second derivative I'' in the direction of $\sqrt{\alpha} \tilde{b}_k \cdot \nabla F$ as the sum of three terms:

$$\begin{aligned} |\nabla(\sqrt{\alpha} \tilde{b}_k \cdot \nabla \log F)|^2 &= \frac{1}{2} \sum_{i=1}^3 |(\partial_{v_i} + \partial_{w_i})(\sqrt{\alpha} \tilde{b}_k \cdot \nabla \log F)|^2 \\ &\quad + |n \cdot \nabla(\sqrt{\alpha} \tilde{b}_k \cdot \nabla \log F)|^2 \\ &\quad + \frac{1}{2|v-w|^2} \sum_{i=1}^3 |\tilde{b}_i \cdot \nabla(\sqrt{\alpha} \tilde{b}_k \cdot \nabla \log F)|^2. \end{aligned}$$

We multiply this identity by F , integrate over \mathbb{R}^6 and add up the result for $k=1, 2, 3$. This, combined with Lemma 8.2, leads to the decomposition

$$\frac{1}{2} \langle I'(F), Q(F) \rangle = -D_{\text{parallel}} - D_{\text{radial}} - D_{\text{spherical}} + \sum_{k=1}^3 \iint_{\mathbb{R}^6} \frac{(\alpha')^2}{2\alpha} F |\tilde{b}_k \cdot \nabla \log F|^2 dw dv,$$

where

$$\begin{aligned} D_{\text{parallel}} &:= \frac{1}{2} \sum_{i,j=1}^3 \iint_{\mathbb{R}^6} \alpha(|v-w|) F |(\partial_{v_i} + \partial_{w_i}) \tilde{b}_j \cdot \nabla \log F|^2 dw dv, \\ D_{\text{radial}} &:= \sum_{i=1}^3 \iint_{\mathbb{R}^6} F |n \cdot \nabla(\sqrt{\alpha} \tilde{b}_i \cdot \nabla \log F)|^2 dw dv, \\ D_{\text{spherical}} &:= \sum_{i,j=1}^3 \iint_{\mathbb{R}^6} \frac{\alpha}{2|v-w|^2} F |\tilde{b}_i \cdot \nabla(\tilde{b}_j \cdot \nabla \log F)|^2 dw dv. \end{aligned} \tag{8.3}$$

Let us also write

$$R_{\text{spherical}} := \sum_{k=1}^3 \iint_{\mathbb{R}^6} \frac{\alpha}{|v-w|^2} F |\tilde{b}_k \cdot \nabla \log F|^2 dw dv. \tag{8.4}$$

Note that, since α depends only on $|v-w|$, then $\tilde{b}_k \cdot \nabla \alpha = 0$ and $(\partial_{v_i} + \partial_{w_i})\alpha = 0$. This is the reason why we can pull the factor α outside of the differentiation in the expressions for D_{parallel} and $D_{\text{spherical}}$.

LEMMA 8.4. *The following inequality holds:*

$$\frac{1}{2} \langle I'(F), Q(F) \rangle \leq -D_{\text{parallel}} - D_{\text{radial}} - D_{\text{spherical}} + \sup_{r>0} \left(\frac{r^2 \alpha'(r)^2}{2\alpha(r)^2} \right) R_{\text{spherical}}.$$

Proof. This is a direct consequence of Lemma 8.2 in terms of the new notation and using that the integrand in $R_{\text{spherical}}$ is non-negative. \square

In order to prove Theorem 1.1, we want to control the positive term $R_{\text{spherical}}$ with the negative terms $-D_{\text{parallel}} - D_{\text{radial}} - D_{\text{spherical}}$. We are going to use only $D_{\text{spherical}}$, that involves second derivatives in the directions \tilde{b}_i .

Recall that, starting from any point $(v, w) \in \mathbb{R}^6$, the flow of the vector fields b_1, b_2 and b_3 stays within the sphere (4.6). The inequality $D_{\text{spherical}} \geq \Lambda R_{\text{spherical}}$ will be deduced as a consequence of an elementary and apparently new inequality for functions on the sphere that we present in the next section.

9. An inequality for functions on the sphere

The objective of this section is to prove the following lemma, which will be crucial for the proof of our main result. It is a Poincaré-like inequality involving the derivatives of the logarithm of a function on the sphere S^2 . Properly interpreted (see Remark 9.14 below), it corresponds to some form of the Γ_2 -criterion of Bakry and Emery (see [4, Proposition 5.7.3]) on the projective space. The lemma is about symmetric functions on S^2 , which is the 2-dimensional sphere in \mathbb{R}^3 . Equivalently, it is a result about functions on the projective space $\mathbb{R}P^2$. The value of the constant (which is $\frac{19}{4}$ below) is not optimal. A more precise value for this constant, as well as a generalization to higher dimensions, is obtained by Sehyun Ji in [38].

LEMMA 9.1. *Let $f: S^2 \rightarrow (0, \infty)$ be a C^2 function on the sphere such that*

$$f(\sigma) = f(-\sigma) \quad \text{for all } \sigma \in S^2.$$

Then, the following inequality holds:

$$\sum_{i,j=1}^3 \int_{S^2} f(b_i \cdot \nabla(b_j \cdot \nabla \log f))^2 d\sigma \geq \frac{19}{4} \sum_{i=1}^3 \int_{S^2} f(b_i \cdot \nabla \log f)^2 d\sigma. \quad (9.1)$$

Remark 9.2. Note that, for $\sigma \in S^2$, the vector fields $b_1(\sigma), b_2(\sigma)$ and $b_3(\sigma)$ are perpendicular to σ , and are therefore vectors on the tangent space of S^2 at every point $\sigma \in S^2$. It is useful to compute their Lie brackets.

$$[b_1, b_2] = -b_3, \quad [b_2, b_3] = -b_1 \quad \text{and} \quad [b_3, b_1] = -b_2. \quad (9.2)$$

Remark 9.3. For functions $\phi, \varphi \in C^1(S^2)$, their gradients $\nabla_\sigma \phi$ and $\nabla_\sigma \varphi$ at a $\sigma \in S^2$ are elements of the tangent plane to S^2 . This makes them vectors in \mathbb{R}^3 perpendicular to σ . The directional derivative $b_i \cdot \nabla \varphi$ is an intrinsic differentiation of the function φ on S^2 . Both vectors $b_i(\sigma)$ and $\nabla_\sigma \varphi$ belong to the tangent space of S^2 at σ . Moreover,

one can check that the b_i 's are also divergence free as vector fields on S^2 . We omit the subindex σ for $\nabla = \nabla_\sigma$ in most of this section.

Applying the decomposition (4.2), we observe that

$$b_1 \otimes b_1 + b_2 \otimes b_2 + b_3 \otimes b_3$$

is the orthogonal projector matrix to the tangent space to S^2 at σ . In particular, it is the identity when applied to vectors on that tangent space. We derive the following elementary but useful identity:

$$\sum_{i=1}^3 (b_i(\sigma) \cdot \nabla \phi)(b_i(\sigma) \cdot \nabla \varphi) = \nabla_\sigma \phi \cdot \nabla_\sigma \varphi \quad \text{for all } \phi, \varphi \in C^1(S^2). \quad (9.3)$$

Another useful identity that follows from this, and that we record here for later use, relates the Laplace–Beltrami operator on S^2 with second derivatives in the directions b_i :

$$\sum_{i=1}^3 (b_i \cdot \nabla(b_i \cdot \nabla f)) = \Delta_\sigma f \quad \text{for all } f \in C^2(S^2). \quad (9.4)$$

To see this, we multiply the left-hand side by $\varphi \in S^2$ and integrate by parts over S^2 using that the b_i 's are divergence free. We use (9.3) to obtain

$$\sum_{i=1}^3 \int_{S^2} (b_i \cdot \nabla(b_i \cdot \nabla f)) \varphi \, d\sigma = - \int_{S^2} \nabla_\sigma f \cdot \nabla_\sigma \varphi \, d\sigma.$$

Integrating by parts on the second integral, it follows that

$$\sum_{i=1}^3 \int_{S^2} (b_i \cdot \nabla(b_i \cdot \nabla f)) \varphi \, d\sigma = \int_{S^2} \Delta_\sigma f \varphi \, d\sigma.$$

Since φ is an arbitrary function in $C^2(S^2)$, we have proved (9.4).

The starting point for the proof of Lemma 9.1 is to express both sides of (9.1) in terms of \sqrt{f} , motivated by the well-known observation that the Fisher information of f is four times the \dot{H}^1 norm of \sqrt{f} . Once this is done, the lemma will follow from the Poincaré inequality on S^2 applied to each $(b_j \cdot \nabla_\sigma \sqrt{f})$.

It is easy to write the right-hand side of (9.1) in terms of \sqrt{f} . Using (9.3), we have

$$\sum_{i=1}^3 \int_{S^2} f(b_i \cdot \nabla \log f)^2 \, d\sigma = 4 \sum_{i=1}^3 \int_{S^2} (b_i \cdot \nabla \sqrt{f})^2 \, d\sigma = 4 \int_{S^2} |\nabla_\sigma \sqrt{f}|^2 \, d\sigma. \quad (9.5)$$

As for the left-hand side of (9.1), for the terms with $i=j$, there is a simple trick to obtain a clean bound in terms of a respective integral involving \sqrt{f} . Indeed, consider the case $i=j$. Writing $b=b_i$, we see that

$$\begin{aligned} & \int_{S^2} (b \cdot \nabla(b \cdot \nabla \log f))^2 f \, d\sigma \\ &= 4 \int_{S^2} \left((b \cdot \nabla(b \cdot \nabla \sqrt{f}))^2 - 2 \frac{(b \cdot \nabla(b \cdot \nabla \sqrt{f}))(b \cdot \nabla \sqrt{f})^2}{\sqrt{f}} + \frac{(b \cdot \nabla \sqrt{f})^4}{f} \right) d\sigma. \end{aligned}$$

Integrating by parts the middle term, we arrive at

$$\begin{aligned} \int_{S^2} (b \cdot \nabla(b \cdot \nabla \log f))^2 f \, d\sigma &= 4 \int_{S^2} \left((b \cdot \nabla(b \cdot \nabla \sqrt{f}))^2 + \frac{1}{3} \frac{(b \cdot \nabla \sqrt{f})^4}{f} \right) d\sigma \\ &\geq 4 \int_{S^2} (b \cdot \nabla(b \cdot \nabla \sqrt{f}))^2 d\sigma. \end{aligned} \quad (9.6)$$

We can apply the same trick for all the terms with $i=j$, but it does not yield a useful expression when $i \neq j$. To overcome this, we will consider pure second derivatives in all possible directions below. Let us first introduce, for a given $g \in C^2(S^2)$, the matrices

$$\begin{aligned} (M_g)_{ij} &= \frac{b_i \cdot \nabla(b_j \cdot \nabla g) + b_j \cdot \nabla(b_i \cdot \nabla g)}{2}, \\ (N_g)_{ij} &= \frac{b_i \cdot \nabla(b_j \cdot \nabla g) - b_j \cdot \nabla(b_i \cdot \nabla g)}{2}. \end{aligned} \quad (9.7)$$

That is, $M_g(\sigma)$ and $N_g(\sigma)$ are the symmetric and antisymmetric parts of the matrix

$$(b_i \cdot \nabla(b_j \cdot \nabla g))_{ij}.$$

Note that $M_g(\sigma)$ and $N_g(\sigma)$ are 3×3 matrices, corresponding to the fact that we are working with three vector fields b_1 , b_2 and b_3 . Moreover, the squared norms of M_g and N_g are of direct interest to us, since

$$\sum_{i,j=1}^3 (b_i \cdot \nabla \sigma(b_j \cdot \nabla \sigma \log f))_{ij}^2 = \sum_{i,j=1}^3 (M_{\log f})_{ij}^2 + (N_{\log f})_{ij}^2 = \|M_{\log f}(\sigma)\|^2 + \|N_{\log f}(\sigma)\|^2.$$

Therefore, we want to understand the integrals of

$$\|M_{\log f}(\sigma)\|^2 f \quad \text{and} \quad \|N_{\log f}(\sigma)\|^2 f.$$

Analyzing the latter is straightforward, since the components of $N_g(\sigma)$ are given by the derivatives $b_i \cdot \nabla g$.

LEMMA 9.4. *Let $g: S^2 \rightarrow \mathbb{R}$ be any C^2 function and $N_g(\sigma)$ be as in (9.7). Then,*

$$\|N_g(\sigma)\|^2 = \frac{1}{2} \sum_{i=1}^3 (b_i \cdot \nabla g)^2.$$

Proof. The entries $(N_g)_{ij}$ correspond to the Lie brackets of b_1 , b_2 and b_3 :

$$(N_g)_{ij} = \frac{1}{2} [b_i, b_j].$$

Thus,

$$\|N\|^2 = \sum_{i,j=1}^3 N_{ij}^2 = \frac{1}{4} \sum_{i,j=1}^3 ([b_i, b_j] \cdot \nabla g)^2.$$

Note that each non-zero term appears twice in the sum. Using (9.2), we get

$$\|N\|^2 = \frac{1}{2} \sum_{i=1}^3 (b_i \cdot \nabla g)^2. \quad \square$$

Lemma 9.4 already gives us a fraction of the inequality of Lemma 9.1. Indeed, by Lemma 9.4, we get

$$\begin{aligned} & \sum_{i,j=1}^3 \int_{S^2} (b_i \cdot \nabla (b_j \cdot \nabla \log f))^2 f \, d\sigma \\ &= \int_{S^2} \|M_{\log f}(\sigma)\|^2 f(\sigma) \, d\sigma + \frac{1}{2} \int_{S^2} \sum_{i=1}^3 (b_i \cdot \nabla \log f)^2 f(\sigma) \, d\sigma. \end{aligned}$$

To obtain our desired inequality (9.1), it remains to understand the integral of $\|M_{\log f}\|^2 f$. Recall that a symmetric matrix $A \in \mathbb{R}^{d \times d}$ is uniquely determined by the quadratic form $\langle Ae, e \rangle$. Therefore, we may try to estimate its norm $\|A\|^2$ in terms of the values of $\langle Ae, e \rangle^2$, as e ranges over all vectors in S^2 . Likewise, in order to compute $\|M_{\log f}(\sigma)\|^2$, we may average the values of second derivatives of $\log f$ along all possible directions and properly normalize it. We need a practical way to determine all possible directional derivatives from a point $\sigma \in S^2$. For any $e \in S^2$, we define the vector field b_e using the cross product:

$$b_e(\sigma) = \sigma \times e.$$

Note that b_{e_1} , b_{e_2} and b_{e_3} conveniently correspond to b_1 , b_2 and b_3 from (4.1). Using these vector fields, we want to recover the quantities on the left- and right-hand sides of Lemma 9.1. It is easy for the right-hand side using the next lemma.

LEMMA 9.5. *For any C^1 function $g: S^2 \rightarrow \mathbb{R}$, the following equality holds at every point $\sigma \in S^2$:*

$$\sum_{i=1}^3 (b_i \cdot \nabla g(\sigma))^2 = \frac{3}{\omega_2} \int_{e \in S^2} (b_e \cdot \nabla g(\sigma))^2 de.$$

Proof. Note that here $\sigma \in S^2$ is a fixed point and $\nabla g(\sigma)$ is the same vector for all values of $e \in S^2$. Let us write $e = (x_1, x_2, x_3)$. We observe that

$$b_e = x_1 b_1 + x_2 b_2 + x_3 b_3,$$

so

$$\begin{aligned} \int_{e \in S^2} (b_e \cdot \nabla g(\sigma))^2 de &= \int_{e \in S^2} \left(\sum_{i=1}^3 x_i b_i \cdot \nabla g \right)^2 de \\ &= \int_{e \in S^2} \sum_{i,j=1}^3 (x_i b_i \cdot \nabla g)(x_j b_j \cdot \nabla g) de. \end{aligned}$$

Note that $b_i = b_i(\sigma)$ and $\nabla g = \nabla g(\sigma)$ are independent of the variable of integration $e \in S^2$. Since $x_i x_j$ integrates to zero on S^2 unless $i = j$, we get

$$\begin{aligned} \int_{e \in S^2} (b_e \cdot \nabla g(\sigma))^2 de &= \int_{e \in S^2} \sum_{i=1}^3 x_i^2 (b_i \cdot \nabla g)^2 de \\ &= \left(\int_{e \in S^2} x_1^2 de \right) \sum_{i=1}^3 (b_i \cdot \nabla g)^2 \\ &= \frac{\omega_2}{3} \sum_{i=1}^3 (b_i \cdot \nabla g)^2. \quad \square \end{aligned}$$

The following is the second-order version of Lemma 9.5. It is slightly more complicated.

LEMMA 9.6. *For any C^2 function $g: S^2 \rightarrow \mathbb{R}$ and any $\sigma \in S^2$, we have the identity*

$$2c_1 \|M_g(\sigma)\|^2 + c_1 |\text{trace}(M_g(\sigma))|^2 = \int_{e \in S^2} (b_e(\sigma) \cdot \nabla (b_e(\sigma) \cdot \nabla g(\sigma)))^2 de.$$

Here,

$$c_1 = \int_{S^2} x_1^2 x_2^2 dS.$$

From identity (9.4), we point out that

$$\text{trace}(M_g) = \Delta_\sigma g.$$

Proof. We write $e = (x_1, x_2, x_3)$ and expand like in the proof of Lemma 9.5:

$$\begin{aligned} \int_{e \in S^2} (b_e \cdot \nabla(b_e \cdot \nabla g))^2 de &= \int_{e \in S^2} \sum_{i,j,k,l=1}^3 (x_i b_i \cdot \nabla(x_j b_j \cdot \nabla g))(x_k b_k \cdot \nabla(x_l b_l \cdot \nabla g)) de \\ &= \sum_{i,j,k,l=1}^3 \left(\int_{e \in S^2} x_i x_j x_k x_l de \right) (b_i \cdot \nabla(b_j \cdot \nabla g))(b_k \cdot \nabla(b_l \cdot \nabla g)). \end{aligned}$$

The integral factors vanish except when (i, j, k, l) consists of either four equal indices, or two pairs of equal indices. We are left with

$$\begin{aligned} \int_{e \in S^2} (b_e \cdot \nabla(b_e \cdot \nabla g))^2 de &= \left(\int_{e \in S^2} x_1^4 de \right) \sum_{i=1}^3 (b_i \cdot \nabla(b_i \cdot \nabla g))^2 \\ &\quad + \left(\int_{e \in S^2} x_1^2 x_2^2 de \right) \sum_{i \neq j} (b_i \cdot \nabla(b_j \cdot \nabla g))^2 \\ &\quad + \left(\int_{e \in S^2} x_1^2 x_2^2 de \right) \sum_{i \neq j} (b_i \cdot \nabla(b_j \cdot \nabla g))(b_j \cdot \nabla(b_i \cdot \nabla g)) \\ &\quad + \left(\int_{e \in S^2} x_1^2 x_2^2 de \right) \sum_{i \neq j} (b_i \cdot \nabla(b_i \cdot \nabla g))(b_j \cdot \nabla(b_j \cdot \nabla g)). \end{aligned}$$

We define $c_1 > 0$ to be the (computable⁽³⁾) value of the integral in the last factor

$$c_1 := \frac{1}{3} \int_{e \in S^2} x_1^4 de = \int_{e \in S^2} x_1^2 x_2^2 de.$$

We continue with our computation, getting

$$\begin{aligned} \int_{e \in S^2} (b_e \cdot \nabla(b_e \cdot \nabla g))^2 de &= 3c_1 \sum_{i=1}^3 (b_i \cdot \nabla(b_i \cdot \nabla g))^2 \\ &\quad + c_1 \sum_{i \neq j} (b_i \cdot \nabla(b_j \cdot \nabla g))^2 \\ &\quad + c_1 \sum_{i \neq j} (b_i \cdot \nabla(b_j \cdot \nabla g))(b_j \cdot \nabla(b_i \cdot \nabla g)) \\ &\quad + c_1 \sum_{i \neq j} (b_i \cdot \nabla(b_i \cdot \nabla g))(b_j \cdot \nabla(b_j \cdot \nabla g)) \end{aligned}$$

⁽³⁾ It can be seen by an elementary computation that $c_1 = 4\pi/15$.

$$\begin{aligned}
&= c_1 \sum_{i,j=1}^3 (b_i \cdot \nabla (b_j \cdot \nabla g))^2 \\
&\quad + c_1 \sum_{i,j=1}^3 (b_i \cdot \nabla (b_j \cdot \nabla g)) (b_j \cdot \nabla (b_i \cdot \nabla g)) \\
&\quad + c_1 \left(\sum_{i=1}^3 b_i \cdot \nabla (b_i \cdot \nabla g) \right)^2 \\
&= 2c_1 \|M_g(\sigma)\|^2 + c_1 |\text{trace}(M_g(\sigma))|^2. \quad \square
\end{aligned}$$

LEMMA 9.7. *For any C^2 function $f: S^2 \rightarrow (0, \infty)$ and any $e \in S^2$, the following inequality holds:*

$$\int_{S^2} (b_e \cdot \nabla (b_e \cdot \nabla \log f))^2 f \, d\sigma \geq 4 \int_{S^2} (b_e \cdot \nabla (b_e \cdot \nabla \sqrt{f}))^2 \, d\sigma.$$

Proof. We have already proved this in the discussion leading to (9.6). In this lemma, we state the inequality for the vector fields b_e . \square

LEMMA 9.8. *For any C^2 function $f: S^2 \rightarrow (0, \infty)$, the following inequality holds:*

$$\begin{aligned}
&\int_{S^2} (2\|M_{\log f}(\sigma)\|^2 + |\text{trace}(M_{\log f}(\sigma))|^2) f(\sigma) \, d\sigma \\
&\geq 4 \int_{S^2} (2\|M_{\sqrt{f}}(\sigma)\|^2 + |\text{trace}(M_{\sqrt{f}}(\sigma))|^2) \, d\sigma.
\end{aligned}$$

Proof. We use Lemma 9.6 to rewrite the left-hand side as follows:

$$\begin{aligned}
&\int_{S^2} (2\|M_{\log f}(\sigma)\|^2 + |\text{trace}(M_{\log f}(\sigma))|^2) f(\sigma) \, d\sigma \\
&= \frac{1}{c_1} \int_{S^2} \int_{S^2} (b_e \cdot \nabla (b_e \cdot \nabla \log f(\sigma)))^2 f(\sigma) \, de \, d\sigma.
\end{aligned}$$

We apply Lemma 9.7 for each individual direction $e \in S^2$ and see that the integral on the right is no smaller than

$$\frac{1}{c_1} \int_{S^2} \int_{S^2} (b_e \cdot \nabla (b_e \cdot \nabla \log f(\sigma)))^2 f(\sigma) \, de \, d\sigma \geq \frac{4}{c_1} \int_{S^2} \int_{S^2} (b_e \cdot \nabla (b_e \cdot \nabla \sqrt{f}))^2 \, d\sigma \, de.$$

The desired inequality then follows by applying Lemma 9.6 again to rewrite this last integral as

$$\frac{4}{c_1} \int_{S^2} \int_{S^2} (b_e \cdot \nabla (b_e \cdot \nabla \sqrt{f}))^2 \, d\sigma \, de = 4 \int_{S^2} (2\|M_{\sqrt{f}}(\sigma)\|^2 + |\text{trace}(M_{\sqrt{f}}(\sigma))|^2) \, d\sigma.$$

Combining the last three displayed formulas, the proof of the lemma is completed. \square

Due to Lemma 9.8, one can express the integral of $\|M_{\log f}\|^2 f$ in terms of that of $\|M_{\sqrt{f}}\|^2$ plus some extra terms involving their respective traces. The next two lemmas will use this to obtain a further inequality, one without the inconvenient trace terms.

LEMMA 9.9. *Let $A \in \mathbb{R}^{3 \times 3}$ and let M and N be its symmetric and anti-symmetric parts, that is*

$$M = \frac{A+A^T}{2} \quad \text{and} \quad N = \frac{A-A^T}{2}.$$

Assume that $\text{rank } A \leq 2$. Then,

$$\|M\|^2 \geq \frac{1}{2} |\text{trace}(M)|^2.$$

Proof. The quantities $\|M\|$ and $\text{trace}(A)$ (which is the same as $\text{trace}(M)$) are invariant by orthonormal changes of coordinates. Since $\text{rank } A \leq 2$, it has a zero eigenvector. We can pick an orthonormal basis that starts with this eigenvector, so that $A_{1,1}=0$. In this basis,

$$\|M\|^2 \geq A_{2,2}^2 + A_{3,3}^2 \geq \frac{1}{2} (A_{2,2} + A_{3,3})^2 = \frac{1}{2} |\text{trace}(A)|^2 = \frac{1}{2} |\text{trace}(M)|^2. \quad \square$$

As we argue below, one can apply Lemma 9.9 to see that, for every $\sigma \in S^2$,

$$2\|M(\sigma)\|^2 \geq |\text{trace}M(\sigma)|^2.$$

We use this fact for the next lemma.

LEMMA 9.10. *Let $f: S^2 \rightarrow (0, \infty)$ be a C^2 function. Then, the following inequality holds:*

$$\int_{S^2} \|M_{\log f}(\sigma)\|^2 f d\sigma \geq \int_{S^2} (2\|M_{\sqrt{f}}(\sigma)\|^2 + |\text{trace}(M_{\sqrt{f}}(\sigma))|^2) d\sigma.$$

Proof. The matrix $A(\sigma)$ given by

$$A_{ij}(\sigma) = b_i \cdot \nabla (b_j \cdot \nabla \log f)$$

is of rank at most 2, because the vectors b_1 , b_2 and b_3 are linearly dependent at every σ .

From Lemma 9.9, we deduce that

$$4\|M_{\log f}\|^2 \geq 2\|M_{\log f}\|^2 + |\text{trace}(M_{\log f})|^2,$$

and then replace it on the left-hand side of Lemma 9.8. \square

The inequality in Lemma 9.10 is how we have managed to turn the trick in Lemma 9.7 into an inequality relating $M_{\log f}$ to $M_{\sqrt{f}}$ (plus an extra term). With (9.10) at hand, what remains in order to prove Lemma 9.1 is using the integral involving $M_{\sqrt{f}}$ to bound the integrals of $(b_i \cdot \nabla \sqrt{f})^2$.

The following lemma is the version for $M_g(\sigma)$ of the well-known fact that the integral of $|D^2g|^2$ and $|\Delta g|^2$ coincide for any twice differentiable function $g: \mathbb{T}^d \rightarrow \mathbb{R}$.

LEMMA 9.11. *Let $g: S^2 \rightarrow \mathbb{R}$ be any C^2 function. Then, the following identity holds:*

$$\begin{aligned} \int_{S^2} \left(\sum_{i=1}^3 b_i \cdot \nabla (b_i \cdot \nabla g) \right)^2 d\sigma &= \int_{S^2} \|M_g(\sigma)\|^2 d\sigma + \frac{1}{2} \int_{S^2} \sum_{i=1}^3 (b_i \cdot \nabla g)^2 d\sigma \\ &= \sum_{i,j=1}^3 \int_{S^2} (b_i \cdot \nabla (b_j \cdot \nabla g))^2 d\sigma. \end{aligned}$$

Proof. We integrate by parts starting from the left-hand side, getting

$$\begin{aligned} I := \int_{S^2} \left(\sum_{i=1}^3 b_i \cdot \nabla (b_i \cdot \nabla g) \right)^2 d\sigma &= \sum_{i,j=1}^3 \int_{S^2} (b_i \cdot \nabla (b_i \cdot \nabla g)) (b_j \cdot \nabla (b_j \cdot \nabla g)) d\sigma \\ &= \sum_{i,j=1}^3 \int_{S^2} [-(b_i \cdot \nabla g) (b_i \cdot \nabla (b_j \cdot \nabla g))] d\sigma. \end{aligned}$$

We introduce commutators to switch the order of differentiation:

$$I = \sum_{i,j=1}^3 \int_{S^2} [-(b_i \cdot \nabla g) (b_j \cdot \nabla (b_i \cdot \nabla (b_j \cdot \nabla g))) - (b_i \cdot \nabla g) ([b_i, b_j] \cdot \nabla (b_j \cdot \nabla g))] d\sigma.$$

We integrate by parts again:

$$I = \sum_{i,j=1}^3 \int_{S^2} [(b_j \cdot \nabla (b_i \cdot \nabla g)) (b_i \cdot \nabla (b_j \cdot \nabla g)) - (b_i \cdot \nabla g) ([b_i, b_j] \cdot \nabla (b_j \cdot \nabla g))] d\sigma.$$

We use that the antisymmetric part of the second derivatives correspond to differentiation along commutators $[b_i, b_j]$, getting

$$I = \sum_{i,j=1}^3 \int_{S^2} \left(|(M_g)_{ij}|^2 - \frac{([b_i, b_j] \cdot \nabla g)^2}{4} - (b_i \cdot \nabla g) ([b_i, b_j] \cdot \nabla (b_j \cdot \nabla g)) \right) d\sigma.$$

We now apply another commutator to the second term and integrate by parts again:

$$\begin{aligned} I = \sum_{i,j=1}^3 \int_{S^2} &\left(|(M_g)_{ij}|^2 - \frac{([b_i, b_j] \cdot \nabla g)^2}{4} - (b_i \cdot \nabla g) ([b_i, b_j] \cdot \nabla (b_j \cdot \nabla g)) \right. \\ &\left. + b_j \cdot \nabla (b_i \cdot \nabla g) ([b_i, b_j] \cdot \nabla g) \right) d\sigma. \end{aligned}$$

Note that $[[b_i, b_j], b_j] = -b_i$ whenever $i \neq j$. Moreover, exchanging i with j in the last term, the symmetric part of $b_j \cdot \nabla(b_i \cdot \nabla g)$ cancels out and we are left with, yielding

$$\begin{aligned} I &= \sum_{i,j=1}^3 \int_{S^2} \left(|(M_g)_{ij}|^2 - \frac{([b_i, b_j] \cdot \nabla g)^2}{4} + \delta_{i \neq j} (b_i \cdot \nabla g)^2 - \frac{([b_i, b_j] \cdot \nabla g)^2}{2} \right) d\sigma \\ &= \int_{S^2} \|M_g\|^2 d\sigma + \frac{1}{2} \int_{S^2} \sum_{i=1}^3 (b_i \cdot \nabla g)^2 d\sigma. \end{aligned}$$

The second identity follows from Lemma 9.4. \square

The last ingredient needed for the proof of Lemma 9.1 is a Poincaré inequality on the sphere.

LEMMA 9.12. *Let $g: S^2 \rightarrow \mathbb{R}$. Assume that $g(\sigma) = g(-\sigma)$ for all $\sigma \in S^2$, and that g has average zero on S^2 . The following inequality holds:*

$$\sum_{i=1}^3 \int_{S^2} (b_i \cdot \nabla_\sigma g)^2 d\sigma \geq 6 \int_{S^2} g^2 d\sigma.$$

Proof. The identity (9.3) says that

$$|\nabla_\sigma g|^2 = \sum_{i=1}^3 (b_i \cdot \nabla_\sigma g)^2,$$

so the desired inequality is equivalent to

$$\int_{S^2} |\nabla_\sigma g|^2 d\sigma \geq 6 \int_{S^2} g^2 d\sigma.$$

This inequality amounts to an elementary observation about the eigenvalues and eigenfunctions of $-\Delta_\sigma$, which are well understood. The eigenfunctions of $-\Delta_\sigma$ correspond to spherical harmonics. The first three eigenvalues are 0, 2 and 6. They correspond to constant functions, first-order spherical harmonics, and second-order spherical harmonics. Since g has average zero and $g(\sigma) = g(-\sigma)$, then it is orthogonal in $L^2(S^2)$ to the eigenspaces corresponding to the first and second eigenfunctions. We obtain the inequality of the lemma because the third eigenvalue of $-\Delta_\sigma$ equals 6. \square

LEMMA 9.13. *Let $f: S^2 \rightarrow (0, \infty)$ be C^2 and even (i.e. $f(\sigma) = f(-\sigma)$), and M_g be as in (9.7). The following inequality holds:*

$$\int_{S^2} \|M_{\log f}(\sigma)\|^2 f d\sigma \geq \frac{17}{4} \sum_{i=1}^3 \int_{S^2} |b_i \cdot \nabla \log f|^2 f d\sigma.$$

Proof. We start from applying Lemma 9.10:

$$\begin{aligned} & \int_{S^2} \|M_{\log f}(\sigma)\|^2 f d\sigma \\ & \geq \int_{S^2} (2\|M_{\sqrt{f}}(\sigma)\|^2 + |\text{trace} M_{\sqrt{f}}(\sigma)|^2) d\sigma \\ & = \int_{S^2} \left(2 \left(\sum_{i,j=1}^3 (b_i \cdot \nabla (b_j \cdot \nabla \sqrt{f}))^2 \right) - 2\|N_{\sqrt{f}}(\sigma)\|^2 + |\text{trace} M_{\sqrt{f}}(\sigma)|^2 \right) d\sigma. \end{aligned}$$

We use Lemma 9.4 to compute the middle term that involves $N_{\sqrt{f}}$:

$$\begin{aligned} & \int_{S^2} \|M_{\log f}(\sigma)\|^2 f d\sigma \\ & \geq \int_{S^2} \left(2 \left(\sum_{i,j=1}^3 (b_i \cdot \nabla (b_j \cdot \nabla \sqrt{f}))^2 \right) + |\text{trace}(M_{\sqrt{f}}(\sigma))|^2 \right) d\sigma - \frac{1}{4} \int_{S^2} \sum_{i=1}^3 (b_i \cdot \nabla \log f)^2 f d\sigma. \end{aligned}$$

We apply Lemma 9.11 to $g = \sqrt{f}$ to replace the second term:

$$\int_{S^2} \|M_{\log f}(\sigma)\|^2 f d\sigma \geq \int_{S^2} 3 \left(\sum_{i,j=1}^3 (b_i \cdot \nabla (b_j \cdot \nabla \sqrt{f}))^2 \right) d\sigma - \frac{1}{4} \int_{S^2} \sum_{i=1}^3 (b_i \cdot \nabla \log f)^2 f d\sigma.$$

Applying Lemma 9.12 to $g = b_j \cdot \nabla \sqrt{f}$, we get

$$\begin{aligned} \int_{S^2} \|M_{\log f}(\sigma)\|^2 f d\sigma & \geq 18 \int_{S^2} \sum_{j=1}^3 (b_j \cdot \nabla \sqrt{f})^2 d\sigma - \frac{1}{4} \int_{S^2} \sum_{i=1}^3 (b_i \cdot \nabla \log f)^2 f d\sigma \\ & \geq \left(\frac{18}{4} - \frac{1}{4} \right) \int_{S^2} \sum_{i=1}^3 (b_i \cdot \nabla \log f)^2 f d\sigma. \end{aligned} \quad \square$$

Proof of Lemma 9.1. We combine Lemma 9.4 with Lemma 9.13. \square

Remark 9.14. The quantities involved in this section correspond to intrinsic geometric objects on the sphere S^2 . The identity (9.4) says that, for any C^2 function $g: S^2 \rightarrow \mathbb{R}$,

$$\sum_{i=1}^3 b_i \cdot \nabla (b_i \cdot \nabla g) = \Delta_\sigma g.$$

It can be seen that other quantities that play a role in this section correspond to the following objects:

$$\|M_g\|^2 = \|\nabla_\sigma^2 g\|^2 + \frac{1}{2} |\nabla_\sigma g|^2 \quad \text{and} \quad \|N_g\|^2 = \frac{1}{2} \sum_{i=1}^3 |b_i \cdot \nabla_\sigma g|^2 = \frac{1}{2} |\nabla_\sigma g|^2.$$

Here ∇_σ and Δ_σ are respectively the gradient and the Laplace-Beltrami operators in S^2 with its standard metric. The proofs in this section can be written in terms of these intrinsic objects. We write it in terms of the vector fields b_i to keep a uniform notation throughout the paper, and to keep the proof more elementary.

Moreover, the integrands in Lemma 9.1 correspond to the *carre-du-champ* operator Γ and the operator Γ_2 in the Bakry–Emery formalism. In fact, from the definition of Γ and Γ_2 one can check that, for any smooth function $g: S^2 \rightarrow \mathbb{R}$,

$$\sum_{i,j=1}^3 (b_i \cdot \nabla g)^2 = \Gamma(g, g) \quad \text{and} \quad \sum_{i,j=1}^3 (b_i \cdot \nabla (b_j \cdot \nabla g))^2 = \Gamma_2(g, g).$$

One only needs to make use of the above expressions for $\nabla_\sigma g$ and $\Delta_\sigma g$ in terms of the b_i 's, expand the resulting terms, and use the commutator identities for all $[b_i, b_j]$ to get the needed cancellations. For the reader's convenience, we recall the definition of the operators Γ and Γ_2 associated to the Laplacian Δ_σ :

$$\begin{aligned} \Gamma(f, g) &:= \frac{\Delta_\sigma(fg) - f(\Delta_\sigma g) - (\Delta_\sigma f)g}{2}, \\ \Gamma_2(f, g) &:= \frac{\Delta_\sigma(\Gamma(f, g)) - \Gamma(f, \Delta_\sigma g) - \Gamma(\Delta_\sigma f, g)}{2}. \end{aligned}$$

Remark 9.15. The optimal constant for the Γ_2 criterion on the sphere is well known to be equal to 2 (see [4, §5.7]), which would not suffice to prove our main result in the case of Coulomb potentials. We get an advantage, due to the symmetry assumption $f(\sigma) = f(-\sigma)$ that is effectively equivalent to consider functions on the projective space $\mathbb{R}P^2$. A discussion of the best known constants in this setting, and an improvement over the constant of Lemma 9.1 is given in [38].

10. The monotonicity of the Fisher information

In this section we complete the proof of Theorem 1.1 as a consequence of the results in the previous sections.

LEMMA 10.1. *Let $F: \mathbb{R}^6 \rightarrow (0, \infty)$ be a smooth function with rapid decay at infinity such that $F(v, w) = F(w, v)$. Let $D_{\text{spherical}}$ and $R_{\text{spherical}}$ be the quantities defined in (8.3) and (8.4), respectively. Then, the following inequality holds:*

$$D_{\text{spherical}} \geq \frac{19}{2} R_{\text{spherical}}.$$

Proof. Let us parametrize v and w in the following way: we let $z := \frac{1}{2}(v+w) \in \mathbb{R}^3$ and we write $\frac{1}{2}(v-w) = r\sigma$ for $r \in [0, \infty)$ and $\sigma \in S^2$. In these coordinates,

$$v = z + r\sigma \quad \text{and} \quad w = z - r\sigma. \quad (10.1)$$

The Jacobian of this change of variables corresponds to

$$dv dw = 8r^2 d\sigma dz dr.$$

Note that $|v-w|=2r$ and $b_k(v-w)=2rb_k(\sigma)$. Moreover, for any vector b perpendicular to σ , we can think of b as a vector on the tangent space of S^2 at σ and we observe that, for any function $G: \mathbb{R}^6 \rightarrow \mathbb{R}$, one has

$$b \cdot \nabla_\sigma G = rb \cdot (\nabla_v - \nabla_w)G.$$

In particular, with $G = \log F$, it leads to

$$\tilde{b}_k(v-w) \cdot \nabla \log F = 2rb_k(\sigma) \cdot (\nabla_v - \nabla_w) \log F = 2b_k(\sigma) \cdot \nabla_\sigma \log F.$$

The symmetry condition $F(v, w) = F(w, v)$ translates to $F(z, r, \sigma) = F(z, r, -\sigma)$ in terms of the new variables.

We rewrite $D_{\text{spherical}}$ and $R_{\text{spherical}}$ in these coordinates:

$$\begin{aligned} D_{\text{spherical}} &= \sum_{i,j=1}^3 \int_{\mathbb{R}^3} \int_0^\infty \int_{S^2} \frac{\alpha(2r)}{8r^2} 16 |b_i(\sigma) \cdot \nabla_\sigma (b_j(\sigma) \cdot \nabla_\sigma \log F)|^2 F (8r^2 d\sigma dr dz) \\ &= 16 \int_{\mathbb{R}^3} \int_0^\infty \alpha(2r) \left(\sum_{i,j=1}^3 \int_{S^2} |b_i(\sigma) \cdot \nabla_\sigma (b_j(\sigma) \cdot \nabla_\sigma \log F)|^2 F d\sigma \right) dr dz, \\ R_{\text{spherical}} &= \sum_{k=1}^3 \int_{\mathbb{R}^3} \int_0^\infty \int_{S^2} \frac{\alpha(2r)}{4r^2} 4 (b_k(\sigma) \cdot \nabla_\sigma \log F)^2 F (8r^2 d\sigma dr dz) \\ &= 8 \int_{\mathbb{R}^3} \int_0^\infty \alpha(2r) \left(\sum_{k=1}^3 \int_{S^2} (b_k(\sigma) \cdot \nabla_\sigma \log F)^2 F d\sigma \right) dr dz. \end{aligned}$$

For each value of $z \in \mathbb{R}^3$ and $r \in (0, \infty)$, the inner integrals with respect to σ (the ones inside the parenthesis) satisfy the inequality of Lemma 9.1. Therefore, we conclude that $\frac{1}{2}D_{\text{spherical}} \geq \frac{19}{4}R_{\text{spherical}}$. Equivalently, $D_{\text{spherical}} \geq \frac{19}{2}R_{\text{spherical}}$. \square

PROPOSITION 10.2. *If $\alpha(r) \geq 0$ is any interaction potential such that, for all $r > 0$,*

$$\frac{r|\alpha'(r)|}{\alpha(r)} \leq \sqrt{19},$$

then (3.5) holds for any smooth function $F: \mathbb{R}^6 \rightarrow (0, \infty)$ with rapid decay at infinity such that $F(v, w) = F(w, v)$.

Proof. Combine Lemma 8.4 with Lemma 10.1. \square

Theorem 1.1 is essentially proved already.

Proof of Theorem 1.1. Note that Proposition 10.2 tells us that (3.5) holds whenever $f: \mathbb{R}^3 \rightarrow (0, \infty)$ is smooth, well-behaved at infinity and strictly positive. Therefore, the right-hand side in Lemma 3.4 is non-positive and Theorem 1.1 follows in that case.

If f has vacuum regions, or if its tails are not sufficiently well-behaved, we can approximate F with a smooth and well-behaved function and pass to the limit. For example, let $\eta: \mathbb{R}^6 \rightarrow [0, 1]$ be a smooth function such that

$$\eta \equiv 1 \text{ in } B_1 \quad \text{and} \quad \eta \equiv 0 \text{ in } \mathbb{R}^6 \setminus B_2.$$

We set

$$F_\varepsilon(v, w) = (F(v, w) + \varepsilon) \eta(\varepsilon v, \varepsilon w) + (1 - \eta(\varepsilon v, \varepsilon w)) e^{-|v|^2}.$$

This function F_ε converges to F as $\varepsilon \rightarrow 0$. Moreover, for each $\varepsilon > 0$, $F_\varepsilon > 0$ and equals a Maxwellian for large values of v and w . We thus know that (3.5) holds for F_ε , and then we deduce that it also holds for F . For a more thorough description of this technical approximation argument, see Appendix B. \square

Remark 10.3. The operator Q has a very simple expression in terms of the variables z, r and σ . From the observation that

$$\tilde{b}_i \cdot \nabla F = 2b_i(\sigma) \cdot \nabla_\sigma F$$

and (9.4), we see that

$$Q(F) = 4\alpha(2r) \Delta_\sigma F.$$

Here, Δ_σ is the Laplace-Beltrami operator with respect to σ on S^2 .

11. The global existence theorem

In this section we explain how Theorem 1.1 is used to obtain Theorem 1.2.

Proof of Theorem 1.2. We recall that the result is already well known in the case $\gamma \in [0, 1]$ from [54], [19]. It is also easy to derive for $\gamma \in [-2, 0]$, as a consequence of the upper bound in [46]. We focus on the case $\gamma < -2$.

Given any initial data as in Theorem 1.2, we construct a solution for a short period of time $[0, T)$, with $T > 0$, using Theorem 2.1. This solution becomes immediately smooth and rapidly decaying. Applying Lemma 2.6, we observe that the Fisher information $i(f)$ also becomes finite for any small $t > 0$.

The continuation criteria in Theorem 2.1 tells us that the solution can be extended for as long as $\|f(t)\|_{L_{k_0}^\infty}$ is bounded. In other words, the solution may blow up at time T only if

$$\lim_{t \rightarrow T} \|f\|_{L_{k_0}^\infty} = +\infty.$$

Also, from Theorem 2.1, during the interval $[0, T)$ the solution is C^∞ , strictly positive and bounded by $C(t)\langle v \rangle^{-k_0}$ (here, $C(t) = \|f(t)\|_{L_{k_0}^\infty}$ might be blowing up as $t \rightarrow T$). Under these conditions, we are able to apply Theorem 1.1 and deduce that its Fisher information is monotone decreasing on the time interval $(0, T)$. In particular, for $t_0 > 0$ small, the Fisher information will remain smaller or equal to $i(f(t_0))$ in the interval $[t_0, T]$. It implies the uniform boundedness of $\|f(t)\|_{L^3}$. Indeed,

$$\|f\|_{L^3} = \|\sqrt{f}\|_{L^6}^2 \leq C \|\sqrt{f}\|_{\dot{H}^1}^2 = \frac{Ci(f)}{4}.$$

From Theorem 2.4, we also know that $\|f(t)\|_{L_q^1}$ remains bounded in $[0, T]$ for all $q > 0$. Interpolating between $\|f(t)\|_{L_q^1}$ and $\|f\|_{L^3}$, we deduce that $\|f(t)\|_{L_k^p}$ is bounded for any p in the range $(1, 3)$ and any large exponent k . We can thus apply Theorem 2.2, since $3/(5+\gamma) < 3$. We deduce that

$$\|f(t)\|_{L^\infty} \leq C_3(1 + (t - t_0)^{-3/2p}),$$

for a constant C_3 that depends only on p and the mass, energy and Fisher information of f_0 . Since the function f is certainly bounded in some short time interval $[0, \delta]$ (from Theorem 2.1), we deduce that

$$\|f(t)\|_{L^\infty} \leq C_4 \quad \text{for } t \in [0, T), \tag{11.1}$$

for a constant C_4 .

We want to apply Theorem 2.3. It propagates a Maxwellian upper bound when the exponent $\beta > 0$ is sufficiently small depending on the mass, energy and entropy of f . This is not a restriction for us because, if $f_0(v) \leq C_0 e^{-\beta|v|^2}$ for some C_0 and $\beta > 0$, then the same inequality also holds with a smaller value of β . We can therefore assume without loss of generality that $\beta > 0$ is small.

Using the upper bound (11.1) and the moment bounds together with Theorem 2.3, we deduce the following uniform Maxwellian upper bound for $t \in [0, T)$:

$$f(t, v) \leq C_5 e^{-\beta|v|^2}.$$

But this means that $\|f\|_{L_k^\infty}$ is uniformly bounded in $[0, T)$ for any exponent k . Thus, the solution can never blow up according to the continuation criteria in Theorem 2.1. \square

The uniqueness of the solution follows from Theorem 2.5. \square

Remark 11.1. Note that the only a-priori estimate on $[0, T]$ used in the proof of Theorem 1.2 that may deteriorate as $T \rightarrow \infty$ is the moment estimate from Theorem 2.4, in the case $\gamma \leq 0$.

Appendix A. The 2-dimensional case

It is slightly simpler to prove a version of Theorem 1.1 in two dimensions, following approximately the same steps in the proof. We state the result here.

THEOREM A.1. *Let $f: [0, T] \times \mathbb{R}^2 \rightarrow [0, \infty)$ be a classical solution to the space-homogeneous Landau equation (1.1). Assume that the interaction potential α satisfies*

$$\frac{r|\alpha'(r)|}{\alpha(r)} \leq 4 \quad \text{for all } r > 0.$$

Then, the Fisher information $i(f)$ is monotone decreasing as a function of time.

The proof of Theorem A.1 follows the same steps as the proof of Theorem 1.1. We sketch the differences here. We only need one vector field b_1 to write $Q(F)$, for a function $F: \mathbb{R}^4 \rightarrow [0, \infty)$. We write

$$b_1(v-w) = \begin{pmatrix} v_2 - w_2 \\ w_1 - v_1 \end{pmatrix} \quad \text{and} \quad \tilde{b}_1 = \begin{pmatrix} b_1 \\ -b_1 \end{pmatrix}.$$

With this notation, we have

$$Q(F) = \alpha \tilde{b}_1 \cdot \nabla(\tilde{b}_1 \cdot \nabla F).$$

Since we have only one vector b_1 instead of the three vectors b_1, b_2 and b_3 , the resulting formulas are simpler and involve no summation. Following the same line of thought as for the 3-dimensional case, we end up with

$$\frac{1}{2} \langle I'(F)Q(F), Q(F) \rangle \leq -D_{\text{parallel}} - D_{\text{radial}} - D_{\text{spherical}} + \sup_{r>0} \left(\frac{r^2 \alpha'(r)^2}{2\alpha(r)^2} \right) R_{\text{spherical}},$$

where

$$D_{\text{parallel}} := \frac{1}{2} \int \alpha(|v-w|) F |(\partial_{v_i} + \partial_{w_i}) \tilde{b}_1 \cdot \nabla \log F|^2 dw dv \quad (\text{A.1})$$

$$D_{\text{radial}} := \int F |a \cdot \nabla(\sqrt{\alpha} \tilde{b}_1 \cdot \nabla \log F)|^2 dw dv \quad (\text{A.2})$$

$$D_{\text{spherical}} := \int \frac{\alpha}{2|v-w|^2} F |\tilde{b}_1 \cdot \nabla(\tilde{b}_1 \cdot \nabla \log F)|^2 dw dv \quad (\text{A.3})$$

$$R_{\text{spherical}} := \int \frac{\alpha}{|v-w|^2} F (\tilde{b}_1 \cdot \nabla \log F)^2 dw dv. \quad (\text{A.4})$$

We still want to control $R_{\text{spherical}}$ with $D_{\text{spherical}}$. It is achieved after the following lemma.

LEMMA A.2. *Let $f: S^1 \rightarrow (0, \infty)$ be a C^2 function on the circle such that $f(\sigma) = f(-\sigma)$ for all $\sigma \in S^1$. Then, the following inequality holds:*

$$\int_{S^1} f(b_1 \cdot \nabla(b_1 \cdot \nabla \log f))^2 d\sigma \geq 4 \int_{S^1} f(b_1 \cdot \nabla \log f)^2 d\sigma.$$

In this case, the factor 4 on the right-hand side is optimal. It is achieved asymptotically for a function $f = 1 + \varepsilon g$ as $\varepsilon \rightarrow 0$, when g is the third eigenfunction of the Laplacian on S^1 .

The proof of Lemma A.2 is significantly easier than Lemma 9.1. Since there is only one direction b_1 , we apply Lemma 9.7 and obtain right away

$$\int_{S^1} f(b_1 \cdot \nabla(b_1 \cdot \nabla \log f))^2 d\sigma \geq 4 \int_{S^1} (b_1 \cdot \nabla(b_1 \cdot \nabla \sqrt{f}))^2 d\sigma$$

Since $f(\sigma) = f(-\sigma)$, then $b_1 \cdot \nabla f$ is orthogonal to the first two eigenspaces. The third eigenvalue of $-\partial_\sigma^2$ in S^1 is equal to 4. Then, we get

$$\int_{S^1} f(b_1 \cdot \nabla(b_1 \cdot \nabla \log f))^2 d\sigma \geq 16 \int_{S^1} (b_1 \cdot \nabla \sqrt{f})^2 d\sigma = 4 \int_{S^1} f(b_1 \cdot \nabla \log f)^2 d\sigma.$$

We conclude that $D_{\text{spherical}} \geq 8R_{\text{spherical}}$, and use it to finish the proof of Theorem A.1.

If we want to carry out the analysis in this paper in \mathbb{R}^d for $d > 3$, we would have to consider $\frac{1}{2}d(d-1)$ vectors b_i 's. For example, in four dimensions, the six vectors would be

$$\begin{aligned} b_1 &= \begin{pmatrix} v_2 - w_2 \\ w_1 - v_1 \\ 0 \\ 0 \end{pmatrix}, & b_2 &= \begin{pmatrix} w_3 - v_3 \\ 0 \\ v_1 - w_1 \\ 0 \end{pmatrix}, & b_3 &= \begin{pmatrix} v_4 - w_4 \\ 0 \\ 0 \\ w_1 - v_2 \end{pmatrix}, \\ b_4 &= \begin{pmatrix} 0 \\ v_3 - w_3 \\ w_2 - v_2 \\ 0 \end{pmatrix}, & b_5 &= \begin{pmatrix} 0 \\ w_4 - v_4 \\ 0 \\ v_2 - w_2 \end{pmatrix}, & b_6 &= \begin{pmatrix} 0 \\ 0 \\ v_4 - w_4 \\ w_3 - v_3 \end{pmatrix}. \end{aligned}$$

There is a result similar to Lemma 9.1 in any dimension. However, the constant factor on the right-hand side depends on the dimension, as well as the range of admissible values of $|r\alpha'(r)/\alpha(r)|$ for the higher-dimensional counterpart of Theorem 1.1. The values we computed in this paper for the 3-dimensional case are probably not optimal. It would require some work to compute the sharp range in arbitrary dimension.

Appendix B. On the decay of the tails of our integrals

Throughout this paper, we work with solutions f that are C^∞ and decay as $|v| \rightarrow \infty$ faster than any algebraic rate. The equation (1.1) is understood in the classical sense. There are several instances where we consider the derivative of the Fisher information and we end up with integrals in \mathbb{R}^6 involving two or three derivatives of $\log f$. All the integrands are homogeneous of degree 1 in f . It is natural to expect them to decay rapidly as $|v| \rightarrow \infty$. However, some justification is required, since $\log f$ and its derivatives have some growth as $|v| \rightarrow \infty$. It is not completely inappropriate (even if admittedly pedantic) to provide a justification that all the integrals in this paper make sense. We describe it in this appendix, so that the reader is not distracted through the main text of the article.

In the proof of Theorem 1.1, we argued that a generic function F can be approximated with a strictly positive function which agrees with a Maxwellian for large velocities. This approximation can be used, whenever necessary, to justify that the lemmas and inequalities throughout this paper apply to a much wider class of functions. In this appendix, we show that the upper and lower Maxwellian bounds introduced in this approximation are propagated in time by the Landau equation. Thus, we show that these solutions to the Landau equation (1.1) will always be well-behaved, C^∞ smooth, strictly positive, and with well-behaved tails for $|v| \rightarrow \infty$. Theorem 2.1 justifies the majority of these statements. The only condition that remains to be justified is that the derivatives of $\log F$ are appropriately bounded for large velocities, so that the tails of the integrals throughout this paper are convergent. In this appendix, we describe the procedure to approximate the whole solution to (1.1) with solutions f^ε for which we verify these bounds.

One sufficient condition that would easily validate all the integral expressions in this article would be when

$$\begin{aligned} f &\lesssim e^{-\beta|v|^2}, \\ (1 + |\nabla \log f| + |D^2 \log f|) &\lesssim e^{\varepsilon|v|^2} \quad \text{for } \varepsilon \ll \beta. \end{aligned}$$

We will show that this condition is satisfied for a general class of initial data f_0 . If f_0 is bounded below and above by a multiple of the same Maxwellian, that is

$$f_0 \approx e^{-\beta|v|^2},$$

then these bounds can be propagated in time following the ideas in [14] for as long as there is a classical solution to the equation. Bounds on the derivatives of $\log f$ follow applying standard parabolic estimates.

Below, we briefly review the propagation of Gaussian bounds following techniques from the literature. It is a completely standard technique, so we only sketch the proofs here.

We explain the propagation of Gaussian bounds in the case $\alpha(r)=r^\gamma$ with $\gamma\in[-3,0]$. In the case $\gamma\in[0,1]$, upper and lower bounds of the same kind are obtained in [19].

PROPOSITION B.1. *Let $f:[0,T]\times\mathbb{R}^3\rightarrow[0,\infty)$ be a solution of (1.1) with $\alpha(r)=r^\gamma$, $\gamma\in[-3,0]$ and initial data f_0 . Assume that, for some $\beta>0$, $\delta_0>0$ and C_0 , one has*

$$\begin{aligned} f_0(v) &\geq \delta_0 e^{-\beta|v|^2}, \\ f(t,v) &\leq C_0 e^{-\beta|v|^2}. \end{aligned}$$

Then, for any $\varepsilon>0$, there is a $\delta_1>0$ such that

$$f(t,v) \geq \delta_1 e^{-(\beta+\varepsilon)|v|^2},$$

for all $t\in[0,T]$ and all $v\in\mathbb{R}^3$.

Sketch of proof. We write the Landau collision operator in non-divergence form:

$$q(f) = \bar{a}_{ij} \partial_{ij} f + \bar{c} f,$$

where

$$\bar{a}_{ij} = \int \alpha(|v-w|) a_{ij}(v-w) f(w) dw, \quad \bar{c} = -\partial_{ij} \bar{a}_{ij} \geq 0.$$

We follow the same idea as in [14, Theorem 4.3], but with a Maxwellian bound from below. We must find a function $\psi(t,v)$ which is a subsolution to

$$\partial_t \psi \leq \bar{a}_{ij} \partial_{ij} \psi + \bar{c} \psi.$$

Following [14], we know that, if $\psi(0,v)\leq f(0,v)$, then we will also have $\psi(t,v)\leq f(t,v)$ for all $t\in[0,T]$ and all $v\in\mathbb{R}^3$. We claim that the function

$$\psi(t,v) = \delta_0 e^{-C_1 t - \varepsilon t |v|^2} e^{-\beta |v|^2}$$

satisfies this differential inequality.

It is well known that the coefficients \bar{a}_{ij} satisfy certain ellipticity bounds. We have, for some constants $\Lambda\geq\lambda>0$,

$$\begin{aligned} \{\bar{a}_{ij}\} &\geq \lambda \langle v \rangle^\gamma ((|v|^2 I - (v \otimes v)) + \langle v \rangle^{-2} (v \otimes v)), \\ \{\bar{a}_{ij}\} &\leq \Lambda \langle v \rangle^\gamma ((|v|^2 I - (v \otimes v)) + \langle v \rangle^{-2} (v \otimes v)). \end{aligned} \tag{B.1}$$

The constant $\lambda > 0$ in the lower bound depends only on the mass, energy and entropy of f_0 . The proof is the same for any value of $\gamma \in [-3, 1]$. It can be found in [46, Lemma 3.1] and [19, Proposition 4]. For a proof of the upper bound, see [14, Lemma 2.1] and the proof of Theorem 2.3 in §2.

Plugging these estimates on our function ψ , we see that

$$\begin{aligned}\partial_t \psi &= -(C_1 + \varepsilon |v|^2) \psi, \\ \bar{a}_{ij} \partial_{ij} \psi + \bar{c} \psi &\geq \bar{a}_{ij} \partial_{ij} \psi \gtrsim -\Lambda \langle v \rangle^{\gamma+2} \psi.\end{aligned}$$

We pick C_1 sufficiently large so that for all $v \in \mathbb{R}^3$ we have

$$C_1 + \varepsilon |v|^2 \gtrsim \Lambda \langle v \rangle^{\gamma+2},$$

and we finish the proof. \square

The derivatives of f can be bounded using standard parabolic estimates.

PROPOSITION B.2. *Let $f: [0, T] \times \mathbb{R}^3 \rightarrow [0, \infty)$ be a solution of (1.1) with $\alpha(r) = r^\gamma$, $\gamma \in [-3, 0]$. Assume that, for some $\beta > 0$ and C_0 ,*

$$f(t, v) \leq C_0 e^{-\beta |v|^2}.$$

Then, for any $\varepsilon > 0$, there is a constant C_1 (depending on T and C_0) such that

$$\begin{aligned}|\nabla_v f(t, v)| &\leq C_1 e^{-(\beta - \varepsilon)|v|^2}, \\ |D_v^2 f(t, v)| &\leq C_1 e^{-(\beta - \varepsilon)|v|^2}.\end{aligned}$$

for all $t \in [\frac{1}{2}T, T]$ and all $v \in \mathbb{R}^3$.

Sketch of proof. The function f satisfies the equation

$$f_t = \bar{a}_{ij} \partial_{ij} f + \bar{c} f.$$

The coefficients a_{ij} satisfy the ellipticity bounds (B.1). Moreover, from the Gaussian upper bound on f , we can deduce that both \bar{a}_{ij} and \bar{c} are Hölder continuous in v . In order to overcome the difficulty that the ellipticity condition degenerates as $|v| \rightarrow \infty$, for every $t_0 > 0$ and every $v_0 \in \mathbb{R}^3$ we use the change of variables T_{v_0} described in [14, §4]. It maps a parabolic ellipsoid around (t_0, v_0) into $(-1, 0] \times B_1$, and the function f into a function \tilde{f} that satisfies a linear parabolic equation whose coefficients are elliptic with parameters uniform with respect to v_0 . We may further rescale it to make the

Hölder norm of the coefficients less than one. Applying the Schauder estimates to this function \tilde{f} , we obtain

$$|D_v^2 \tilde{f}(0)| \leq C \|\tilde{f}\|_{L^\infty((-1,0] \times B_1)} \lesssim e^{-\beta|v_0|^2}.$$

Rewriting the estimate above in terms of the original function f , we obtain, for some computable exponent $m \in \mathbb{N}$,

$$|D_v^2 f(t_0, v_0)| \leq C \langle v_0 \rangle^m e^{-\beta|v_0|^2} \leq C_1 e^{-(\beta-\varepsilon)|v_0|^2}.$$

Some early references to the Schauder estimates depending on the Hölder norm of the coefficients in space only are [11], [40]. \square

Combining the upper bound from Theorem 2.3, the lower bound of Proposition B.1 and the bounds for the derivatives of Proposition B.2, we see that, whenever f is a solution of (1.1) with initial data f_0 such that

$$\delta_0 e^{-\beta|v|^2} \leq f_0(v) \leq C_0 e^{-\beta|v|^2},$$

then for any $\varepsilon > 0$ and $t > 0$ there exists a constant C_2 such that

$$\begin{aligned} |\nabla_v \log f(t, v)| &\leq C_2 e^{\varepsilon|v|^2}, \\ |D_v^2 \log f(t, v)| &\leq C_2 e^{\varepsilon|v|^2}. \end{aligned}$$

Not only do the solutions of the Landau equation decay at large velocities, but so do their derivatives. We can use that to deduce that the coefficients \bar{a}_{ij} are also C^∞ , and deduce estimates like those of Proposition B.2 for higher-order derivatives of f .

The potential growth of $\nabla \log f$ and $D^2 \log f$ is of lower order than the decay of f , as $|v| \rightarrow \infty$. This is enough to conclude that every integrand considered in this paper decays faster than some Gaussian rate as $|v| \rightarrow \infty$. It shows that all our integrals are well defined, and our manipulations (such as integration by parts) are fully justified at least when the initial data f_0 satisfies these Gaussian bounds.

If f_0 is a generic function with arbitrary decay and perhaps some vacuum regions, we may approximate f_0 following the same rule as in the proof of Theorem 1.1. More precisely, let f^ε be the solution to (1.1) with initial data

$$f^\varepsilon(0, v) = (f_0(v) + \varepsilon) \eta(\varepsilon v) + (1 - \eta(\varepsilon v)) e^{-|v|^2}.$$

Here, $\eta: \mathbb{R}^3 \rightarrow [0, 1]$ is a smooth function supported in B_2 such that $\eta \equiv 1$ in B_1 .

As we discussed above, this solution f^ε propagates in time the Gaussian bounds from above and below. For this approximate solution, all our computations are justified

and Theorems 1.1 and 1.2 hold. We deduce that, for every $\varepsilon > 0$, there exists a global-in-time smooth solution whose Fisher information is monotone decreasing. As $\varepsilon \rightarrow 0$, f^ε converges to f , which is the unique (from Theorem 2.5) solution to (1.1) with initial data f_0 . The convergence holds for example in $L^{\infty, k}$ for f , for any exponent k , and in $L^\infty([0, T], H^1(\mathbb{R}^3))$ for \sqrt{f} . We conclude that the Fisher information of an arbitrary solution f of (1.1) is monotone decreasing in time, assuming only that $f_0 \in L^{\infty, k}$ for k as in Theorem 2.1 and $\sqrt{f_0} \in H^1$.

References

- [1] ALONSO, R. J., BAGLAND, V., DESVILLETTES, L. & LODS, B., A priori estimates for solutions to Landau equation under Prodi–Serrin like criteria. *Arch. Ration. Mech. Anal.*, 248 (2024), Paper No. 42, 63 pp.
- [2] ALONSO, R. J., BAGLAND, V. & LODS, B., Uniform estimates on the Fisher information for solutions to Boltzmann and Landau equations. *Kinet. Relat. Models*, 12 (2019), 1163–1183.
- [3] ARSEN'EV, A. A. & PESKOV, N. V., The existence of a generalized solution of Landau's equation. *Zh. Vychisl. Mat. i Mat. Fiz.*, 17 (1977), 1063–1068, 1096 (Russian), English translation in *Comput. Math. Math. Phys.*, 18 (1978), 54–60.
- [4] BAKRY, D., GENTIL, I. & LEDOUX, M., *Analysis and Geometry of Markov Diffusion Operators*. Grundlehren der mathematischen Wissenschaften, 348. Springer, Cham, 2014.
- [5] BEDROSSIAN, J., GUALDANI, M. P. & SNELSON, S., Non-existence of some approximately self-similar singularities for the Landau, Vlasov–Poisson–Landau, and Boltzmann equations. *Trans. Amer. Math. Soc.*, 375 (2022), 2187–2216.
- [6] BOBYLEV, A. V., The method of the Fourier transform in the theory of the Boltzmann equation for Maxwell molecules. *Dokl. Akad. Nauk SSSR*, 225 (1975), 1041–1044 (Russian), English translation in *Sov. Phys. Dokl.*, 20 (1975), 820–822.
- [7] — The theory of the nonlinear spatially uniform Boltzmann equation for Maxwell molecules, in *Mathematical Physics Reviews*, Soviet Sci. Rev. Sect. C: Math. Phys. Rev., 7, pp. 111–233. Harwood Academic Publ., Chur, 1988.
- [8] — Radially symmetric models of the Landau kinetic equation and high energy tails. *J. Stat. Phys.*, 190 (2023), Paper No. 48, 24 pp.
- [9] BOBYLEV, A. V., GAMBA, I. & POTAPENKO, I., On some properties of the Landau kinetic equation. *J. Stat. Phys.*, 161 (2015), 1327–1338.
- [10] BOLLEY, F. & CARRILLO, J. A., Tanaka theorem for inelastic Maxwell models. *Comm. Math. Phys.*, 276 (2007), 287–314.
- [11] BRANDT, A., Interior schauder estimates for parabolic differential- (or difference-) equations via the maximum principle. *Israel J. Math.*, 7 (1969), 254–262.
- [12] BUET, C. & CORDIER, S., Conservative and entropy decaying numerical scheme for the isotropic Fokker–Planck–Landau equation. *J. Comput. Phys.*, 145 (1998), 228–245.
- [13] CABRERA, R., GUALDANI, M. P. & GUILLEN, N., Regularization estimates of the Landau–Coulomb diffusion. *Nonlinear Analysis*, 251 (2025), Paper No. 113695.
- [14] CAMERON, S., SILVESTRE, L. & SNELSON, S., Global a priori estimates for the inhomogeneous Landau equation with moderately soft potentials. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 35 (2018), 625–642.

- [15] CARLEN, E. A., Superadditivity of Fisher's information and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 101 (1991), 194–211.
- [16] DESVILLETTES, L., Entropy dissipation estimates for the Landau equation in the Coulomb case and applications. *J. Funct. Anal.*, 269 (2015), 1359–1403.
- [17] DESVILLETTES, L., GOLDING, W., GUALDANI, M. P. & LOHER, A., Production of the Fisher information for the Landau–Coulomb equation with L^1 initial data. Preprint, 2024. [arXiv:2410.10765 \[math.AP\]](https://arxiv.org/abs/2410.10765).
- [18] DESVILLETTES, L., HE, L.-B. & JIANG, J.-C., A new monotonicity formula for the spatially homogeneous Landau equation with Coulomb potential and its applications. *J. Eur. Math. Soc.*, (2023).
- [19] DESVILLETTES, L. & VILLANI, C., On the spatially homogeneous Landau equation for hard potentials. I. Existence, uniqueness and smoothness. *Comm. Partial Differential Equations*, 25 (2000), 179–259.
- [20] — On the spatially homogeneous Landau equation for hard potentials. II. H -theorem and applications. *Comm. Partial Differential Equations*, 25 (2000), 261–298.
- [21] EL SAFADI, M., Smoothness of weak solutions of the spatially homogeneous Landau equation. *Anal. Appl. (Singap.)*, 5 (2007), 29–49.
- [22] FOURNIER, N., Uniqueness of bounded solutions for the homogeneous Landau equation with a Coulomb potential. *Comm. Math. Phys.*, 299 (2010), 765–782.
- [23] FOURNIER, N. & GUÉRIN, H., Well-posedness of the spatially homogeneous Landau equation for soft potentials. *J. Funct. Anal.*, 256 (2009), 2542–2560.
- [24] GOLDING, W., GUALDANI, M. P. & LOHER, A., Nonlinear regularization estimates and global well-posedness for the Landau–Coulomb equation near equilibrium. *SIAM J. Math. Anal.*, 56 (2024), 8037–8069.
- [25] GOLDING, W. & LOHER, A., Local-in-time strong solutions of the homogeneous Landau–Coulomb equation with L^p initial datum. *La Matematica*, 3 (2024), 337–369.
- [26] GOLSE, F., Partial regularity in time for the Landau equation (with Coulomb interaction), in *From Particle Systems to Partial Differential Equations*, Springer Proc. Math. Stat., 352, pp. 283–300. Springer, Cham, 2021.
- [27] GOLSE, F., GUALDANI, M. P., IMBERT, C. & VASSEUR, A., Partial regularity in time for the space-homogeneous Landau equation with Coulomb potential. *Ann. Sci. Éc. Norm. Supér.*, 55 (2022), 1575–1611.
- [28] GOLSE, F., IMBERT, C., JI, S. & VASSEUR, A. F., Local regularity for the space-homogenous Landau equation with very soft potentials. *J. Evol. Equ.*, 24 (2024), 1–81.
- [29] GRESSMAN, P. T., KRIEGER, J. & STRAIN, R. M., A non-local inequality and global existence. *Adv. Math.*, 230 (2012), 642–648.
- [30] GUALDANI, M. P. & GUILLEN, N., Estimates for radial solutions of the homogeneous Landau equation with Coulomb potential. *Analysis & PDE*, 9 (2016), 1772–1809.
- [31] — On A_p weights and the Landau equation. *Calc. Var. Partial Differential Equations*, 58 (2019), Paper No. 17, 55 pp.
- [32] — Hardy's inequality and the isotropic Landau equation. *J. Funct. Anal.*, 283 (2022), Paper No. 109559, 25 pp.
- [33] GUO, Y., The Landau equation in a periodic box. *Comm. Math. Phys.*, 231 (2002), 391–434.
- [34] HENDERSON, C., SNELSON, S. & TARFULEA, A., Local solutions of the Landau equation with rough, slowly decaying initial data. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 37 (2020), 1345–1377.
- [35] HWANG, H. J. & JANG, J. W., Compactness properties and local existence of weak solutions to the Landau equation. *Proc. Amer. Math. Soc.*, 148 (2020), 5141–5157.

- [36] IMBERT, C., SILVESTRE, L. & VILLANI, C., On the monotonicity of the Fisher information for the Boltzmann equation. Preprint, 2024. [arXiv:2409.01183 \[math.AP\]](https://arxiv.org/abs/2409.01183).
- [37] JI, S., Entropy dissipation estimates for the Landau equation with Coulomb potentials. *Kinet. Relat. Models*, 18 (2025), 368–388.
- [38] —— Bounds for the optimal constant of the Bakry–Émery γ_2 criterion inequality on $\mathbb{R}\mathbb{P}^{d-1}$. Preprint, 2024. [arXiv:2408.13954 \[math.AP\]](https://arxiv.org/abs/2408.13954).
- [39] —— Dissipation estimates of the Fisher information for the Landau equation. Preprint, 2024. [arXiv:2410.09035 \[math.AP\]](https://arxiv.org/abs/2410.09035).
- [40] KNERR, B. F., Parabolic interior Schauder estimates by the maximum principle. *Arch. Rational Mech. Anal.*, 75 (1980/81), 51–58.
- [41] KRIEGER, J. & STRAIN, R. M., Global solutions to a non-local diffusion equation with quadratic non-linearity. *Comm. Partial Differential Equations*, 37 (2012), 647–689.
- [42] LANDAU, L. D., Kinetic equation for the case of Coulomb interaction. *Phys. Z. Sowjetunion*, (1936), 154–164 (Russian), English translation in *Collected Papers of L. D. Landau*, Pergamon Press (1965), 163–170.
- [43] MATTTHES, D. & TOSCANI, G., Variation on a theme by Bobylev and Villani. *C. R. Math. Acad. Sci. Paris*, 350 (2012), 107–110.
- [44] MCKEAN, H. P., Speed of approach to equilibrium for Kac’s caricature of a Maxwellian gas. *Arch. Ration. Mech. Anal.*, 21 (1966), 343–367.
- [45] MENG, F., WANG, H., MIN, L. & JIN, Z., Uniform estimates for the Fisher information of the Landau equation for soft potentials. *J. Math. Anal. Appl.*, 523 (2023), Paper No. 126992, 21 pp.
- [46] SILVESTRE, L., Upper bounds for parabolic equations and the Landau equation. *J. Differential Equations*, 262 (2017), 3034–3055.
- [47] —— Regularity estimates and open problems in kinetic equations, in *A^3N^2M : Approximation, Applications, and Analysis of Nonlocal, Nonlinear Models* (Proceedings of the 50th John H. Barrett Memorial Lectures), pp. 101–148. Springer, Berlin–Heidelberg, 2023.
- [48] SNELSON, S., Global existence for an isotropic modification of the Boltzmann equation. *J. Funct. Anal.*, 286 (2024), Paper No. 110423.
- [49] TANAKA, H., Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete*, 46 (1978), 67–105.
- [50] TOSCANI, G., New a priori estimates for the spatially homogeneous Boltzmann equation. *Contin. Mech. Thermodyn.*, 4 (1992), 81–93.
- [51] TOSCANI, G. & VILLANI, C., On the trend to equilibrium for some dissipative systems with slowly increasing a priori bounds. *J. Stat. Phys.*, 98 (2000), 1279–1309.
- [52] VILLANI, C., Fisher information estimates for Boltzmann’s collision operator. *J. Math. Pures Appl.*, 77 (1998), 821–837.
- [53] —— On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Rational Mech. Anal.*, 143 (1998), 273–307.
- [54] —— On the spatially homogeneous Landau equation for Maxwellian molecules. *Math. Models Methods Appl. Sci.*, 8 (1998), 957–983.
- [55] —— Decrease of the Fisher information for solutions of the spatially homogeneous Landau equation with Maxwellian molecules. *Math. Models Methods Appl. Sci.*, 10 (2000), 153–161.
- [56] —— A review of mathematical topics in collisional kinetic theory, in *Handbook of Mathematical Fluid Dynamics*, Vol. I, pp. 71–305. North-Holland, Amsterdam, 2002.
- [57] —— Fisher information in kinetic theory. Preprint, 2025. [arXiv:2501.00925 \[math.AP\]](https://arxiv.org/abs/2501.00925).
- [58] WU, K.-C., Global in time estimates for the spatially homogeneous Landau equation with soft potentials. *J. Funct. Anal.*, 266 (2014), 3134–3155.

NESTOR GUILLEN
Department of Mathematics
Texas State University
601 University Drive
San Marcos, TX 78666
U.S.A.
nestor@txstate.edu

LUIS SILVESTRE
Department of Mathematics
The University of Chicago
5734 S. University Ave.
Chicago, IL 60637
U.S.A.
luiss@uchicago.edu

Received November 20, 2023

Received in revised form January 10, 2025