On some nonlinear and nonlocal effective equations in kinetic theory and nonlinear optics

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der KIT-Fakultät für Mathematik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

Tobias Ried, M.Sc.

Tag der mündlichen Prüfung: 18.10.2017

1. Referent: Prof. Dr. Dirk Hundertmark
2. Referent: Prof. Dr. Jean-Marie Barbaroux
3. Referent: Prof. Dr. Mathieu Lewin

This document is licensed under a Creative Commons
Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0)
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en

To my parents

'Everything can be described on different levels'. There may be a simpler but coarser description which conveys the point one wants to make, though in some ways it is oversimplified or even wrong. There will be a more accurate and detailed description which however is too complicated to bring one's point into a focus.

Walter E. Thirring[†]

[†] *God's Traces in the Laws of Nature*, in *The Cultural Values of Science*, Pontifical Academy of Sciences, *Scripta Varia* **105** (2003), 362–372.

Abstract

This thesis deals with some nonlinear and nonlocal effective equations arising in kinetic theory and nonlinear optics.

First, it is shown that the homogeneous non-cutoff Boltzmann equation for Maxwellian molecules enjoys strong smoothing properties: In the case of power-law type particle interactions, we prove the Gevrey smoothing conjecture. For Debye-Yukawa type interactions, an analogous smoothing effect is shown. In both cases, the smoothing is exactly what one would expect from an analogy to certain heat equations of the form $\partial_t u = f(-\Delta)u$, with a suitable function f, which grows at infinity, depending on the interaction potential. The results presented work in arbitrary dimensions, including also the one-dimensional Kac-Boltzmann equation.

In the second part we study the entropy decay of certain solutions of the Kac master equation, a probabilistic model of a gas of interacting particles. It is shown that for initial conditions corresponding to N particles in a thermal equilibrium and $M \leq N$ particles out of equilibrium, the entropy relative to the thermal state decays exponentially to a fraction of the initial relative entropy, with a rate that is essentially independent of the number of particles.

Finally, we investigate the existence of dispersion management solitons. Using variational techniques, we prove that there is a threshold for the existence of minimisers of a nonlocal variational problem, even with saturating nonlinearities, related to the dispersion management equation.

Acknowledgements

First and foremost, I want to thank Dirk Hundertmark for accepting me as his PhD student, and for all his guidance and support throughout the years at KIT. I am truly thankful that he introduced me to so many interesting mathematical physics problems, and that he took the time to sometimes sit through hours of calculations just to realise that the constant didn't get smaller... it really stayed constant. Thank you for your trust and for always valuing my opinion!

Secondly, I would like to thank Jean-Marie Barbaroux for co-supervising me, for all the excellent discussions, and the invitations to Toulon. Thank you also for taking the time to carefully read a draft of my thesis, and for all your useful comments and questions.

I am greatly indebted to both of you for making the years of my PhD studies inspiring and pleasurable, both on a scientific and a personal level.

Special thanks go to Mathieu Lewin for agreeing to review this thesis.

I most highly value the collaboration with Semjon Vugalter, his experience, knowledge, and ideas. Without him, the Boltzmann project and this thesis wouldn't have become what they are now.

I also would like to thank Peer Kunstmann for some valuable discussions and his expertise in harmonic analysis, and Roland Schnaubelt for very useful comments and remarks on the dispersion management project and my CRC reports.

For the harmonious and friendly atmosphere I want to thank all my other colleagues in the *Applied Analysis* workgroup: Leo, Johanna, Silvana, Jannis, Nikos, Elena, Tobias^2, Michael^2, Jasmin, Michal, Jonas and Simon. In particular, I want to mention Kaori Nagato-Plum for keeping the place running, always with great joy and happiness, even when things got chaotic.

A big thank you to Vadim Zharnitsky for inviting me to Champaign-Urbana to work on dispersion management solitons, and to Young-Ran Lee as well for the excellent collaboration on the project.

For introducing me to the Kac model and for the wonderful time in Atlanta I would like to express my deepest gratitude towards Michael Loss. Thank you for taking the time to explain so many things to me and all the discussions. Many thanks also to Federico Bonetto and Alissa Geisinger for the great collaboration on the Kac project. It was a pleasure working with you!

I would also like to thank Johannes Eilinghoff for meticulously proof-reading a draft of the manuscript and for making some helpful comments.

Parts of the research contained in my thesis are the result of visits to various research institutions. It is therefore a pleasure to thank the *Centre International de Rencontres Mathématiques* in Luminy and the *Mathematisches Forschungsinstitut Oberwolfach* for their excellent “research in pairs” programmes, as well as the *Université de Toulon*, *Aix-Marseille Univ*, *Georgia Institute of Technology*, and the *University of Illinois at Urbana-Champaign* for their warm hospitality.

For an invitation to the University of Cambridge to discuss Gevrey smoothing for the Boltzmann equation I want to thank Clément Mouhot. I would also like to thank Helge Dietert for organising a talk at the C.A.K.E. seminar in Cambridge, for showing so much interest in my PhD project and for the very interesting discussions.

For the possibility to attend the 4^{th} Heidelberg Laureate Forum I want to thank the scientific committee of the Heidelberg Laureate Forum, the Heidelberg Laureate Forum Foundation, and the Klaus Tschira Stiftung. It was a truly exceptional experience to meet the laureates of such prestigious prizes and all the fellow young scientists.

I am grateful to the Karlsruhe House of Young Scientists (KHYS) for a Research Travel Grant supporting the stay at Georgia Tech. Financial support by the Deutsche Forschungsgemeinschaft (DFG) through CRC 1173 “Wave Phenomena” is gratefully acknowledged.

Finally, I would like to thank all my friends and family for being there when needed and their patience in times when there was more work than expected. In particular, I want to thank Sandra for being part of my life. I am looking forward to all the things to come.

Tobias Ried
September 2017

Contents

List of Figures

CHAPTER 0

Introduction

The goal of this thesis is to study the properties of solutions to effective equations from kinetic theory and non-linear fibre optics. These equations are usually obtained as approximations of the fundamental physical equations in a suitable limit.

The first of these equations is the Boltzmann equation, describing the evolution of a dilute gas of particles that interact by colliding with one another. Devised by Ludwig Boltzmann in 1872, it has given rise to important developments in statistical physics, probability theory and analysis. But even today, due to its enormous richness and complexity, many questions regarding the Boltzmann equation have still not been answered with full mathematical rigour.

An alternative approach due to Mark Kac, based purely upon probabilistic assumptions, has proved to be very effective in the study of convergence of the microscopic dynamics to an effective equation describing the dynamics on a mesoscopic scale (Kac-Boltzmann equation). It is also useful in the study of convergence to equilibrium, especially when one is interested in the rate of convergence. Including an interaction with thermal reservoirs may lead to a better understanding of non-equilibrium stationary states.

The second effective equation considered in this thesis is the dispersion management (DM) equation. Dispersion, the spreading of initially well-localised wave packets, limits the maximal bandwidth of optical communication systems. The dispersion management technique, based on a strong local variation of the dispersion along the cable, has been successfully applied in optical fibre cables to reliably send vast amounts of data over large distances, for instance through transatlantic cables. The equation is nonlinear and highly nonlocal and contains only strongly oscillating terms, which makes the analysis rather challenging.

The results presented in Part I are based upon a collaboration with Prof. Dr. Jean-Marie Barbaroux (Aix-Marseille Univ et Université de Toulon) and Prof. Dr. Dirk Hundertmark (KIT), as well as Dr. Semjon Vugalter (KIT). The contents of Chapter 3, Appendix C, and parts of Chapter 2 have been published in

- J.-M. Barbaroux, D. Hundertmark, T. Ried, and S. Vugalter, Gevrey smoothing for weak solutions of the fully nonlinear homogeneous Boltzmann and Kac equations without cutoff for Maxwellian molecules. *Archive for Rational Mechanics and Analysis* **225** (2017), 601–661. MR 3665667. Zbl 06759780.

The contents of Chapter 4, Appendix D, and parts of Chapter 2 have been published in

- J.-M. Barbaroux, D. Hundertmark, T. Ried, and S. Vugalter, Strong smoothing for the non-cutoff homogeneous Boltzmann equation for Maxwellian molecules with Debye-Yukawa type interaction. *Kinetic and Related Models* **10** (2017), 901–924. MR 3622094. Zbl 06694337.

Part II is the result of a collaboration with Prof. Federico Bonetto, Ph.D. (Georgia Institute of Technology), Alissa Geisinger, M.Sc. (Eberhard Karls Universität Tübingen), and Prof. Dr. Michael Loss (Georgia Institute of Technology), while the author was on a research stay at Georgia Tech. The contents of Chapters 7 and 8 have been submitted to a peer-reviewed journal, with a manuscript available on arXiv,

- F. Bonetto, A. Geisinger, M. Loss, and T. Ried, Entropy decay for the Kac evolution. *Preprint* arXiv 1707.09584.

Part III is based on a collaboration with Prof. Dr. Dirk Hundertmark (KIT), Prof. Young-Ran Lee, Ph.D. (Sogang University), and Prof. Vadim Zharnitsky, Ph.D. (University of Illinois at Urbana-Champaign). The contents of Chapter 10 have been published in

- D. Hundertmark, Y.-R. Lee, T. Ried, and V. Zharnitsky, Dispersion managed solitons in the presence of saturated nonlinearity. *Physica D: Nonlinear Phenomena* **356–357** (2017), 65–69. MR 3689774.

The contents of Chapter 11 have been published in

- D. Hundertmark, Y.-R. Lee, T. Ried, and V. Zharnitsky, Solitary waves in nonlocal NLS with dispersion averaged saturated nonlinearities. *Journal of Differential Equations* (2017), doi:10.1016/j.jde.2017.08.028.

Part I

The Boltzmann Equation

CHAPTER 1

Introduction

The Boltzmann equation [Bol72] is one of the most important equations in the kinetic theory of gases. It describes the time evolution of the probability distribution function $f : \mathbb{R} \times \Omega_x \times \mathbb{R}^d_v \to [0, \infty)$, $(t, x, v) \mapsto f(t, x, v)$ on the phase space $\Omega_x \times \mathbb{R}^d_v$, where $\Omega_x \subset \mathbb{R}^d_x$. If no external forces are acting on the particles, it is given by

$$\partial_t f + v \cdot \nabla_x f = Q(f, f) \tag{1.1}$$

supplemented by suitable boundary conditions on $\partial\Omega_x$.

The Boltzmann collision operator Q models the influence of the collisions between particles on the density f, and will be discussed in more detail in the following section. The left hand side of the *inhomogeneous* Boltzmann equation (1.1), including the classical transport operator $v \cdot \nabla_x$, describes the streaming of particles between collisions.

For simplicity, we shall assume that the gas is *spatially homogeneous*, i.e., $f(t, x, v) = f(t, v)$, and comment on the inhomogeneous case in Chapter 5.

Our goal is to study the regularity of weak solutions of the Cauchy problem

$$\begin{cases} \partial_t f = Q(f, f) \\ f|_{t=0} = f_0 \end{cases} \tag{1.2}$$

for the fully nonlinear homogeneous Boltzmann equation in $d \geq 2$ dimensions, with physically reasonable initial data f_0. Our results also hold for the homogeneous Kac equation [Kac59], which can be thought of a one-dimensional version of the Boltzmann equation. It has its origin in the Kac model, a probabilistic sketch of true physical collisions, and will be discussed in more detail in Part II.

The next section introduces the Boltzmann collision operator as derived by Boltzmann [Bol72], together with the main assumptions on the interactions between particles undergoing collisions. Most of the material is taken from [Vil02], with some details on the collision kernel from [Cer88].

1.1 The Boltzmann collision operator

The collision operator Q takes into account the change in the particles' velocities due to scattering. In a dilute gas, the scattering process can be assumed to involve only two particles at a time (*binary* collisions). That is, the density of the gas is low enough so that interactions of more than two particles at once are negligible.

It is further assumed that the collisions between the identical particles are *elastic*, i.e., momentum and kinetic energy are conserved in the collision process. If v' and v'_* denote the velocities of the two particles before a collision happens, and v, v_* the velocities after the collision, then

$$v' + v'_* = v + v_* \qquad \text{(momentum conservation)}$$
$$|v'|^2 + |v'_*|^2 = |v|^2 + |v_*|^2. \qquad \text{(energy conservation)}$$

A convenient way of parametrising the solutions to the above conservation laws is given by the σ-representation (centre-of-mass coordinates), which parametrises the pre-collisional velocities in terms of the post-collisional velocities by a direction $\sigma \in \mathbb{S}^{d-1}$,

$$v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2}\sigma, \quad v'_* = \frac{v + v_*}{2} - \frac{|v - v_*|}{2}\sigma,$$

see Figure 1.1.

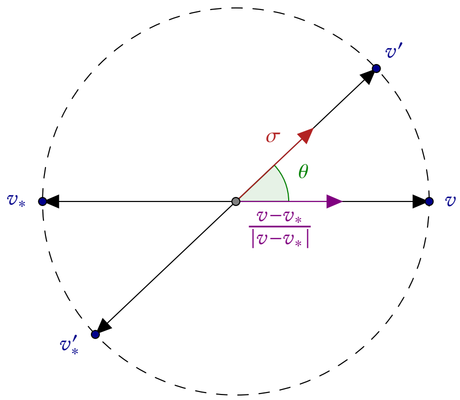

Figure 1.1: Geometry of the collision process in the σ-representation.

The collisions are further assumed to be *localised* in space and time, in the sense that they occur on a much smaller time and space scale than the typical scales in the model, and are assumed to be *micro-reversible*. Micro-reversibility means that the microscopic dynamics are time-reversible, or, put differently (interpreted probabilistically), that the probability of a pair of velocities (v', v'_*) being changed to (v, v_*) in a collision is the same as the probability of (v, v_*) being changed to (v', v'_*).

Under the *molecular chaos assumption* (Boltzmann's *Stoßzahlansatz*)[1], which states that prior to a collision the velocities of the two colliding particles are uncorrelated,

[1] we will not discuss the far-reaching consequences of this assumption here, but refer to the review article [Vil02] and references therein.

the Boltzmann collision operator Q takes the form

$$Q(f,f) = \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} B\left(|v-v_*|, \cos\theta\right)\left(f(v_*')f(v') - f(v_*)f(v)\right)\,\mathrm{d}\sigma\mathrm{d}v_*, \tag{1.3}$$

where $\cos\theta = \frac{v-v_*}{|v-v_*|}\cdot\sigma$. Notice that the quadratic nature of Q results from the fact that collisions are binary, while the chaos assumption is reflected by the dependence of Q on the tensor product $f\otimes f$. The non-negative function B, whose precise form depends on the potential by which the particles interact, is called the *collision kernel*.

This describes the general form of the Boltzmann collision operator in $d \geq 2$ dimensions. Our results also apply to the Kac equation in $d = 1$ dimension, where

$$Q(f,f) = K(f,f) = \int_{\mathbb{R}}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} b_1(\theta)\left(f(w_*')f(w') - f(w_*)f(w)\right)\,\mathrm{d}\theta\mathrm{d}w_*, \tag{1.4}$$

with collision kernel $b_1 \geq 0$. The pre- and post-collisional velocities are related by

$$\begin{pmatrix} w' \\ w_*' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} w \\ w_* \end{pmatrix}, \qquad \text{for } \theta \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}],$$

that is, energy is conserved. Notice however, that momentum is no longer conserved in this model (otherwise the scattering would be trivial...).

The Kac model and its relation to the Kac-Boltzmann equation

$$\partial_t f = K(f,f)$$

will be described in more detail in Part II.

A few details on the collision kernel

In this section we briefly review the classical scattering of two particles (in a non-rigorous manner), which yields the precise dependence of the collision kernel B on the relative velocity $|v-v_*|$ and the deviation angle $\cos\theta = \frac{v-v_*}{|v-v_*|}\cdot\sigma$. For more details we refer to Cercignani [Cer88] for a Boltzmann equation oriented approach, and the classical books by Landau-Lifshitz [LL97] and Newton[2] [New82] for a more general exposition.

Consider the classical scattering between two identical particles of mass $m = 1$, interacting by a repulsive radial interaction potential $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$. By a change of coordinates (centre-of-mass and relative coordinates), the two-body problem is equivalent to the scattering of one particle coming from infinity (the *bullet* particle), with reduced mass $\mu = \frac{m}{2}$, in the field Φ generated by the second particle (the *target* particle) sitting at the centre-of-mass location. Let q and $p = \mu v_{\mathrm{rel}}$ be position and momentum of the bullet particle.

Since the potential is radial, the angular momentum (2-form) $L = q\wedge p$ is conserved and the motion therefore takes place in a plane. It is convenient to introduce polar

[2] *not* Sir Isaac Newton, but the theoretical physicist Roger G. Newton.

coordinates r and φ in the plane of motion of the bullet particle. Then conservation of angular momentum and energy can be expressed as

$$\begin{aligned} L &= \mu r^2 \dot{\varphi} = \text{const.} \\ E &= \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) + \Phi(r) = \frac{\mu \dot{r}^2}{2} + \frac{L^2}{2\mu r^2} + \Phi(r) = \text{const.} \end{aligned} \tag{1.5}$$

An immediate consequence of these conservation laws is that $\dot{\varphi}$ has a fixed sign, so φ is monotone in time. Further,

$$\dot{r}^2(t) = \frac{2}{\mu}\left[E - \Phi(r(t))\right] - \frac{L^2}{\mu^2 r^2(t)},$$

which means that the trajectory is restricted by

$$\frac{2}{\mu}\left[E - \Phi(r(t))\right] - \frac{L^2}{\mu^2 r^2(t)} \geq 0. \tag{1.6}$$

Equality in (1.6) determines the extremal points of the trajectory, as in those points we have $\dot{r} = 0$. To describe scattering states we assume that $r(t) \in [r_{\min}, \infty)$ and that $r(t) \to \infty$ for $t \to \pm\infty$. The minimal distance $r_{\min}$ (in the point of closest approach) to the bullet particle at the centre-of-mass is then the (positive) root of the equation

$$\Phi(r) + \frac{L^2}{2\mu r^2} = E.$$

Let φ_0 be the polar angle corresponding to the point of closest approach. Integration of the conservation laws (1.5) then yields the orbit equation for r as a function of φ in an implicit form,

$$\varphi - \varphi_0 = \pm \int_{r_{\min}}^{r} \frac{\frac{L}{\mu s^2}}{\sqrt{\frac{2}{\mu}[E - \Phi(s)] - \frac{L^2}{\mu^2 s^2}}}\, \mathrm{d}s.$$

Notice that $\dot{r} = \pm\sqrt{\frac{2}{\mu}[E - \Phi(s)] - \frac{L^2}{\mu^2 s^2}}$ changes its sign at $r_{\min}$, which implies that the trajectory is symmetric around $r_{\min}$. Indeed, any two points on the orbit equidistant from the origin only differ by the sign of $\varphi - \varphi_0$. Therefore, the deviation angle θ of the bullet particle on its way past the target particle is given by

$$\theta = \pi - 2\varphi_0 = \pi - 2\int_{r_{\min}}^{\infty} \frac{\frac{L}{\mu s^2}}{\sqrt{\frac{2}{\mu}[E - \Phi(s)] - \frac{L^2}{\mu^2 s^2}}}\, \mathrm{d}s,$$

see also Figure 1.2.

For our purposes it is convenient to introduce v_∞ as the asymptotic velocity of the particle at infinity and the *impact parameter* ρ, which is the distance at which the

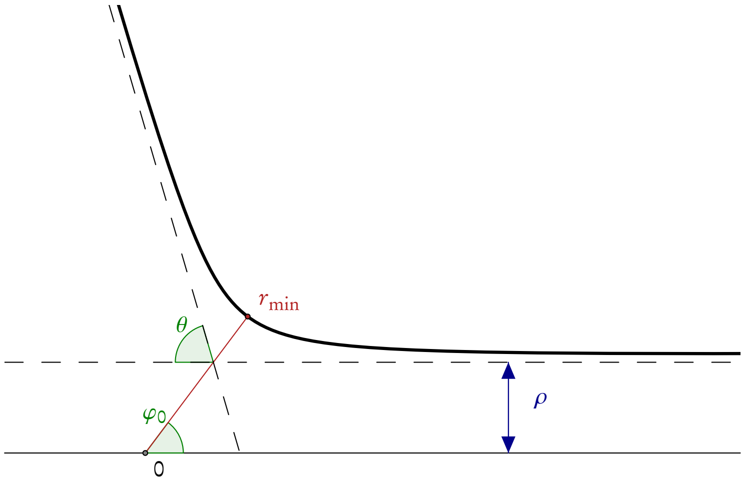

Figure 1.2: Scattering of two particles

particle would fly by the centre-of-mass if there were no field acting on it. Those two quantities can be used to express the conserved energy and momentum as

$$E = \frac{\mu}{2} v_\infty^2, \quad L = \mu \rho v_\infty.$$

Then the deviation angle can be computed as

$$\begin{aligned} \theta &= \pi - 2\rho \int_{r_{\min}}^{\infty} \left(1 - \frac{2\Phi(s)}{\mu v_\infty^2} - \frac{\rho^2}{s^2}\right)^{-1/2} \frac{\mathrm{d}s}{s^2} \\ &= \pi - 2 \int_0^{\frac{\rho}{r_{\min}}} \left(1 - x^2 - \frac{2}{\mu v_\infty^2} \Phi\left(\frac{\rho}{x}\right)\right)^{-1/2} \mathrm{d}x. \end{aligned} \tag{1.7}$$

We will later need the dependence of the impact parameter ρ on the deviation angle θ. Unfortunately, even for such easy interaction potentials as inverse power laws, the equation (1.7) cannot in general be inverted explicitly to yield $\rho = \rho(\theta)$. However, it is possible to study the asymptotics of $\rho(\theta)$ for angles close to 0 and $\frac{\pi}{2}$, see [Cer88, p.71] and [MUXY09, Appendix].

Let $x_0 = \frac{\rho}{r_{\min}}$ be the positive root of

$$1 - x_0^2 - \frac{2}{\mu v_\infty^2} \Phi\left(\frac{\rho}{x_0}\right) = 0, \tag{1.8}$$

and write with $\frac{\pi}{2} = \int_0^1 (1+y^2)^{-1/2}\,\mathrm{d}y$,

$$\begin{aligned}\frac{\theta}{2} &= \frac{\pi}{2} - \int_0^{x_0} \left(1 - x^2 - \frac{2}{\mu v_\infty^2}\Phi\left(\frac{\rho}{x}\right)\right)^{-1/2} \mathrm{d}x \\ &= \int_0^1 (1+y^2)^{-1/2}\,\mathrm{d}y - \int_0^1 \left(1 - y^2 + x_0^{-2}\left[1 - x_0^2 - \frac{2}{\mu v_\infty^2}\Phi\left(\frac{\rho}{x_0 y}\right)\right]\right)^{-1/2} \mathrm{d}y.\end{aligned}$$

By the definition (1.8) of x_0, we obtain

$$\begin{aligned}&\int_0^1 \left(1 - y^2 + x_0^{-2}\left[1 - x_0^2 - \frac{2}{\mu v_\infty^2}\Phi\left(\frac{\rho}{x_0 y}\right)\right]\right)^{-1/2} \mathrm{d}y \\ &= \int_0^1 \left(1 - y^2 + \frac{2}{\mu v_\infty^2 x_0^2}\left[\Phi\left(\frac{\rho}{x_0}\right) - \Phi\left(\frac{\rho}{x_0 y}\right)\right]\right)^{-1/2} \mathrm{d}y,\end{aligned}$$

and hence

$$\frac{\theta}{2} = \int_0^1 \frac{1}{\sqrt{1-y^2}}\left[1 - \left(1 + \frac{2}{\mu v_\infty^2 x_0^2}\frac{\Phi\left(\frac{\rho}{x_0}\right) - \Phi\left(\frac{\rho}{x_0 y}\right)}{1-y^2}\right)^{-1/2}\right] \mathrm{d}y.$$

In the limit $\theta \longrightarrow 0$, that is, $\rho \longrightarrow \infty$, we can approximate $x_0^2 = 1 + \mathcal{O}(\rho^{-1})$ and

$$\left(1 + \frac{2}{\mu v_\infty^2 x_0^2}\frac{\Phi\left(\frac{\rho}{x_0}\right) - \Phi\left(\frac{\rho}{x_0 y}\right)}{1-y^2}\right)^{-1/2} = 1 - \frac{1}{\mu v_\infty^2}\frac{\Phi(\rho) - \Phi\left(\frac{\rho}{y}\right)}{1-y^2} + \mathcal{O}(\rho^{-1}),$$

so,

$$\frac{\theta}{2} = \frac{1}{\mu v_\infty^2}\int_0^1 (1-y^2)^{-3/2}\left[\Phi(\rho) - \Phi\left(\frac{\rho}{y}\right)\right] \mathrm{d}y + \mathcal{O}(\rho^{-1}) \tag{1.9}$$

for small θ.

Example (Inverse power-law potentials, [Cer88]). For repulsive inverse power-law interaction potentials $\Phi(r) = \kappa r^{1-n}$, $n > 2$, $\kappa > 0$, we obtain the equation

$$\theta = \pi - 2\int_0^{x_0} \left(1 - x^2 - \left(\frac{x}{\beta}\right)^{n-1}\right)^{-1/2} \mathrm{d}x, \tag{1.10}$$

with $\beta = \rho v_\infty^{\frac{2}{n-1}} \left(\frac{\mu}{2\kappa}\right)^{\frac{1}{n-1}}$, and x_0 being the positive root of

$$1 - x_0^2 - \left(\frac{x_0}{\beta}\right)^{n-1} = 0 \tag{1.11}$$

for $x_0 = \frac{\rho}{r_{\min}}$. Equation (1.10) implicitly defines the dependence of β (which is directly related to the impact parameter ρ) on θ, so that we may write

$$\rho(\theta) = \left(\frac{2\kappa}{\mu}\right)^{\frac{1}{n-1}} v_\infty^{-\frac{2}{n-1}} \beta(\theta). \tag{1.12}$$

Except for the case $n = 3$ equation (1.10) cannot be solved explicitly for β. However, the function β is locally smooth and for small deviation angles $\theta \to 0$ we obtain with (1.9) the asymptotics

$$\frac{\theta}{2} = \frac{\kappa}{\mu v_\infty^2} c_n \rho^{1-n} + \mathcal{O}(\rho^{-1})$$

where $c_n = \int_0^1 (1-y^2)^{-3/2}(1-y^{n-1})\,\mathrm{d}y = \pi^{1/2}\Gamma(n/2)/\Gamma(\frac{n-1}{2})$. It follows that

$$\rho = \mathcal{O}\left(\theta^{-\frac{1}{n-1}}\right)$$

and

$$\frac{\mathrm{d}\rho^{d-1}}{\mathrm{d}\theta} = \mathcal{O}\left(\theta^{-1-\frac{d-1}{n-1}}\right)$$

as $\theta \to 0$.

In the limit $\theta \to \pi$ we must have $x_0 \to 0$ by (1.10) and thus $\beta \simeq x_0 \to 0$ by equation (1.11). In particular,

$$\theta = \pi - 2\int_0^{x_0} \left(1 - x^2 - \left(\frac{x}{\beta}\right)^{n-1}\right)^{-1/2} \mathrm{d}x \simeq \pi - 2\beta \int_0^1 \left(1 - y^{n-1}\right)^{-1/2} \mathrm{d}y,$$

so that $\beta = \mathcal{O}(\pi - \theta)$ and

$$\frac{\mathrm{d}\rho^{d-1}}{\mathrm{d}\theta} = \mathcal{O}\left((\pi - \theta)^{d-2}\right).$$

♦

Example (Debye-Yukawa type potentials, [MUXY09]). As a second example we consider the family of screened Coulomb potentials

$$\Phi(r) = \frac{\mathrm{e}^{-r^s}}{r}, \quad 0 < s < 2.$$

The case $s = 1$ is the classical *Debye-Yukawa* potential. As in the power-law case, we have

$$\theta = \pi - \int_0^{x_0} \left(1 - x^2 - \frac{2}{\mu v_\infty^2}\frac{x}{\rho}\mathrm{e}^{-\left(\frac{\rho}{x}\right)^s}\right)^{-1/2} \mathrm{d}x$$

by equation (1.7), with x_0 being the positive root of (1.8). This time, however, the dependence of ρ on θ does not factorise into a function of v_∞ and θ as before.

We therefore contend ourselves with the study of the asymptotics for *grazing collisions*, that is, for small deviation angles $\theta \to 0$. By equation (1.9), we find

$$\frac{\theta}{2} \approx \frac{1}{\mu v_\infty^2} \frac{\mathrm{e}^{-\rho^s}}{\rho} \int_0^1 (1-y^2)^{-3/2} \left(1 - y\mathrm{e}^{-\rho^s(y^{-s}-1)}\right) \mathrm{d}y.$$

Let $f(\rho) := \int_0^1 (1-y^2)^{-3/2} \left(1 - y\mathrm{e}^{-\rho^s(y^{-s}-1)}\right) \mathrm{d}y$. Then

$$f(0) = \int_0^1 (1-y^2)^{-3/2}(1-y)\,\mathrm{d}y = 1$$

and

$$f'(\rho) = s\rho^{s-1} \int_0^1 (1-y^2)^{-3/2} y^{1-s}(1-y^s)\,\mathrm{e}^{-\rho^s(y^{-s}-1)}\,\mathrm{d}y \geq 0.$$

Further, by estimating the exponential in the integrand by one, we obtain the bounds

$$0 \leq f'(\rho) \leq c_s s \rho^{s-1},$$

with $c_s = \int_0^1 (1-y^2)^{-3/2} y^{1-s}(1-y^s)\,\mathrm{d}y < \infty$ for $0 < s < 2$. Integrating the inequality yields

$$1 \leq f(\rho) \leq c_s \rho^s + 1,$$

so for small θ, that is large ρ, we can approximate $\log\theta \approx -K\rho^s$, and hence

$$\rho(\theta) \approx \kappa \left(\log\theta^{-1}\right)^{\frac{1}{s}} \tag{1.13}$$

as $\theta \to 0$. ♦

An important quantity in scattering theory is the differential scattering cross-section, as we usually do not have any information on the impact parameter in the scattering process. It describes the flux of particles scattered in a particular direction. Assume that a beam of bullet particles with energy E enters into the scattering uniformly distributed over the azimuthal direction $\omega \in \mathbb{S}^{d-2}$ and impact parameters ρ. Then all the bullet particles with impact parameter in the infinitesimal area element $\rho^{d-2}\,\mathrm{d}\rho\,\mathrm{d}\omega$ will be scattered into an infinitesimal cone of solid angle $\mathrm{d}\Omega = \sin^{d-2}\theta\,\mathrm{d}\theta\,\mathrm{d}\omega$, where the dependence of θ and ρ is given by (1.7). Here, $\mathrm{d}\omega$ denotes the uniform measure on $\mathbb{S}^{d-2}$ and $\mathrm{d}\Omega$ the uniform measure on $\mathbb{S}^{d-1}$. The differential scattering cross-section $\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$ is defined by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \sin^{d-2}\theta\,\mathrm{d}\theta\,\mathrm{d}\omega = \rho^{d-2}\,\mathrm{d}\rho\,\mathrm{d}\omega,$$

so

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{\sin^{d-2}\theta}\rho^{d-2}\frac{\mathrm{d}\rho}{\mathrm{d}\theta} = \frac{1}{\sin^{d-2}\theta}\frac{1}{d-1}\frac{\mathrm{d}\rho^{d-1}}{\mathrm{d}\theta}. \tag{1.14}$$

The Boltzmann kernel is related to the differential scattering cross-section by

$$B(z,\cos\theta) = |z|\,\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}(z,\theta), \tag{1.15}$$

where $z = v - v_*$ is the relative velocity, in particular, $|z| = v_\infty$, see, for instance [Cer88].

Example (Inverse power-law potentials, continued)**.** In the case of inverse power-law interactions, we therefore get with (1.14) and the dependence of the impact parameter on the deflection angle from (1.12), that

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{d-1}\left(\frac{2\kappa}{\mu}\right)^{\frac{d-1}{n-1}} v_\infty^{-\frac{2(d-1)}{n-1}}\frac{1}{\sin^{d-2}\theta}\frac{\mathrm{d}\beta^{d-1}(\theta)}{\mathrm{d}\theta}.$$

In particular, the differential scattering cross-section, and therefore also the Boltzmann kernel for power-law interactions, factorises into a product of a function of the asymptotic velocity $v_\infty = |v - v_*|$ (*kinetic factor*) and a function of the deviation angle θ (*angular collision kernel*). If we write

$$b(\cos\theta) := \frac{1}{d-1}\left(\frac{2\kappa}{\mu}\right)^{\frac{d-1}{n-1}}\frac{1}{\sin^{d-2}\theta}\frac{\mathrm{d}\beta^{d-1}(\theta)}{\mathrm{d}\theta},$$

then the Boltzmann kernel takes the form

$$B(|v-v_*|,\cos\theta) = |v-v_*|^\gamma b(\cos\theta)$$

with $\gamma = \frac{n-(2d-1)}{n-1}$. The angular collision kernel b is a locally smooth, non-negative function, which is well-behaved for $\theta \to \pi$, but has a *non-integrable singularity*

$$\sin^{d-2}\theta\, b(\cos\theta) \overset{\theta\to 0}{\sim} \frac{K}{\theta^{1+2\nu}} \tag{1.16}$$

for some $K > 0$ and $\nu = \frac{d-1}{2(n-1)}$.

Notice also that

$$\int_0^\pi \sin^d\theta\, b(\cos\theta)\,\mathrm{d}\theta < +\infty \tag{1.17}$$

as long as $\nu < 1$. The quantity $\int_0^\pi \sin^d\theta\, b(\cos\theta)\,\mathrm{d}\theta$ is related to the cross-section for momentum transfer M, defined by

$$\int_{\mathbb{S}^{d-1}} (v-v')\,B(|v-v_*|,\cos\theta)\,\mathrm{d}\sigma = \frac{v-v_*}{2}M(|v-v_*|),$$

see [Vil02, p.48]. Splitting the direction σ into its components along $v - v_*$ and an orthogonal direction $\omega \in \mathbb{S}^{d-2}$, we may write

$$\sigma = \cos\theta \, \frac{v - v_*}{|v - v_*|} + \sin\theta\, \omega,$$

and obtain with $v' = \frac{v+v_*}{2} + \frac{|v-v_*|}{2}\sigma$,

$$\begin{aligned}
&\int_{\mathbb{S}^{d-1}} (v - v')\, B(|v - v_*|, \cos\theta)\, \mathrm{d}\sigma \\
&= \frac{v - v_*}{2} |v - v_*|^\gamma \int_0^\pi \int_{\mathbb{S}^{d-2}} b(\cos\theta)(1 - \cos\theta) \sin^{d-2}\theta \, \mathrm{d}\omega \, \mathrm{d}\theta \\
&= \frac{v - v_*}{2} |v - v_*|^\gamma |\mathbb{S}^{d-2}| \int_0^\pi b(\cos\theta)(1 - \cos\theta) \sin^{d-2}\theta \, \mathrm{d}\theta
\end{aligned}$$

as the integral over ω vanishes by symmetry in the second term. So finiteness of the integral (1.17) guarantees that the cross-section for momentum transfer is finite.

An important simplification occurs for $n = 2d - 1$ ($= 5$ in three dimensions), where the exponent $\gamma = 0$ in the kinetic factor, so the Boltzmann kernel is a function of the deviation angle θ only. This situation is generally referred to as the *Maxwellian molecules* case. ♦

Example (Debye-Yukawa type potentials, continued)**.** For Debye-Yukawa type potentials, we obtain from (1.13) the following asymptotics of the Boltzmann kernel for small angles $\theta \to 0$:

$$B(|v - v_*|, \cos\theta) \approx |v - v_*| b(\cos\theta),$$

with

$$\sin^{d-2}\theta\, b(\cos\theta) \approx \kappa_{d,s} \theta^{-1} \left(\log\theta^{-1}\right)^{\frac{d-1}{s}-1}.$$

Again, the singularity at $\theta = 0$ is not integrable, but we have

$$\int_0^\pi \sin^d\theta\, b(\cos\theta)\, \mathrm{d}\theta < \infty$$

as long as $s < d - 1$. ♦

Assumptions on the scattering kernel

After this short excursion into classical scattering theory, we shall formulate the basic assumptions on the collision kernels we want to investigate in the following chapters. These are modelled after the two examples of inverse power-law interactions and Debye-Yukawa type interactions.

We start with the assumptions on B for $d \geq 2$:

(B1) We only consider the case of **Maxwellian molecules**, that is,

$$B(|v - v_*|, \cos\theta) = b(\cos\theta).$$

By symmetry properties of the Boltzmann collision operator $Q(f, f)$, the function b can be assumed to be supported on angles $\theta \in [0, \frac{\pi}{2}]$; for otherwise (see [Vil02]) it can be replaced by

$$\widetilde{b}(\cos\theta) = \big(b(\cos\theta) + b(\cos(\pi - \theta)\big)\, \mathbb{1}_{\{0 \le \theta \le \frac{\pi}{2}\}}.$$

(B2) The angular collision kernel b has a **non-integrable singularity** for grazing collisions $\theta \longrightarrow 0$ of the form

(a) *Inverse power-law type*

$$\sin^{d-2}\theta\, b(\cos\theta) \sim \frac{\kappa}{\theta^{1+2\nu}}, \quad \text{as } \theta \longrightarrow 0^+ \tag{1.18}$$

for some $\kappa > 0$ and $0 < \nu < 1$,

(b) *Debye-Yukawa type*

$$\sin^{d-2}\theta\, b(\cos\theta) \sim \kappa\theta^{-1}\left(\log\theta^{-1}\right)^{\mu} \tag{1.19}$$

for some $\kappa, \mu > 0$.

(B3) The cross-section for momentum transfer is finite,

$$\int_0^{\pi/2} \sin^d\theta\, b(\cos\theta)\, \mathrm{d}\theta < \infty. \tag{1.20}$$

In the one-dimensional Boltzmann-Kac case, b_1 was originally chosen to be constant. We will assume, as in [Des03], that

(K1) b_1 is an even function supported on angles $\theta \in [-\frac{\pi}{4}, \frac{\pi}{4}]$. The restriction on the smaller set of deviation angles θ is again possible due to symmetry properties of $K(f, f)$, for otherwise it can be replaced by its symmetrised version

$$\widetilde{b}_1(\theta) = \big(b_1(\theta) + b_1(\tfrac{\pi}{2} - \theta)\big)\, \mathbb{1}_{\{0 \le \theta \le \frac{\pi}{4}\}} + \big(b_1(\theta) + b_1(-\tfrac{\pi}{2} - \theta)\big)\, \mathbb{1}_{\{-\frac{\pi}{4} \le \theta \le 0\}}.$$

(K2) b_1 has the non-integrable singularity

$$b_1(\theta) \sim \frac{\kappa}{|\theta|^{1+2\nu}}, \quad \text{for } \theta \longrightarrow 0, \tag{1.21}$$

with $0 < \nu < 1$ and some $\kappa > 0$.

(K3) b_1 satisfies

$$\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} b_1(\theta) \sin^2\theta\, \mathrm{d}\theta < \infty. \tag{1.22}$$

1.2 Weak solutions: definition, existence, and uniqueness

The non-integrability of the angular collision kernel produces some difficulties. In particular, it is a priori not clear if the expressions (1.3) and (1.4) are even well-defined.

In this section we discuss the above issue and define the class of weak solutions for which our regularity results hold. We will mainly work with the weighted L^p spaces, defined as

$$L^p_\alpha(\mathbb{R}^d) := \left\{ f \in L^p(\mathbb{R}^d) : \langle \cdot \rangle^\alpha f \in L^p(\mathbb{R}^d) \right\}, \quad p \geq 1, \alpha \in \mathbb{R},$$

with norm

$$\|f\|_{L^p_\alpha(\mathbb{R}^d)} = \left(\int_{\mathbb{R}^d} |f(v)|^p \langle v \rangle^{\alpha p} \, \mathrm{d}v \right)^{1/p}, \quad \langle v \rangle := (1 + |v|^2)^{1/2}.$$

We will also use the weighted (L^2 based) Sobolev spaces

$$H^k_\ell(\mathbb{R}^d) = \left\{ f \in \mathcal{S}'(\mathbb{R}^d) : \langle \cdot \rangle^\ell f \in H^k(\mathbb{R}^d) \right\}, \quad k, \ell \in \mathbb{R},$$

where $H^k(\mathbb{R}^d)$ are the usual Sobolev spaces given by

$$H^k(\mathbb{R}^d) = \left\{ f \in \mathcal{S}'(\mathbb{R}^d) : \langle \cdot \rangle^k \hat{f} \in L^2(\mathbb{R}^d) \right\}, \quad k \in \mathbb{R}.$$

The inner product on $L^2(\mathbb{R}^d)$ is given by $\langle f, g \rangle = \int_{\mathbb{R}^d} \overline{f(v)} g(v) \, \mathrm{d}v$.

It will be assumed that the initial datum $f_0 \not\equiv 0$ is a non-negative density with finite mass, energy and entropy, which is equivalent to

$$f_0 \geq 0, \quad f_0 \in L^1_2(\mathbb{R}^d) \cap L \log L(\mathbb{R}^d), \tag{1.23}$$

where

$$L \log L(\mathbb{R}^d) = \left\{ f : \mathbb{R}^d \to \mathbb{R} \text{ measurable} : \|f\|_{L \log L} < \infty \right\},$$

and

$$\|f\|_{L \log L} = \int_{\mathbb{R}^d} |f(v)| \log \left(1 + |f(v)|\right) \, \mathrm{d}v,$$

and the negative of the entropy is given by $H(f) := \int_{\mathbb{R}^d} f \log f \, \mathrm{d}v$.

The space $L^1_2(\mathbb{R}^d) \cap L \log L(\mathbb{R}^d)$ is very natural, since we have

Lemma 1.1. *Let $f \geq 0$. Then*

$$f \in L^1_2(\mathbb{R}^d) \cap L \log L(\mathbb{R}^d) \quad \Leftrightarrow \quad f \in L^1_2(\mathbb{R}^d) \text{ and } H(f) \text{ is finite.}$$

Proof. Let $f \in L^1_2(\mathbb{R}^d) \cap L \log L(\mathbb{R}^d)$. Then

$$|H(f)| = \int_{\mathbb{R}^d} f \log_+ f \,\mathrm{d}v + \int_{\mathbb{R}^d} f \log_- f \,\mathrm{d}v.$$

The positive part is bounded by $\int f \log(1+f)\,\mathrm{d}v = \|f\|_{L \log L}$. The negative part can be controlled by

$$\begin{aligned}\int_{\mathbb{R}^d} f \log_- f \,\mathrm{d}v &= \int_{\{f\le 1\}} f \log \frac{1}{f} \,\mathrm{d}v \le C_\delta \int_{\{f \le 1\}} f^{1-\delta} \,\mathrm{d}v \\ &\le C_\delta \left(\int_{\mathbb{R}^d} (1+|v|^2)^{-\frac{1-\delta}{\delta}} \,\mathrm{d}v \right)^{\delta} \|f\|_{L^1_2}^{1-\delta},\end{aligned}$$

which is finite for $0 < \delta < \frac{2}{d+2}$, having used that for any $\delta > 0$ there exists a constant C_δ such that $\log t \le C_\delta t^\delta$ for all $t \ge 1$.

Conversely, let $f \in L^1_2(\mathbb{R}^d)$ with finite entropy $H(f)$. Then

$$\int_{\mathbb{R}^d} f \log(1+f)\,\mathrm{d}v = \int_{\{f \le 1\}} f \log(1+f)\,\mathrm{d}v + \int_{\{f>1\}} f \log(1+f)\,\mathrm{d}v$$

On the set where $f \le 1$, we replace f by 1 and where $f > 1$, we bound $1+f$ by $2f$, leading to

$$\int_{\mathbb{R}^d} f \log(1+f)\,\mathrm{d}v \le \log 2 \int_{\mathbb{R}^d} f \,\mathrm{d}v + \int_{\mathbb{R}^d} f \log f \,\mathrm{d}v + \int_{\mathbb{R}^d} f \log_- f \,\mathrm{d}v.$$

As above, we conclude

$$\int_{\mathbb{R}^d} f \log(1+f)\,\mathrm{d}v \le \log 2 \|f\|_{L^1(\mathbb{R}^d)} + H(f) + C_{\delta,d} \|f\|_{L^1_2(\mathbb{R}^d)}^{1-\delta}, \tag{1.24}$$

with a finite constant $C_{\delta,d}$ for $0 < \delta < \frac{2}{d+2}$. ■

The following is the precise definition of weak solutions which we use:

Definition 1.2 (Weak Solutions of the Cauchy Problem (1.2) [Ark81, Vil98, Des95]). Assume that the initial datum f_0 is in $L^1_2(\mathbb{R}^d) \cap L \log L(\mathbb{R}^d)$. $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ is called a weak solution to the Cauchy problem (1.2), if it satisfies the following conditions[3]:

(i) $f \ge 0, \quad f \in \mathscr{C}(\mathbb{R}_+; \mathscr{D}'(\mathbb{R}^d)) \cap L^\infty(\mathbb{R}_+; L^1_2(\mathbb{R}^d) \cap L \log L(\mathbb{R}^d))$;

(ii) $f(0,\cdot) = f_0$;

(iii) For all $t \ge 0$, mass is conserved, $\int_{\mathbb{R}^d} f(t,v)\,\mathrm{d}v = \int_{\mathbb{R}^d} f_0(v)\,\mathrm{d}v$, kinetic energy is decreasing, $\int_{\mathbb{R}^d} f(t,v)\, v^2 \,\mathrm{d}v \le \int_{\mathbb{R}^d} f_0(v)\, v^2 \,\mathrm{d}v$, and the entropy is increasing, that is, $H(f(t,\cdot)) \le H(f_0)$;

[3] Throughout the text, whenever not explicitly mentioned, we will drop the dependence on t of a function, i.e. $f(v) := f(t,v)$ etc.

(iv) For all $\varphi \in \mathscr{C}^1(\mathbb{R}_+;\mathscr{C}_0^\infty(\mathbb{R}^d))$ one has

$$\begin{aligned}\langle f(t,\cdot),\varphi(t,\cdot)\rangle - \langle f_0,\varphi(0,\cdot)\rangle - \int_0^t \langle f(\tau,\cdot),\partial_\tau\varphi(\tau,\cdot)\rangle\,\mathrm{d}\tau \\ = \int_0^t \langle Q(f,f)(\tau,\cdot),\varphi(\tau,\cdot)\rangle\,\mathrm{d}\tau, \quad \text{for all } t\geq 0,\end{aligned} \tag{1.25}$$

where the latter expression involving Q is defined by

$$\begin{aligned}\langle Q(f,f),\varphi\rangle = \frac{1}{2}\int_{\mathbb{R}^{2d}}\int_{\mathbb{S}^{d-1}} b\left(\frac{v-v_*}{|v-v_*|}\cdot\sigma\right) f(v_*)f(v) \\ \times\left(\varphi(v')+\varphi(v'_*)-\varphi(v)-\varphi(v_*)\right)\,\mathrm{d}\sigma\mathrm{d}v\mathrm{d}v_*,\end{aligned} \tag{1.26}$$

for test functions $\varphi \in W^{2,\infty}(\mathbb{R}^d)$ in dimension $d\geq 2$, and in one dimension

$$\begin{aligned}\langle Q(f,f),\varphi\rangle &= \langle K(f,f),\varphi\rangle \\ &= \int_{\mathbb{R}^2}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} b_1(\theta)\,g(w_*)g(w)\,(\varphi(w')-\varphi(w))\,\mathrm{d}\theta\mathrm{d}w\mathrm{d}w_*\end{aligned} \tag{1.27}$$

for test functions $\varphi \in W^{2,\infty}(\mathbb{R})$, making use of symmetry properties of the Boltzmann and Kac collision operators and cancellation effects.

Remark. In the study of the Boltzmann operator Q it is often convenient to use its *bilinear form*

$$Q(g,f) = \int_{\mathbb{R}^d\times\mathbb{S}^{d-1}} B(|v-v_*|,\cos\theta)\left(g(v'_*)f(v')-g(v_*)f(v)\right)\,\mathrm{d}v_*\mathrm{d}\sigma,$$

which, for $f,g\in L^1_2(\mathbb{R}^d)$ is well-defined in the weak sense

$$\begin{aligned}\langle Q(g,f),\varphi\rangle := \int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} B(|v-v_*|,\cos\theta)\,g(v_*)f(v) \\ \times\left[\varphi(v'_*)+\varphi(v')-\varphi(v_*)-\varphi(v)\right]\,\mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma\end{aligned}$$

for $\varphi\in W^{2,\infty}(\mathbb{R}^d)$. ♦

A few remarks regarding Definition 1.2 are in order:[4]

I. Maxwell's weak formulation. Formally, equations (1.26) and (1.27) can be derived from the original expressions (1.3) and (1.4) by suitable coordinate transformations. Indeed, the *pre-post-collisional change of variables*

$$(v,v_*,\sigma)\to(v',v'_*,k) \quad \text{with} \quad k=\frac{v-v_*}{|v-v_*|}$$

[4] We refer to the reviews by Villani [Vil02] and Alexandre [Ale09] for more details.

is involutive by momentum conservation, has unit Jacobian by energy conservation, and leaves B invariant by micro-reversibility. Further, the change of coordinates $(v, v_*) \to (v_*, v)$ defines an involution with unit Jacobian. It follows that for suitable test functions φ the following equalities hold, at least formally[5]:

$$\begin{aligned}
&\int_{\mathbb{R}^d} Q(f,f)(v)\varphi(v)\,\mathrm{d}v \\
&= \int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} B(|v-v_*|,\cos\theta)\left(f'f'_* - ff_*\right)\varphi\,\mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma \\
&= \int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} B(|v-v_*|,\cos\theta) ff_*\left(\varphi'-\varphi\right)\,\mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma &(1.28)\\
&= \frac{1}{2}\int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} B(|v-v_*|,\cos\theta) ff_*\left(\varphi'+\varphi'_*-\varphi-\varphi_*\right)\,\mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma &(1.29)\\
&= -\frac{1}{4}\int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} B(|v-v_*|,\cos\theta)\left(f'f'_*-ff_*\right)\left(\varphi'+\varphi'_*-\varphi-\varphi_*\right)\,\mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma \\
& &(1.30)
\end{aligned}$$

Expressions (1.28) and (1.29) are particularly important in the case of singular collision kernels B, because the differences $\varphi' - \varphi$ and $\varphi' + \varphi'_* - \varphi - \varphi_*$ vanish for small deviation angles θ (for which $v' \approx v$ and $v'_* \approx v_*$) under some smoothness assumption on φ. Indeed, if $\varphi \in W^{2,\infty}(\mathbb{R}^d)$, then we can Taylor expand it to obtain

$$\begin{aligned}
\varphi(v') - \varphi(v) =& \nabla\varphi(v)\cdot(v'-v) \\
&+ \int_0^1 D^2\varphi(v+\tau(v'-v))(1-\tau)\,\mathrm{d}\tau \;:\; (v'-v)\otimes(v'-v),
\end{aligned}$$

and similarly

$$\begin{aligned}
\varphi(v'_*) - \varphi(v_*) =& \nabla\varphi(v_*)\cdot(v'_*-v_*) \\
&+ \int_0^1 D^2\varphi(v_*+\tau(v'_*-v_*))(1-\tau)\,\mathrm{d}\tau \;:\; (v'_*-v_*)\otimes(v'_*-v_*).
\end{aligned}$$

Notice that $v'_* - v_* = -(v' - v)$ and, decomposing the direction σ into its components along $v - v_*$ and an orthogonal direction $\omega \in \mathbb{S}^{d-2}$, we may write

$$\sigma = \cos\theta\,\frac{v-v_*}{|v-v_*|} + \sin\theta\,\omega,$$

[5] we use the common notation $f_* = f(v_*)$, $f' = f(v')$ etc to make the rather lengthy expressions a bit more readable

and obtain

$$
\begin{aligned}
v' - v &= \frac{v + v_*}{2} + \frac{|v - v_*|}{2}\sigma - v = -\frac{v - v_*}{2} + \frac{|v - v_*|}{2}\sigma \\
&= -\frac{v - v_*}{2}(1 - \cos\theta) + \frac{|v - v_*|}{2}\sin\theta\,\omega \\
&= -\sin^2\frac{\theta}{2}(v - v_*) + \frac{|v - v_*|}{2}\sin\theta\,\omega.
\end{aligned}
$$

With these observations we can estimate

$$|\varphi' + \varphi'_* - \varphi - \varphi_*| \le C(\varphi)|v - v_*|^2\theta,$$

where we also used that $|\nabla\varphi(v) - \nabla\varphi(v_*)| \le \|\varphi\|_{W^{2,\infty}}|v - v_*|$ to get a second-order contribution in $|v - v_*|$.

Recall, though, that we need a factor of θ^2 in order to make the angular singularity in $b(\cos\theta)$ integrable for small θ. Notice that the "bad", i.e., linear in θ, contribution comes from the term proportional to ω, which disappears by symmetry when integrating over $\mathbb{S}^{d-2}$! In this case, one indeed gets the desired smoothing effect of the singularity

$$\left|\int_{\mathbb{S}^{d-2}}(\varphi' + \varphi'_* - \varphi - \varphi_*)\,\mathrm{d}\omega\right| \le C(\varphi)|v - v_*|^2\theta^2$$

with a constant $C(\varphi)$ depending on the $W^{2,\infty}$ norm of φ. This is the reason why the definition of $\langle Q(f,f),\varphi\rangle$ in (1.26) makes sense by duality in $(W^{2,\infty})^*$, with

$$
\begin{aligned}
&|\langle Q(f,f),\varphi\rangle| \\
&\le \int_{\mathbb{R}^{2d}}\int_0^{\frac{\pi}{2}} b(\cos\theta)\sin^{d-2}\theta\, f f_* \left|\int_{\mathbb{S}^{d-2}}(\varphi' + \varphi'_* - \varphi - \varphi_*)\,\mathrm{d}\omega\right| \mathrm{d}\theta\mathrm{d}v\mathrm{d}v_* \\
&\le C(\varphi)\int_0^{\frac{\pi}{2}} b(\cos\theta)\sin^{d-2}\theta\,\theta^2\,\mathrm{d}\theta \int_{\mathbb{R}^{2d}} f(v)f(v_*)\,|v - v_*|^2\,\mathrm{d}v\mathrm{d}v_* \\
&\le C(\varphi, b)\,\|f\|^2_{L^1_2}.
\end{aligned} \tag{1.31}
$$

As noted by Alexandre and Villani [AV02], a more careful estimate shows that one does not really need to define the Boltzmann operator in the fully symmetric weak form (1.29), but can work with (1.28). More precisely, they obtained the bound

$$
\begin{aligned}
&\langle Q(f,f),\varphi\rangle \\
&\le \frac{1}{4}\|\varphi\|_{W^{2,\infty}(\mathbb{R}^d)}\int_{\mathbb{R}^{2d}} f_* f\,|v - v_*|(1 + |v - v_*|)M(|v - v_*|)\,\mathrm{d}v\,\mathrm{d}v_*,
\end{aligned} \tag{1.32}
$$

where M is the cross-section for momentum transfer introduced in Section 1.1. However, this does not make any difference when discussing the homogeneous Maxwellian problem.

II. Formal conservation laws. Maxwell's weak formulation (1.29) also implies the formal conservation laws

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} f(t,v)\begin{pmatrix}1\\ v_j\\ \frac{|v|^2}{2}\end{pmatrix}\mathrm{d}v = 0, \quad j = 1,\dots,d,$$

of total mass, momentum, and kinetic energy. Indeed, if f solves the (homogeneous) Boltzmann equation, then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} f(t,v)\varphi(v)\,\mathrm{d}v &= \int_{\mathbb{R}^d} Q(f,f)(t,v)\varphi(v)\,\mathrm{d}v\\ &= \frac{1}{2}\int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} B(|v-v_*|,\cos\theta)ff_*\left(\varphi'+\varphi'_*-\varphi-\varphi_*\right)\,\mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma.\end{aligned}$$

So if

$$\varphi(v')+\varphi(v'_*) = \varphi(v)+\varphi(v_*) \quad \text{for a.e. } (v,v_*,\sigma)\in\mathbb{R}^{2d}\times\mathbb{S}^{d-1}, \tag{1.33}$$

then, at least formally, $\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} f(t,v)\varphi(v)\,\mathrm{d}v = 0$. Solutions to the functional equation 1.33 are called *(summational) collisional invariants* and under the very weak assumption that φ is measurable and finite almost everywhere, one can show that[6]

$$\varphi(v) = A + B\cdot v + C\frac{|v|^2}{2}, \quad A, C\in\mathbb{R},\ B\in\mathbb{R}^d,$$

see [AC90].

III. Formal increase of the entropy, Boltzmann's H theorem. Entropy plays an extremely important role in kinetic theory, and there is obviously not enough time nor space to treat it in proper detail in this thesis. We will discuss some aspects again in Part II in the simpler Kac model, and refer for a more detailed treatise in Villani's review article [Vil02, p.32ff] and references therein, in particular [Kac59].

The important thing for us is that it is physically reasonable to assume that the entropy of the initial datum is finite, and that it is increasing under the Boltzmann evolution. To see this, at least at a formal level, we take a closer look at the Boltzmann H functional

$$H(f(t,\cdot)) = \int_{\mathbb{R}^d} f(t,v)\log f(t,v)\,\mathrm{d}v,$$

[6] Smooth $\mathscr{C}^2$ solutions to the equation (1.33) had already been shown by Boltzmann to be linear combinations of 1, v, and $|v|^2$.

which is the negative of the physical (dynamical, i.e. non-equilibrium) entropy. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}H(f(t,\cdot)) = \int_{\mathbb{R}^d} Q(f,f)(t,v)\log f(t,v)\,\mathrm{d}v + \int_{\mathbb{R}^d} Q(f,f)(t,v)\,\mathrm{d}v.$$

Using the weak formulation (1.30) (Boltzmann's weak formulation), and

$$\int Q(f,f)\,\mathrm{d}v = 0,$$

we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}H(f(t,\cdot)) &= \int_{\mathbb{R}^d} Q(f,f)(t,v)\log f(t,v)\,\mathrm{d}v \\ &= -D(f(t,\cdot))\end{aligned}$$

with the *entropy dissipation functional*

$$D(f) := -\frac{1}{4}\int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} B(|v-v_*|,\cos\theta)\left(f'f'_* - ff_*\right)\log\frac{f'f'_*}{ff_*}\,\mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma.$$

Notice that the function $(x,y)\mapsto(x-y)\log\frac{x}{y}$ is non-negative, so that along solutions of the homogeneous Boltzmann equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}H(f(t,\cdot)) = -D(f(t,\cdot)) \leq 0. \tag{1.34}$$

If the Boltzmann collision kernel B is strictly positive, then equality in (1.34) occurs if and only if

$$f(v')f(v'_*) = f(v)f(v_*) \quad \text{for a.e. } (v,v_*,\sigma)\in\mathbb{R}^{2d}\times\mathbb{S}^{d-1}.$$

Taking logarithms, we see that $\varphi = \log f$ has to be a collisional invariant, so f has to be a Maxwellian distribution

$$f(v) = \left(\frac{\beta}{2\pi}\right)^{d/2} \mathrm{e}^{-\frac{\beta}{2}|v-u|^2}$$

for some inverse temperature $\beta > 0$ and momentum $u\in\mathbb{R}^d$, see also (1.33). The Maxwellian is normalised to have total mass $\int f\,\mathrm{d}v = 1$, and the parameters u and β are uniquely determined by

$$\begin{aligned}\int_{\mathbb{R}^d} vf(v)\,\mathrm{d}v &= u \\ \int_{\mathbb{R}^d}|v|^2 f(v)\,\mathrm{d}v &= |u|^2 + \frac{d}{\beta}.\end{aligned}$$

Collecting results from the literature, the following is known regarding the existence, uniqueness and further properties of weak solutions:

$d \geq 2$: The existence of weak solutions of the Cauchy problem (1.2) with initial conditions satisfying (1.23) for the homogeneous Boltzmann equation was first proved by Arkeryd [Ark72, Ark81] (see also the articles by Goudon [Gou97], Villani [Vil98], and Desvillettes [Des01, Des03]). Uniqueness in this case was shown by Toscani and Villani [TV99], see also the review articles by Mischler and Wennberg [MW99] (for the cut-off case) and Desvillettes [Des01].

$d = 1$: For the homogeneous non-cutoff Kac equation for Maxwellian molecules existence of weak solutions was established by Desvillettes [Des95].

For the sake of completeness, we present the relevant existence theorems for our smoothing result in the following two theorems.

Theorem 1.3 (Existence of weak solutions for the Boltzmann equation, Arkeryd [Ark72], Villani [Vil98]). *Let $f_0 \geq 0$, $f_0 \in L^1_2 \cap L \log L(\mathbb{R}^d)$, and assume that the collision kernel satisfies Assumptions* **(B1)–(B3)**. *There exists a weak solution f of the Cauchy problem (1.2) for the Boltzmann equation with initial datum f_0 in the sense of Definition 1.2, which in addition conserves momentum and energy.*

Moreover, for all $\varphi \in W^{2,\infty}(\mathbb{R}^d)$, the map

$$(0,\infty) \ni t \mapsto \int_{\mathbb{R}^d} f(t)\varphi \, \mathrm{d}v$$

is Lipschitz continuous.

We give a short proof of the existence result for Maxwellian molecules, which is contained in the more general results due to Arkeryd [Ark72], with extensions due to Villani [Vil98].

Proof. Fix some arbitrary $T > 0$ and let $f_0 \geq 0$, $f_0 \in L^1_2 \cap L \log L(\mathbb{R}^d)$. Assume that the collision kernel satisfies (1.18) and (1.20).

The first step is to construct a sequence

$$b_n(\cos\theta) := b(\cos\theta) \wedge n$$

of cut-off collision kernels, with corresponding Boltzmann operators Q_n satisfying the Grad cut-off assumption.

By Arkeryd's existence theorem in the cut-off case [Ark72], for each $n \in \mathbb{N}$ there exists a classical solution $f_n \geq 0$ of

$$\partial_t f_n = Q_n(f_n, f_n), \quad t \in [0,T], \tag{1.35}$$

with $f_n(0) = f_0$ for all $n \in \mathbb{N}$, such that mass, momentum and kinetic energy are preserved, and $H(f_n)$ is decreasing. That is, for all $t \in [0,T]$, we have the a priori bounds

$$\begin{gathered} \int_{\mathbb{R}^d} f_n(t,v)\left(1+|v|^2\right) \mathrm{d}v = \|f_0\|_{L^1_2}, \\ H(f_n(t)) \leq H(f_0), \quad n \in \mathbb{N}. \end{gathered} \tag{1.36}$$

So the sequence $\{f_n(t)\}_{n\in\mathbb{N}}$ is bounded in $L^1_2 \cap L\log L$, see Lemma 1.1. In particular, by De la Vallée Poussin's criterion, $\{f_n(t)\}$ is equi-integrable. Let $\phi : \mathbb{R}^d \to [0,\infty)$ be a function such that $\langle\cdot\rangle^{-\kappa}\phi \in L^\infty(\mathbb{R}^d)$ for some $0 \le \kappa < 2$. Then

$$\begin{aligned}&\sup_{n\in\mathbb{N}} \int_{|v|\ge R} f_n(t,v)\phi(v)\,\mathrm{d}v \le \|\langle\cdot\rangle^{-\kappa}\phi\|_{L^\infty} \sup_{n\in\mathbb{N}} \int_{|v|\ge R} f_n(t,v)\langle v\rangle^{\kappa}\,\mathrm{d}v \\ &\le \frac{\|\langle\cdot\rangle^{-\kappa}\phi\|_{L^\infty}}{(1+R^2)^{2-\kappa}} \sup_{n\in\mathbb{N}} \int_{|v|\ge R} f_n(t,v)\langle v\rangle^2\,\mathrm{d}v \le \frac{\|\langle\cdot\rangle^{-\kappa}\phi\|_{L^\infty}}{(1+R^2)^{2-\kappa}}\|f_0\|_{L^1_2},\end{aligned}$$

hence

$$\lim_{R\to\infty} \sup_{n\in\mathbb{N}} \int_{|v|\ge R} f_n(t,v)\phi(v)\,\mathrm{d}v = 0. \tag{1.37}$$

It follows that the sequence $\{f_n(t)\}$ is weakly compact in $L^1(\mathbb{R}^d)$ by the Dunford-Pettis Theorem[7]. By a diagonal-sequence argument, we can find a non-negative function $f \in L^\infty\left([0,T]; L^1_2 \cap L\log L(\mathbb{R}^d)\right)$ such that, along a subsequence, $f_n(t) \rightharpoonup f(t)$ weakly in L^1 for all $t \in [0,T] \cap \mathbb{Q}$.

Note that for any test function $\psi \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ and $0 \le s < t \le T$ we have

$$\begin{aligned}&\left|\int_{\mathbb{R}^d} f_n(t,v)\psi(v)\,\mathrm{d}v - \int_{\mathbb{R}^d} f_n(s,v)\psi(v)\,\mathrm{d}v\right| \\ &\le \int_s^t \left|\int_{\mathbb{R}^d} Q_n(f_n,f_n)(\tau,v)\psi(v)\,\mathrm{d}v\right|\mathrm{d}\tau \\ &\le C(\psi,b)\int_s^t \|f_n(\tau,\cdot)\|^2_{L^1_2(\mathbb{R}^d)}\,\mathrm{d}\tau \\ &\le C(\psi,b)|t-s|\|f_0\|^2_{L^1_2(\mathbb{R}^d)},\end{aligned} \tag{1.38}$$

see also inequality (1.31). The constant $C(\psi,b)$ is independent of n, since

$$\int_0^{\frac{\pi}{2}} b_n(\cos\theta)\,\sin^{d-2}\theta\,\theta^2\,\mathrm{d}\theta \le \int_0^{\frac{\pi}{2}} b(\cos\theta)\,\sin^{d-2}\theta\,\theta^2\,\mathrm{d}\theta,$$

for $n \in \mathbb{N}$, and depends only on the test function ψ (through its $W^{2,\infty}$ norm) and on $\int_0^{\frac{\pi}{2}} b(\cos\theta)\,\sin^{d-2}\theta\,\theta^2\,\mathrm{d}\theta < \infty$. We also used that the functions f_n conserve mass and kinetic energy.

[7] **Theorem** (Weak compactness in L^1, Dunford-Pettis) Let $(\Omega,\mathscr{A},\mu)$ be a measure space. $\mathscr{F} \subset L^1(\Omega)$ is weakly sequentially precompact if and only if

(i) $\mathscr{F}$ is bounded in $L^1(\Omega)$,

(ii) $\mathscr{F}$ is equi-integrable and for every $\epsilon > 0$ there exists $A \in \mathscr{A}$, $\mu(A) < \infty$, such that

$$\sup_{f\in\mathscr{F}} \int_{A^c} |f|\,\mathrm{d}\mu \le \epsilon.$$

We refer to, for instance, Theorem 2.54 in [FL07] for a proof.

The sequence $\{\int f_n(t)\psi \,\mathrm{d}v\}_{n\in\mathbb{N}}$ is therefore uniformly equicontinuous in t, and converges pointwise to $\int_{\mathbb{R}^d} f(t)\psi \,\mathrm{d}v$ for all $t \in [0,T]\cap\mathbb{Q}$. Hence, for each test function $\psi \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ the map $t \mapsto \int_{\mathbb{R}^d} f(t)\psi \,\mathrm{d}v$ is continuous (even Lipschitz continuous by (1.38)) on $[0,T]$.

Further, by (1.37),

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f_n(t,v)\phi(v)\,\mathrm{d}v = \int_{\mathbb{R}^d} f(t,v)\phi(v)\,\mathrm{d}v$$

for all $t \in [0,T]$ along a subsequence, if $\langle\cdot\rangle^{-\kappa}\phi \in L^\infty(\mathbb{R}^d)$ for some $\kappa \in [0,2)$. In particular, mass and momentum are conserved.

Now if $\varphi \in \mathscr{C}^1(\mathbb{R}_+;\mathscr{C}_0^\infty(\mathbb{R}^d))$, and $t \in [0,T]$, then for each $n \in \mathbb{N}$

$$\begin{aligned}\int_0^t \langle Q_n(f_n,f_n)(\tau),\varphi(\tau)\rangle\,\mathrm{d}\tau &= \int_0^t \langle \partial_\tau f_n(\tau),\varphi(\tau)\rangle \\ = \langle f_n(t),\varphi(t)\rangle - \langle f_0,\varphi(0)\rangle &- \int_0^t \langle f_n(\tau),\partial_\tau\varphi(\tau)\rangle\,\mathrm{d}\tau,\end{aligned}$$

because f_n is a classical solution of (1.35) with initial datum $f_n(0) = f_0$. Thus, taking the limit $n \to \infty$, the above continuity properties guarantee that f satisfies the weak formulation (1.26) of the Boltzmann equation.

The increase of entropy (i.e. decrease of H) follows by convexity of the functional H, together with the weak-L^1 convergence of $f_n(t) \rightharpoonup f(t)$ along a subsequence, and the corresponding property of the approximation f_n,

$$H(f_n(t)) = H(\text{w-lim}_{n\to\infty} f_n(t)) \le \liminf_{n\to\infty} H(f_n(t)) \le H(f_0).$$

To check energy conservation we use that the weak formulation implies for any $\psi \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ (time independent), and any $0 \le s < t \le T$,

$$\begin{aligned}&\int_{\mathbb{R}^d} f(t,v)\psi(v)\,\mathrm{d}v - \int_{\mathbb{R}^d} f(s,v)\psi(v)\,\mathrm{d}v = \int_s^t \langle Q(f,f)(\tau),\psi\rangle\,\mathrm{d}\tau \\ &= \int_s^t \int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} b(\cos\theta) f(\tau,v)f(\tau,v_*)\left[\psi(v') + \psi(v_*') - \psi(v) - \psi(v_*)\right]\mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma\mathrm{d}\tau.\end{aligned} \tag{1.39}$$

Let $\psi(v) = |v|^2\chi(\epsilon v)$ with χ a smooth cut-off function, say equal to 1 on $\{|v| \le 1\}$ and zero for $\{|v| \ge 2\}$. Then by our a priori bounds on the energy (1.36) (which carry over to the weak limit f) and Lebesgue's dominated convergence theorem, the expression in brackets on the right hand side of (1.39) tends to

$$|v'|^2 + |v_*'|^2 - |v|^2 - |v_*|^2 = 0$$

as $\epsilon \to 0$. ∎

In an analogous fashion, one can prove

Theorem 1.4 (Existence of weak solutions for the Kac equation, Desvillettes [Des95]). *Let $f_0 \geq 0$, $f_0 \in L^1_2 \cap L \log L(\mathbb{R})$, and assume that the collision kernel satisfies Assumptions* **(K1)–(K3)**. *Then there exists a weak solution of the Boltzmann-Kac equation*

$$\partial_t f = K(f, f) = Q(f, f)$$

with initial datum f_0 in the sense of Definition 1.2.

Moreover, if $f_0 \in L^1_{2p}(\mathbb{R})$ for some $p \geq 2$, then $f \in L^\infty(\mathbb{R}_+; L^1_{2p}(\mathbb{R}))$ and the conservation of energy

$$\int_{\mathbb{R}} f(t, v)|v|^2 \, \mathrm{d}v = \int_{\mathbb{R}} f_0(v)|v|^2 \, \mathrm{d}v$$

holds for all $t \geq 0$.

For more details on the proof we refer to [Des97].

1.3 Bobylev identity

We close this chapter with a calculation of the Fourier transform of the Boltzmann operator in the Maxwellian case, which goes back to Bobylev [Bob84], see also the appendix of [ADVW00].

The following convention regarding the Fourier transform of a Schwartz function $f \in \mathcal{S}(\mathbb{R}^d)$ will be used throughout Part I,

$$(\mathcal{F} f)(\eta) = \hat{f}(\eta) = \int_{\mathbb{R}^d} f(v) \, \mathrm{e}^{-2\pi \mathrm{i} v \cdot \eta} \, \mathrm{d}v,$$

and extended by duality to the space of tempered distributions $\mathcal{S}'$.

Theorem 1.5 (Bobylev Identity for Maxwellian molecules). *The (distributional) Fourier transform of the Boltzmann collision operator for Maxwellian molecules is given by*

$$\begin{aligned}&\widehat{Q(g, f)}(\xi) \\ &= \frac{1}{2} \int_{\mathbb{S}^{d-1}} b\left(\frac{\xi}{|\xi|} \cdot \sigma\right) \left[\hat{g}(\xi^-)\hat{f}(\xi^+) + \hat{g}(\xi^+)\hat{f}(\xi^-) - \hat{g}(\xi)\hat{f}(0) - \hat{g}(0)\hat{f}(\xi)\right] \mathrm{d}\sigma,\end{aligned}$$

where $\xi^\pm = \frac{\xi \pm |\xi|\sigma}{2}$.

Remark. In view of inequality (1.32), we can also define weak solutions by the weak formulation (1.28) of the Boltzmann collision operator. This leads to the Bobylev identity

$$\widehat{Q(g, f)}(\xi) = \int_{\mathbb{S}^{d-1}} b\left(\frac{\xi}{|\xi|} \cdot \sigma\right) \left(\hat{g}(\xi^-)\hat{f}(\xi^+) - \hat{g}(0)\hat{f}(\xi)\right) \mathrm{d}\sigma \tag{1.40}$$

for Maxwellian molecules. This is actually the form of Bobylev's identity we will be using in the following.

A simple, but in a sense important, consequence of Bobylev's identity (1.40) is that, for all $d \geq 1$,

$$P_\Lambda Q(g, f) = P_\Lambda Q(P_\Lambda g, P_\Lambda f) \tag{1.41}$$

where, for convenience, we put $P_\Lambda := \mathbb{1}_\Lambda(D_v)$ for the orthogonal projection onto Fourier 'modes' $|\eta| \leq \Lambda$. This (quasi-)locality property will turn out to be crucial in our inductive strategy to prove the strong smoothing properties of the homogeneous Boltzmann equation.

We will get back to Bobylev's identity in Chapter 5, where the non-Maxwellian version of Theorem 1.5 will be briefly discussed. In particular, we will give some details on why the proof of our main result cannot simply be extended to this case. ♦

Proof. Let $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and assume that $f, g \in L^1_2(\mathbb{R}^d)$. Then by the definition of the distributional Fourier transform and the weak formulation of Boltzmann's collision operator, we have

$$\begin{aligned}
&\langle \widehat{Q(g, f)}, \varphi\rangle = \langle Q(g, f), \widehat{\varphi}\rangle \\
&= \frac{1}{2}\int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} b\left(\frac{v - v_*}{|v - v_*|}\cdot\sigma\right) g(v_*)f(v)\left[\widehat{\varphi}(v_*') + \widehat{\varphi}(v') - \widehat{\varphi}(v_*) - \widehat{\varphi}(v)\right] \mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma \\
&= \frac{1}{2}\int_{\mathbb{R}^d} \mathrm{d}\xi \int_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} \mathrm{d}v\mathrm{d}v_*\mathrm{d}\sigma\, b\left(\frac{v - v_*}{|v - v_*|}\cdot\sigma\right) g(v_*)f(v)\varphi(\xi) \\
&\qquad\qquad \times\left[e^{-2\pi \mathrm{i}\xi\cdot v_*'} + e^{-2\pi\mathrm{i}\xi\cdot v'} - e^{-2\pi\mathrm{i}\xi\cdot v} - e^{-2\pi\mathrm{i}\xi\cdot v}\right] \\
&= \frac{1}{2}\int_{\mathbb{R}^d} \mathrm{d}\xi \int_{\mathbb{R}^{2d}} \mathrm{d}v\mathrm{d}v_*\, g(v_*)f(v)\varphi(\xi)\, e^{-2\pi\mathrm{i}\xi\cdot\frac{v+v_*}{2}} \\
&\quad \int_{\mathbb{S}^{d-1}} \mathrm{d}\sigma\, b\left(\frac{v - v_*}{|v - v_*|}\cdot\sigma\right)\left[\mathrm{e}^{2\pi\mathrm{i}\xi\cdot\frac{|v-v_*|}{2}\sigma} + e^{-2\pi\mathrm{i}\xi\cdot\frac{|v-v_*|}{2}\sigma} - e^{-2\pi\mathrm{i}\xi\cdot\frac{v-v_*}{2}} - \mathrm{e}^{2\pi\mathrm{i}\xi\cdot\frac{v-v_*}{2}}\right].
\end{aligned}$$

The key observation of Bobylev was that

$$\int_{\mathbb{S}^{d-1}} F(k\cdot\sigma, \ell\cdot\sigma)\,\mathrm{d}\sigma = \int_{\mathbb{S}^{d-1}} F(\ell\cdot\sigma, k\cdot\sigma)\,\mathrm{d}\sigma$$

for any integrable F and unit vectors $k, l \in \mathbb{S}^{d-1}$. Notice that due to cancellations,

$$\int_{\mathbb{S}^{d-2}} \left[\mathrm{e}^{2\pi\mathrm{i}\xi\cdot\frac{|v-v_*|}{2}\sigma} + e^{-2\pi\mathrm{i}\xi\cdot\frac{|v-v_*|}{2}\sigma} - e^{-2\pi\mathrm{i}\xi\cdot\frac{v-v_*}{2}} - \mathrm{e}^{2\pi\mathrm{i}\xi\cdot\frac{v-v_*}{2}}\right]\mathrm{d}\omega$$

is of order θ^2 for small θ, which regularises the singularity in b, see the calculation leading to (1.31). In particular, the above remark for $k = \frac{v-v_*}{|v-v_*|}$ and $\ell = \frac{\xi}{|\xi|}$ yields

$$\begin{aligned}
&\int_{\mathbb{S}^{d-1}} \mathrm{d}\sigma\, b\left(\frac{v - v_*}{|v - v_*|}\cdot\sigma\right)\left[\mathrm{e}^{2\pi\mathrm{i}\xi\cdot\frac{|v-v_*|}{2}\sigma} + e^{-2\pi\mathrm{i}\xi\cdot\frac{|v-v_*|}{2}\sigma} - e^{-2\pi\mathrm{i}\xi\cdot\frac{v-v_*}{2}} - \mathrm{e}^{2\pi\mathrm{i}\xi\cdot\frac{v-v_*}{2}}\right] \\
&= \int_{\mathbb{S}^{d-1}} \mathrm{d}\sigma\, b\left(\frac{\xi}{|\xi|}\cdot\sigma\right)\left[\mathrm{e}^{2\pi\mathrm{i}\frac{|\xi|}{2}(v-v_*)\cdot\sigma} + e^{-2\pi\mathrm{i}\frac{|\xi|}{2}(v-v_*)\cdot\sigma} - e^{-2\pi\mathrm{i}\xi\cdot\frac{v-v_*}{2}} - \mathrm{e}^{2\pi\mathrm{i}\xi\cdot\frac{v-v_*}{2}}\right].
\end{aligned}$$

Introducing $\xi^\pm = \frac{\xi \pm |\xi|\sigma}{2}$, it then follows that

$$
\begin{aligned}
\langle \widehat{Q(g,f)}, \varphi \rangle = \frac{1}{2} \int_{\mathbb{R}^d} \mathrm{d}\xi \int_{\mathbb{R}^{2d} \times \mathbb{S}^{d-1}} \mathrm{d}v \mathrm{d}v_* \mathrm{d}\sigma \, b\left(\frac{\xi}{|\xi|} \cdot \sigma \right) g(v_*) f(v) \varphi(\xi) \\
\times \left[e^{-2\pi \mathrm{i} v \cdot \xi^-} e^{-2\pi \mathrm{i} v_* \cdot \xi^+} + e^{-2\pi \mathrm{i} v \cdot \xi^+} e^{-2\pi \mathrm{i} v_* \cdot \xi^-} - e^{-2\pi \mathrm{i} v \cdot \xi} - e^{-2\pi \mathrm{i} v_* \cdot \xi} \right] \\
= \frac{1}{2} \int_{\mathbb{R}^d} \mathrm{d}\xi \, \varphi(\xi) \int_{\mathbb{S}^{d-1}} \mathrm{d}\sigma \, b\left(\frac{\xi}{|\xi|} \cdot \sigma \right) \\
\times \left[\hat{g}(\xi^+) \hat{f}(\xi^-) + \hat{g}(\xi^-) \hat{f}(\xi^+) - \hat{g}(0) \hat{f}(\xi) - \hat{g}(\xi) \hat{f}(0) \right].
\end{aligned}
$$

This identifies the distributional Fourier transform of $Q(g,f)$. ■

CHAPTER 2

Smoothing properties of the Boltzmann collision operator

It has long been suspected that the non-cutoff Boltzmann operator with a singular cross section kernel of the form (1.16) has similar coercivity properties to the fractional Laplacian $(-\Delta)^\nu$, for suitable $0 < \nu < 1$:

$$-Q(g, f) \approx (-\Delta)^\nu f + \text{lower order terms},$$

that is, it behaves similar to a singular integral operator with leading term proportional to a fractional Laplacian. The intuition has been made precise by ALEXANDRE, DESVILLETTES, VILLANI, and WENNBERG [ADVW00], see also the reviews by ALEXANDRE [Ale09] and by VILLANI [Vil02] for the idea's history. In terms of compactness properties this has been noticed for the linearised Boltzmann kernel as early as in [Pa074] and for the nonlinear Boltzmann kernel in [Lio94a, Lio94b]. The suspicion has led to the hope that the fully nonlinear homogenous Boltzmann equation enjoys regularity properties similar to the heat equation with a fractional Laplacian given by

$$\begin{cases} \partial_t u + (-\Delta)^\nu u &= 0 \\ \qquad\quad u|_{t=0} &= u_0 \in L^1(\mathbb{R}^d). \end{cases}$$

Using the Fourier transform one immediately sees that

$$\widehat{u}(t, \xi) = \mathrm{e}^{-t(2\pi|\xi|)^{2\nu}} \widehat{u_0}(\xi) \quad \text{with} \quad \widehat{u_0} \in L^\infty(\mathbb{R}^d),$$

so

$$\sup_{t>0} \sup_{\xi\in\mathbb{R}^d} \mathrm{e}^{t|\xi|^{2\nu}} |\widehat{u}(t, \xi)| \leq \|u_0\|_{L^1(\mathbb{R}^d)} < \infty,$$

that is, the Fourier transform of the solution is extremely fast decaying for strictly positive times, that is, solutions of the fractional heat equation gain a high amount of regularity for arbitrary positive times.

This is in sharp contrast to the fact that in the Grad cutoff case there cannot be any smoothing effect. Instead, regularity and singularities of the initial datum get propagated in this case, see, for example, [MV04].

The discussion about solutions of the heat equation with a fractional Laplacian motivates the following definition of Gevrey spaces, which give a convenient framework to describe this smoothing by interpolating between smooth and (ultra-)analytic functions.

Definition 2.1. Let $s > 0$. A function $f \in L^1(\mathbb{R}^d)$ belongs to the Gevrey class $G^s(\mathbb{R}^d)$, if there exists an $\epsilon_0 > 0$ such that[1]

$$e^{\epsilon_0 \langle D_v \rangle^{1/s}} f \in L^2(\mathbb{R}^d), \quad \text{where} \quad \langle D_v \rangle = \left(1 + |D_v|^2\right)^{1/2}.$$

Thus, $G^1(\mathbb{R}^d)$ is the space of real analytic functions, and $G^s(\mathbb{R}^d)$ for $s \in (0,1)$ the space of ultra-analytic functions.

Equivalently[2], $f \in G^s(\mathbb{R}^d)$ if $f \in \mathscr{C}^\infty(\mathbb{R}^d)$ and there exists a constant $C > 0$ such that for all $k \in \mathbb{N}_0$ one has

$$\|D^k f\|_{L^2(\mathbb{R}^d)} \leq C^{k+1}(k!)^s,$$

where $\|D^k f\|_{L^2}^2 = \sup_{|\beta|=k} \|\partial^\beta f\|_{L^2}^2$.

It is therefore natural to believe, as conjectured in [DW04], that weak solutions to the non-cutoff Boltzmann equation gain a certain amount of smoothness, and even analyticity, for any $t > 0$.

Conjecture (Gevrey smoothing)**.** *Any weak solution of the non-cutoff homogenous Boltzmann equation with a singular cross section kernel of order ν and with initial datum in $L^1_2(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$, i.e., finite mass, energy and entropy, belongs to the Gevrey class $G^{\frac{1}{2\nu}}(\mathbb{R}^d)$ for strictly positive times.*

The central result of this part of the thesis is a proof of this conjecture for Maxwellian molecules. In particular, we prove

Theorem. *Assume that the non-cutoff Boltzman cross section has a singularity $1+2\nu$ with $0 < \nu < 1$ and obeys some further technical conditions, which are true in all physically relevant cases, for details see (1.18) and (2.5). Then, for initial conditions $f_0 \in L\log L \cap L^1_m$ with an integer*

$$m \geq \max\left(2, \frac{2^\nu - 1}{2(2-2^\nu)}\right)$$

any weak solution of the fully non-linear homogenous Boltzmann equation for Maxwellian molecules belongs to the Gevrey class $G^{\frac{1}{2\nu}}$ for strictly positive times.

[1] We denote $D_v = -\frac{i}{2\pi}\nabla$ and for a suitable function $G : \mathbb{R}^d \to \mathbb{C}$ we define the operator $G(D_v)$ as a Fourier multiplier, that is, $G(D_v)f := \mathscr{F}^{-1}[G\hat{f}]$.

[2] Regarding equivalency, see, for example, Theorem 4 in [LO97].

In particular, for $\nu \le \log(9/5)/\log(2) \simeq 0.847996$ *we have* $m = 2$ *and the theorem does not require anything except the physically reasonable assumptions of finite mass, energy, and entropy. If* $\log(9/5)/\log(2) < \nu < 1$ *and we assume only that* $f_0 \in L \log L \cap L^1_2$*, then we prove that the solution is in* $G^{\frac{\log 2}{2\log(9/5)}}$*, in particular, that it is ultra-analytic.*

1. For a more precise formulation of our results, see Theorems 2.2, 2.4, and 2.5 for the case $m = 2$ and Theorems 3.27, 3.28, and 3.29 below.

2. We would like to stress that our results cover both the weak and strong singularity regimes, where $0 < \nu < 1/2$, respectively $1/2 \le \nu < 1$.

3. The theorem above applies to all dimensions $d \ge 1$. The physical case for Maxwellian molecules in dimension $d = 3$ is $\nu = 1/4$.

The main problem for establishing Gevrey regularity is that, in order to use the coercivity results of Alexandre, Desvillettes, Villani and Wennberg [ADVW00], one has to bound a non-linear and non-local commutator of the Boltzmann kernel with certain sub-Gaussian Fourier multipliers. The main ingredient in our proof is a new way of estimating this non-local and nonlinear commutator.

In a similar way, one can look at the homogeneous Boltzmann equation for Maxwellian molecules and an angular singularity of Debye-Yukawa type (1.19).

In this case, the singularity and thus the coercive effects are much weaker and of the form

$$-Q(g, f) \approx (\log(1 - \Delta))^{\mu+1} f + \text{lower order terms}, \tag{2.1}$$

as was noticed in [MUXY09], see also Section 2.4 below. However, one should still expect a smoothing effect similar to the logarithmic type heat equation

$$\begin{cases} \partial_t u + (\log(1-\Delta))^{\mu+1} u & = 0 \\ \qquad\qquad\qquad u|_{t=0} & = u_0 \in L^1(\mathbb{R}^d), \end{cases}$$

where the solutions satisfy

$$\sup_{t>0} \sup_{\xi \in \mathbb{R}^d} \mathrm{e}^{t \log(1+|\xi|^2)^{\mu+1}} |\widehat{u}(t,\xi)| \le \|u_0\|_{L^1(\mathbb{R}^d)} < \infty.$$

Higher regularity of weak solutions

The first regularisation results in this direction were due to Desvillettes for the spatially homogeneous non-cutoff Kac equation [Des95] and the homogeneous non-cutoff Boltzmann equation for Maxwellian molecules in two dimensions [Des97], where $\mathscr{C}^\infty$ regularisation is proved. Later, Desvillettes and Wennberg [DW04] proved under rather general assumptions on the collision cross-section (excluding Maxwellian molecules, though) regularity in Schwartz space of weak solutions to the non-cutoff homogeneous Boltzmann equation. By quite different methods, using Littlewood-Paley decompositions, Alexandre and El Safadi [AE05] showed that the assumptions on the cross-section (1.18)-(1.20) imply that the solutions are in H^∞ for any positive time $t > 0$. By moment propagation results for Maxwellian molecules (see Truesdell [Tru56]) this cannot be improved to regularity in Schwartz space.

For collision cross-sections corresponding to Debye-Yukawa-type interaction potentials, Morimoto, Ukai, Xu and Yang [MUXY09] proved the same H^∞ regularising effect using suitable test functions in the weak formulation of the problem.

The question of the local existence of solutions in Gevrey spaces for Gevrey regular initial data with additional strong decay at infinity was first addressed in 1984 by Ukai [Uka84], both in the spatially homogeneous and inhomogeneous setting.

We are interested in the Gevrey *smoothing effect*, namely that under the (physical) assumptions of finite mass, energy and entropy of the initial data, weak solutions of the homogeneous Boltzmann equation without cutoff are Gevrey functions for any strictly positive time. This question was treated in the case of the *linearised* Boltzmann equation in the homogeneous setting by Morimoto *et al.* [MUXY09], where they proved that, given $0 < \nu < 1$, weak solutions of the linearised Boltzmann equation belong to the space $G^{\frac{1}{\nu}}(\mathbb{R}^3)$ for any positive time. In [LMPX14], radially symmetric perturbations $g = g(|v|)$ around a global Maxwellian $\mu(v) = (2\pi)^{-\frac{3}{2}} e^{-\frac{|v|^2}{2}}$, that is, for f in (1.2),

$$f(v) = \mu(v) + \sqrt{\mu(v)}\, g(v), \quad g(v) = g(|v|),$$

were studied by using eigenfunctions of the linearised Boltzmann operator $\mathscr{L}$, where

$$\mathscr{L} g = -\mu^{-\frac{1}{2}} Q(\mu, \mu^{\frac{1}{2}} g) - \mu^{-\frac{1}{2}} Q(\mu^{\frac{1}{2}} g, \mu).$$

In this setting, the authors obtained a Gelfand-Shilov smoothing effect, which includes Gevrey regularity.

For the non-Maxwellian Boltzmann operator, Gevrey regularity was proved under very strong unphysical decay assumptions on the initial datum in [Lin14].

For radially symmetric solutions, the homogeneous non-cutoff Boltzmann equation for Maxwellian molecules is related to the homogeneous non-cutoff Kac equation. The non-cutoff Kac equation was introduced by Desvillettes in [Des95], where first regularity results were established, see also Desvillettes' review [Des03]. For this equation, the best available results so far are due to Lekrine and Xu [LX09] and Glangetas and Najeme [GN13]: Lekrine and Xu [LX09] proved Gevrey regularisation of order $\frac{1}{2\alpha}$ for mild singularities $0 < \nu < \frac{1}{2}$ and all $0 < \alpha < \nu$. Strong singularities $\frac{1}{2} \leq \nu < 1$ were treated by Glangetas and Najeme [GN13], where they prove that for $\nu = \frac{1}{2}$ the solution becomes Gevrey regular of order $\frac{1}{2\alpha}$ for any $0 < \alpha < \frac{1}{2}$ and Gevrey regular of order 1, that is, analytic, when $\frac{1}{2} < \nu < 1$. Thus, in the critical case $\nu = \frac{1}{2}$, the result of [GN13] misses the analyticity of weak solutions and does not prove ultra-analyticity in the range $\frac{1}{2} < \nu < 1$. Moreover, both results are obtained under the *additional* moment assumption $f_0 \in L^1_{2+2\nu}(\mathbb{R})$.

Ultra-analyticity results have previously been obtained by Morimoto and Xu [MX09] for the homogeneous Landau equation in the Maxwellian molecules case and related simplified models in kinetic theory. The analysis of smoothing properties of Landau equation is quite different from the Boltzmann and Kac equations. The Landau equation explicitly contains a second order elliptic term, which yields coercivity, and, more importantly, certain commutators with weights in Fourier space are identically zero, which simplifies the analysis tremendously, see Proposition 2.2 in [MX09].

For the nonlinear non-cutoff homogeneous Boltzmann equation some partial results regarding Gevrey regularisation were obtained by Morimoto and Ukai [MU10], including the non-Maxwellian molecules case, but under the strong additional assumptions of Maxwellian decay and smoothness of the solution. Still with these strong decay assumptions, Yin and Zhang [YZ12, YZ14] extended this result to a larger class of kinetic cross-sections.

We stress that for the main result of our paper the initial datum is *only* assumed to obey the *natural* assumptions coming from physics, i.e., finiteness of mass, energy and entropy.

2.1 Absence of smoothing in the Grad cutoff case

Before we formulate the main results about smoothing properties of the Boltzmann and Kac equations, we quickly discuss the Grad cut-off case, that is

$$\int_{\mathbb{S}^{d-1}} b(\cos\theta)\,\mathrm{d}\sigma =: a < \infty.$$

In this case, the Boltzmann collision operator can be split into a *gain part* $Q^+(f,f)$ and a *loss part* $Q^-(f,f)$, according to the sign in the definition of the Boltzmann operator (1.3),

$$Q(f,f) = Q^+(f,f) - Q^-(f,f) = Q^+(f,f) - f(Lf),$$

where $Lf = a\int_{\mathbb{R}^d} f(v)\,\mathrm{d}v$. Solutions of the homogeneous Boltzmann equation

$$\partial_t f + f(Lf) = Q^+(f,f)$$

are then given by Duhamel's formula

$$f(t,v) = \mathrm{e}^{-\int_0^t Lf(\tau,v)\,\mathrm{d}\tau} f_0(v) + \int_0^t \mathrm{e}^{-\int_s^t Lf(\tau,v)\,\mathrm{d}\tau} Q^+(f,f)(s,v)\,\mathrm{d}s. \tag{2.2}$$

While one can show that the operator $Q^+(f,f)$ has smoothing properties, equation (2.2) shows that the solution f in the Grad cut-off case will at most be as regular as the initial datum f_0, that is $f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$ in our case.[3]

[3]However, one can show that the solution can be split into a part that lies in the Sobolev space H^s, s arbitrarily large, and one that has the regularity of the inital datum, but decays exponentially, see [MV04] for more details in a non-Maxwellian setting and the classical treatise by Wild [Wil51].

This observation shows the importance of the singularity in the angular collision kernel for grazing collisions $\theta \to 0$ when it comes to smoothing properties of the Boltzmann equation.

2.2 Gevrey smoothing for Maxwellian Molecules with Power-law Interaction

In this section we formulate the main theorems regarding Gevrey smoothing for the homogeneous Boltzmann and Kac equations for Maxwellian molecules with singularity of power-law type.

Given $\beta > 0$ and $\alpha \in (0,1)$ we define the Gevrey multiplier $G : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ by

$$G(t,\eta) := \mathrm{e}^{\beta t \langle \eta \rangle^{2\alpha}}$$

and for $\Lambda > 0$ the cut-off Gevrey multiplier $G_\Lambda : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ by

$$G_\Lambda(t,\eta) := G(t,\eta)\mathbb{1}_\Lambda(|\eta|),$$

where $\mathbb{1}_\Lambda$ is the characteristic function of the interval $[0,\Lambda]$. The associated Fourier multiplication operator is denoted by $G_\Lambda(t, D_v)$,

$$(G_\Lambda(t,D_v)f)(t,v) := \int_{\mathbb{R}^d} G_\Lambda(t,\eta)\hat{f}(t,\eta)\,\mathrm{e}^{2\pi i\eta\cdot v}\,\mathrm{d}\eta = \mathscr{F}^{-1}\left[G_\Lambda(t,\cdot)\hat{f}(t,\cdot)\right].$$

Note that, since $G_\Lambda(t,\cdot)$ has compact support in $\mathbb{R}^d_\eta$ for any $t > 0$, one has

$$G_\Lambda f, G_\Lambda^2 f \in L^\infty([0,T_0]; H^\infty(\mathbb{R}^d))$$

for any finite $T_0 > 0$ and $\Lambda > 0$, if $f \in L^\infty([0,T_0]; L^1(\mathbb{R}^d))$. This holds since

$$\|G_\Lambda f\|^2_{H^s(\mathbb{R}^d_v)} \le \|\hat{f}\|^2_{L^\infty(\mathbb{R}^d_\eta)} \|\langle\cdot\rangle^s G_\Lambda(t,\cdot)\|^2_{L^2(\mathbb{R}^d_\eta)} \le \|f\|^2_{L^1(\mathbb{R}^d_v)} \|\langle\cdot\rangle^s G_\Lambda(T_0,\cdot)\|^2_{L^2(\mathbb{R}^d_\eta)}$$

for all $s \ge 0$. These functions, due to the cut-off in Fourier space, are even analytic in a strip containing $\mathbb{R}^d_v$.

Theorem 2.2 (Gevrey smoothing I). *Let $0 < \nu < 1$ and assume that the cross-section b satisfies conditions (B1)–(B3) with power-law type singularity (1.18) for $d \ge 2$. For $d = 1$, assume that b_1 satisfies conditions (K1)–(K3). Let f be a weak solution of the Cauchy problem (1.2) with initial datum $0 \le f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$. Then, for all $0 < \alpha \le \min\{\alpha_{2,d}, \nu\}$,*

$$f(t,\cdot) \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d) \tag{2.3}$$

for all $t > 0$, where $\alpha_{2,d} = \frac{\log[(8+d)/(4+d)]}{\log 2}$.

Remarks 2.3. (i) In numbers,

$$\alpha_{2,1} \simeq 0.847997, \quad \alpha_{2,2} \simeq 0.736966, \quad \text{and} \quad \alpha_{2,3} \simeq 0.652077.$$

This means, that under *only* physically reasonable assumptions of finite mass, energy, and entropy, weak solutions are analytic for $\nu \geq \frac{1}{2}$ and even ultra-analytic if $\nu > \frac{1}{2}$. It is easy to see that $\alpha_{2,d}$ is decreasing in d and for $d = 6$, $\alpha_{2,6} \simeq 0.485427$. Hence, for $d \geq 6$, analyticity (respectively ultra-analyticity) does not follow from this theorem.

(ii) For the proof of Theorem 2.2 (and also 2.4 and 2.5 below) it is important that the energy of f is bounded, which enters in the technical Lemma 3.7 and its Corollary 3.8. A considerably simpler proof could be given using only that $f \in L^1_1(\mathbb{R}^d)$. In this case, $\alpha_{2,d}$ is replaced by $\alpha_{1,d} = \frac{\log[(4+d)/(2+d)]}{\log 2}$ (see also Remark 2.2 below). However, $\alpha_{1,3} < 0.4855$ in three dimensions, thus we would not be able to conclude (ultra-)analytic smoothing of weak solutions for strong singularities $\frac{1}{2} \leq \nu < 1$.

(iii) As our theorem above shows, weak solutions of the homogenous Kac equation become Gevrey regular for strictly positive times for moderately singular collision kernels with singularity $\nu \in (0, \frac{1}{2})$, see (1.21) for the precise description of the singularity. For $\nu = \frac{1}{2}$ they become analytic, which improves the result of Glangetas and Najeme [GN13] in this critical case. They even become ultra-analytic for $\nu \in (\frac{1}{2}, 1)$.

(iv) Rotationally symmetric solutions f corresponding to rotationally symmetric initial conditions f_0 are Gevrey regular for strictly positive times under the same conditions as in the one-dimensional case $d = 1$. The proof is exactly as the proof of Theorem 3.27 with some small changes in the proof of Lemma 3.15, where the independence of the solution f on the angular coordinates can be explicitly used with the $n = 1$ version of Corollary 3.8.

As already remarked, the result of Theorem 2.2 deteriorates in the dimension. Under the same assumptions, but using quite a bit more structure of the Boltzmann operator, we can prove a dimension independent version. Its proof, however, is considerably more involved than the proof of Theorem 2.2.

Theorem 2.4 (Gevrey smoothing II). *Let $d \geq 2$. Assume that the cross-section b satisfies the conditions of Theorem 2.2. Let f be a weak solution of the Cauchy problem (1.2) with initial datum $0 \leq f_0 \in L^1_2 \cap L \log L(\mathbb{R}^d)$. Then, for all $0 < \alpha \leq \min\{\alpha_{2,2}, \nu\}$,*

$$f(t, \cdot) \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d) \tag{2.4}$$

for all $t > 0$, where $\alpha_{2,2} = \frac{\log(5/3)}{\log 2} \simeq 0.736966$. In particular, in contrast to Theorem 2.2, the weak solution is real analytic if $\nu = \frac{1}{2}$ and ultra-analytic if $\nu > \frac{1}{2}$ in any dimension.

If the integrability condition (1.20) is replaced by the slightly stronger condition that $b(\cos\theta)$ is bounded away from $\theta = 0$, that is,

$$\begin{aligned} &\text{for any } 0 < \theta_0 < \tfrac{\pi}{2} \text{ there exists } C_{\theta_0} < \infty \text{ such that} \\ &0 \le b(\cos\theta) \le C_{\theta_0} \text{ for all } \theta_0 \le \theta \le \tfrac{\pi}{2}, \end{aligned} \tag{2.5}$$

which is true in all physically relevant cases, we can prove an even stronger result.

Theorem 2.5 (Gevrey smoothing III)**.** *Let $d \ge 2$. Assume that the cross-section b satisfies the conditions of Theorem 2.2 and the condition (2.5), that is, it is bounded away from the singularity. Let f be a weak solution of the Cauchy problem (1.2) with initial datum $0 \le f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$. Then, for all $0 < \alpha \le \min\{\alpha_{2,1}, \nu\}$,*

$$f(t,\cdot) \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d) \tag{2.6}$$

for all $t > 0$, where $\alpha_{2,1} = \frac{\log(9/5)}{\log 2} \simeq 0.847997$.

Remark. (i) Since we do not rely on interpolation inequalities between Sobolev spaces, our results also include the *limiting case* $\alpha = \nu$, at least if $\nu \le \alpha_{2,n}$ $(n = d, 2, 1)$. This is in contrast to all previous results on smoothing properties of the Boltzmann and Kac equations.

(ii) If higher moments of the initial datum are bounded (and thus stay bounded eternally due to moment propagation results, see, for instance, Villani's review [Vil02]), the results in Theorem 2.4 and Theorem 2.5 can be improved in the high singularity case, where ν is close to one. Namely, let $f_0 \in L\log L \cap L^1_m(\mathbb{R}^d)$ for some integer $m > 2$, then the constants $\alpha_{2,d}$, $\alpha_{2,2}$, respectively $\alpha_{2,1}$, are replaced by $\alpha_{m,n} = \frac{\log[(4m+n)/(2m+n)]}{\log 2}$ $(n = d, 2, 1)$, which are strictly increasing towards the limit $\alpha_{\infty,n} = 1$ as m becomes large. See Theorems 3.27, 3.28 and 3.29 below.

♦

Moreover, we prove that for very strong singularities ν, we can prescribe precise conditions on the initial datum such that we have $f \in G^{\frac{1}{2\nu}}(\mathbb{R}^d)$.

Theorem 2.6. *Given $0 < \nu < 1$, there is $m(\nu)$ such that, if $m \in \mathbb{N}$ and $m \ge m(\nu)$ and $f_0 \in L\log L \cap L^1_m$, the weak solution is in $G^{\frac{1}{2\nu}}(\mathbb{R}^d)$ for all $t > 0$.*

More precisely, under the conditions of Theorem 2.2 having $m \ge \max\left(2, \frac{2^\nu - 1}{2 - 2^\nu}\right)$ yields Gevrey smoothing of order $\frac{1}{2\nu}$ and under the slightly stronger conditions of Theorem 2.5 having $m \ge \max\left(2, \frac{2^\nu - 1}{2(2 - 2^\nu)}\right)$ is enough.

Remark. The proof of this theorem follows directly from the results of Theorems 3.27, 3.28, and 3.29 in Section 3.7, which extend Theorems 2.2, 2.4, and 2.5 to the case of finite moments $m \ge 2$. ♦

The strategy of the proofs of our main results Theorems 2.2, 2.4, and 2.5 is as follows: we start with the additional assumption $f_0 \in L^2$ on the initial datum. We use the known H^∞ smoothing of the non-cutoff Boltzmann and Kac equation to allow this. This yields an L^2 reformulation of the weak formulation of the Boltzmann and Kac equations which includes suitable growing Fourier multipliers.

The inclusion of sub-Gaussian Fourier multipliers leads to a nonlocal and nonlinear commutator of the Boltzmann and Kac kernels, which turns out to be a three-linear expression in the weighted solution $\hat{f}$ on the Fourier side. In order to bound this expression with L^2 norms, one of the three terms has to be controlled pointwise, *including* a sub-Gaussian growing factor, see Proposition 3.5. The problem is that one has to control the pointwise bound with an L^2 norm, which is in general impossible. To overcome this obstacle there are several important technical steps:

(1) When working on a ball of radius Λ, we need this uniform control only on a ball of radius $\Lambda/\sqrt{2}$, which enables an inductive procedure.

(2) Using the additional a priori information that the kinetic energy is finite, or, depending on the initial condition, even higher moments are finite, we transform weighted L^2 bounds into pointwise bounds on slightly smaller balls with an additional loss of power in the weights in Fourier space. Here we rely on Kolmogorov-Landau type inequalities, see Lemma 3.9 and Appendix C.

(3) We use the strict concavity of the Fourier multipliers, see Lemma 3.3, in order to compensate for this loss of power.

(4) Averaging over a codimension 2 sphere in the proof of Theorem 2.4, we get, in any dimension, the same results as for the two dimensional Boltzmann equation.

(5) Averaging over a codimension 1 set constructed from a codimension 2 sphere and the collision angles θ away from the singularity, and using the fact that near the singularity one of the three Fourier weights is not big due to Lemma 3.3, we can even get, in any dimension, the same results as for the one-dimensional Kac equation under the conditions of Theorems 2.5 and 3.29.

2.3 Strong smoothing for Maxwellian Molecules with Debye-Yukawa type Interaction

Based upon the proof of Gevrey smoothing for the homogeneous Boltzmann equation with Maxwellian molecules and angular singularity of the inverse-power law type (1.18), we can show a stronger than H^∞ regularisation property of weak solutions in the Debye-Yukawa case.

In [MUXY09] it has been shown that weak solutions to the Cauchy problem (1.2) with Debye-Yukawa type interactions enjoy an H^∞ smoothing property, i.e. starting with arbitrary initial datum $f_0 \geq 0$, $f_0 \in L^1_2 \cap L \log L(\mathbb{R}^d)$, one has $f(t, \cdot) \in H^\infty$ for any positive time $t > 0$.

Proposition 2.7 (H^∞ smoothing of weak solutions, Theorem 1.1 in [MUXY09])**.** *Assume that* $0 \leq f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$ *and the collision kernel satisfies Assumptions* **(B1)–(B3)** *with Debye-Yukawa type singularity* (1.19)*. Let* f *be a weak solution of the Cauchy problem* (1.2) *satisfying*

$$\sup_{t\in[0,T]} \int f(t,v)\left(1+|v|^2+\log(1+f(t,v))\right)\,\mathrm{d}v < \infty$$

for some $T > 0$*. Then for any* $0 < t \leq T$*, we have*

$$f(t,\cdot) \in H^\infty.$$

Remark. (1) Even though the above theorems are stated in dimension $d = 3$ in the references, their proofs can be done in any dimension $d \geq 2$.

(2) Also the proof of H^∞ smoothing, valid in dimension $d \geq 2$ for the Boltzmann equation, respectively, the Kac equation in dimension $d = 1$, in Appendix B can be easily transferred to the Debye-Yukawa case.

(3) This kind of logarithmic regularity effect was first observed by Morimoto for infinitely degenerate elliptic operators [Mor87].

♦

We will need the following function spaces to describe the gain of smoothness in the Debye-Yukawa case:

Definition 2.8. Let $\mu > 0$. A function $f \in H^\infty(\mathbb{R}^d)$ is defined to be in the space $\mathcal{A}^\mu(\mathbb{R}^d)$ if there exist constants $C > 0$ and $b > 0$ such that

$$\left\|\partial^\alpha f\right\|_{L^2} \leq C^{|\alpha|+1} \mathrm{e}^{b|\alpha|^{1+1/\mu}} \quad \text{for all } \alpha \in \mathbb{N}_0^d. \tag{2.7}$$

For $\mu > 0$ we define the family of function spaces, parametrised by $\tau > 0$,

$$\mathcal{D}\left(\mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}} : L^2(\mathbb{R}^d)\right) := \left\{f \in L^2(\mathbb{R}^d) : \mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}} f \in L^2(\mathbb{R}^d)\right\}$$

Lemma 2.9. *Let* $\mu > 0$*. Then*

$$\mathcal{A}^\mu(\mathbb{R}^d) = \bigcup_{\tau>0} \mathcal{D}\left(\mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}} : L^2(\mathbb{R}^d)\right).$$

The proof is rather technical and is deferred to Appendix D.

In view of the coercivity property (2.1) and the regularisation properties of the logarithmic heat equation

$$\partial_t f = -(\log(1-\Delta))^{\mu+1} f, \tag{2.8}$$

the spaces $\mathcal{A}^\mu$, through their Fourier characterisation in Lemma 2.9, capture exactly the gain of regularity that is to be expected for the Boltzmann equation with Debye-Yukawa type angular singularity. Indeed, our main result is

Theorem 2.10. *Let f be a weak solution of the Cauchy problem* (1.2) *for the homogeneous Bolzmann equation for Maxwellian molecules with angular collision kernel satisfying Assumptions* **(B1)–(B3)** *with Debye-Yukawa type singularity* (1.19), *and initial datum $f_0 \geq 0$, $f_0 \in L^1_2(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$. Then for any $T_0 > 0$ there exists $\beta > 0$ such that*

$$e^{\beta t \left(\log\langle D_v\rangle\right)^{\mu+1}} f(t,\cdot) \in L^2(\mathbb{R}^d) \tag{2.9}$$

for all $t \in (0, T_0]$. In particular, $f(t,\cdot) \in \mathcal{A}^\mu$ for all $t > 0$.

Furthermore, for all $0 \leq t \leq T_0$ there exists $M > 0$ such that with β as in (2.9),

$$\sup_{\eta\in\mathbb{R}^d} e^{\frac{2}{d+2}\beta t \left(\log\langle\eta\rangle\right)^{\mu+1}} |\hat{f}(t,\eta)| \leq M. \tag{2.10}$$

Remark. The bound (2.10) holds *uniformly* in $0 \leq t \leq T_0$, whereas (2.9) does not give uniform control on the L^2 norm of $e^{\beta t \left(\log\langle D_v\rangle\right)^{\mu+1}} f(t,\cdot)$ as $t \to 0$.

As $t \to 0$, $\|f(t,\cdot)\|_{L^2(\mathbb{R}^d)} \to \infty$ since the initial datum is only assumed to satisfy $f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$. Indeed, our proof strategy uses H^∞ smoothing of the weak solution first (see Proposition 2.7) and then proceeds with the additional assumption $f_0 \in L^2 \cap L^1_2 \cap L\log L(\mathbb{R}^d)$, see Theorem 4.4. As in Appendix B, we can show that the L^2 norm of f blows up at most polynomially though as $t \to 0$. ♦

Remark. This regularity is much weaker than the Gevrey regularity we prove for singular kernels of the form (1.18), but it is much stronger than the H^∞ smoothing shown in [MUXY09]. Moreover, it is *exactly* the right type of regularity one would expect for a coercive term of the form (2.1) from the analogy with the heat equation (2.8). We therefore expect this regularity result to be sharp. ♦

For our proof we have to choose β small if T_0 is large and our bounds on β deteriorate to zero in the limit $T_0 \to \infty$, so our Theorem 2.10 does not give a uniform result for all $t > 0$. Nevertheless, by propagation results due to Desvillettes, Furiolo and Terraneo [DFT09] we even get a uniform bound:

Corollary 2.11. *Under the same assumptions as in Theorem 2.10, for any weak solution f of the Cauchy problem* (1.2) *with initial datum $f_0 \geq 0$ and $f_0 \in L^1_2(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$, there exist constants $0 < K, C < \infty$ such that*

$$\sup_{0\leq t<\infty} \sup_{\eta\in\mathbb{R}^d} e^{K \min(t,1) \left(\log\langle\eta\rangle\right)^{\mu+1}} |\hat{f}(t,\eta)| \leq C. \tag{2.11}$$

The strategy of the proofs of our main result Theorem 2.10 is as follows: We start with the additional assumption $f_0 \in L^2$ on the initial datum (Theorem 4.4). We use the known H^∞ smoothing [MUXY09] of the non-cutoff Boltzmann equation to allow for this. Within an L^2 framework, a reformulation of the weak formulation of the Boltzmann equation is possible which includes suitable growing Fourier multipliers. As in [MUXY09], the inclusion of Fourier multipliers leads to a nonlocal and nonlinear commutator with the Boltzmann kernel. For non-power-type Fourier multipliers this commutator is considerably more complicated than the one encountered in the H^∞ smoothing case. To overcome this, we follow the strategy we developed in the power-law case (see Chapter 3), where an inductive procedure is invented to control the commutation error, in order to prove the Gevrey smoothing conjecture in the Maxwellian molecules case.

2.4 Coercivity of the Boltzmann collision operator

The coercive properties of the non-cutoff Boltzmann bilinear operator, which play the crucial role in the smoothing of solutions, are made precise in the following sub-elliptic estimate by Alexandre, Desvillettes, Villani and Wennberg [ADVW00]. We remark that, while the proof there is given for the Boltzmann equation, it equally applies to the Kac equation.

Proposition 2.12 (Sub-elliptic Estimate, [ADVW00, MUXY09]). *Let $g \in L^1_2(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$, $g \geq 0$ ($g \not\equiv 0$). Assume that the collision cross-section b satisfies Assumptions* **(B1)–(B3)***, respectively* **(K1)–(K3)***.*

Power-law case. *If the singularity of b for grazing collisions is of the type (1.18), then there exists a constant $C_g > 0$ (depending only on the dimension d, the collision kernel b, $\|g\|_{L^1_2}$ and $\|g\|_{L\log L}$) and a constant $C > 0$ (depending only on d and b), such that for any $f \in H^1(\mathbb{R}^d)$ one has*

$$-\langle Q(g,f), f\rangle \geq C_g \|f\|^2_{H^\nu} - \left(2^\nu C_g + C\|g\|_{L^1_2}\right)\|f\|^2_{L^2}.$$

Debye-Yukawa case. *If the singularity of b for grazing collisions is of the type (1.21), then there exists a positive constant C_g depending only on the dimension d, the collision kernel b, $\|g\|_{L^1_1}$ and $\|g\|_{L\log L}$ and constants $C > 0$, $R \geq \sqrt{e}$, depending only on the dimension d and on the collision kernel b, such that for all $\alpha \geq 0$ and all $0 \leq f \in H^1(\mathbb{R}^d)$ one has*

$$\begin{aligned}-\langle Q(g,f), f\rangle \geq & \frac{C_g}{(\log(\alpha+e))^{\mu+1}}\left\|\left(\log\langle D_v\rangle_\alpha\right)^{\frac{\mu+1}{2}} f\right\|^2_{L^2} \\ & - \left(C_g\left(\log R\right)^{\mu+1} + C\|g\|_{L^1}\right)\|f\|^2_{L^2}.\end{aligned}$$

Remark. As explained, for instance, in [AMUXY10], the constant C_g is an increasing function of $\|g\|_{L^1}$, $\|g\|^{-1}_{L^1_2}$ and $\|g\|^{-1}_{L\log L}$. In particular, if g is a weak solution of the Cauchy problem (1.2) with initial datum $g_0 \in L^1_2(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$, we have $\|g\|_{L^1} = \|g_0\|_{L^1}$, $\|g\|_{L^1_2} \leq \|g_0\|_{L^1_2}$ and $\|g\|_{L\log L} \leq \log 2\|g_0\|_{L^1} + H(g_0) + C_{\delta,d}\|g_0\|^{1-\delta}_{L^1_2}$, for small enough $\delta > 0$ (see (1.24)). This implies $C_g \geq C_{g_0}$ and thus

$$-\langle Q(g,f), f\rangle \geq C_{g_0}\|f\|^2_{H^\nu} - \left(2^\nu C_{g_0} + C\|g_0\|_{L^1_2}\right)\|f\|^2_{L^2},$$

respectively,

$$-\langle Q(g,f),f\rangle \geq \frac{C_{g_0}}{(\log(\alpha+\mathrm{e}))^{\mu+1}} \left\| (\log\langle D_v\rangle_\alpha)^{\frac{\mu+1}{2}} f \right\|_{L^2}^2 - \left(C_{g_0} (\log R)^{\mu+1} + C\|g_0\|_{L^1} \right) \|f\|_{L^2}^2,$$

uniformly in $t \geq 0$. ♦

As a first step in the proof of Proposition 2.12 we show

Lemma 2.13. *Assume that the assumptions of Proposition 2.12 hold. Then*

Power-law case. *There exists a constant $C_g' > 0$, depending only on b, the dimension d, and $\|g\|_{L^1}$, $\|g\|_{L^1_2}$, and $\|g\|_{L\log L}$, such that*

$$\int_{\mathbb{S}^{d-2}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right) \left(\hat{g}(0) - |\hat{g}(\eta^-)|\right) \,\mathrm{d}\sigma \geq C_g' |\eta|^{2\nu}\, \mathbb{1}_{\{|\eta|\geq 1\}}$$

for all $\eta \in \mathbb{R}^d$.

Debye-Yukawa case. *There exists a constant $C_g' > 0$, depending only on b, the dimension d, and $\|g\|_{L^1}$, $\|g\|_{L^1_2}$, and $\|g\|_{L\log L}$, as well as a constant $R \geq \sqrt{\mathrm{e}}$ depending only on d and b, such that*

$$\int_{\mathbb{S}^{d-2}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right) \left(\hat{g}(0) - |\hat{g}(\eta^-)|\right) \,\mathrm{d}\sigma \geq C_g' \left(\frac{\log\langle\eta\rangle_\alpha}{\log(\alpha+\mathrm{e})}\right)^{\mu+1} \mathbb{1}_{\{|\eta|\geq R\}}$$

for all $\eta \in \mathbb{R}^d$.

Proof of Lemma 2.13. Since $g \geq 0$, $g \in L^1_2 \cap L\log L$, there exists a constant $\widetilde{C}_g > 0$ such that for all $\eta \in \mathbb{R}^d$

$$\hat{g}(0) - |\hat{g}(\eta)| \geq \widetilde{C}_g \left(|\eta|^2 \wedge 1\right).$$

It is therefore enough to bound $\int_{\mathbb{S}^{d-1}} b(\frac{\eta}{|\eta|}\cdot\sigma)(|\eta^-|^2 \wedge 1)\,\mathrm{d}\sigma$. Recall that $|\eta^-|^2 = \frac{|\eta|^2}{2}\left(1 - \frac{\eta}{|\eta|}\cdot\sigma\right)$, and, choosing spherical coordinates with pole $\frac{\eta}{|\eta|}$ such that $\frac{\eta}{|\eta|}\cdot\sigma = \cos\theta$, we obtain

$$\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right) \left(|\eta^-|^2 \wedge 1\right) \,\mathrm{d}\sigma = |\mathbb{S}^{d-2}| \int_0^{\frac{\pi}{2}} \sin^{d-2}\theta\, b(\cos\theta) \left(|\eta|^2 \sin^2\frac{\theta}{2} \wedge 1\right) \,\mathrm{d}\theta$$
$$\geq \frac{|\mathbb{S}^{d-2}|}{4\pi^2} \int_0^{\frac{\pi}{2}} \sin^{d-2}\theta\, b(\cos\theta) \left(|\eta|^2\theta^2 \wedge 1\right) \,\mathrm{d}\theta.$$

Now, for the inverse power-law case, we use assumption (1.16) on the singularity: there exists a deviation angle $\theta_0 > 0$ small enough such that

$$\begin{aligned}
&\frac{|\mathbb{S}^{d-2}|}{4\pi^2}\int_0^{\frac{\pi}{2}} \sin^{d-2}\theta\, b(\cos\theta)\, \left(|\eta|^2\theta^2 \wedge 1\right)\, \mathrm{d}\theta \\
&\qquad \geq \frac{\kappa}{2}\frac{|\mathbb{S}^{d-2}|}{4\pi^2}\int_0^{\theta_0} \theta^{-1-2\nu}\, \left(|\eta|^2\theta^2 \wedge 1\right)\, \mathrm{d}\theta \\
&\qquad = |\eta|^{2\nu}\frac{\kappa}{2}\frac{|\mathbb{S}^{d-2}|}{4\pi^2}\int_0^{|\eta|\theta_0} \theta^{-1-2\nu}\, \left(\theta^2 \wedge 1\right)\, \mathrm{d}\theta.
\end{aligned}$$

For $|\eta| \geq 1$ we thus obtain

$$\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(|\eta^-|^2 \wedge 1\right)\, \mathrm{d}\sigma \geq C_g'|\eta|^{2\nu},$$

where

$$C_g' = \frac{\kappa}{2}\frac{|\mathbb{S}^{d-2}|}{4\pi^2}\int_0^{\theta_0} \theta^{-1-2\nu}\, \left(\theta^2 \wedge 1\right)\, \mathrm{d}\theta < \infty$$

in view of the momentum-transfer assumption **(B3)**.

In the Debye-Yukawa case, by assumption (1.19) on the singularity for grazing collisions on b, there exists a $\theta_0 > 0$ small enough such that

$$\begin{aligned}
&\frac{|\mathbb{S}^{d-2}|}{4\pi^2}\int_0^{\frac{\pi}{2}} \sin^{d-2}\theta\, b(\cos\theta)\, \left(|\eta|^2\theta^2 \wedge 1\right)\, \mathrm{d}\theta \\
&\qquad \geq \frac{\kappa}{2}\frac{|\mathbb{S}^{d-2}|}{4\pi^2}\int_0^{\theta_0} \theta^{-1}\left(\log\theta^{-1}\right)^{\mu}\, \left(|\eta|^2\theta^2 \wedge 1\right)\, \mathrm{d}\theta.
\end{aligned}$$

Let $R > 0$ be large enough, such that $\frac{1}{R} < \theta_0$. Then for $|\eta| \geq R$ we have

$$\begin{aligned}
&\frac{\kappa}{2}\frac{|\mathbb{S}^{d-2}|}{4\pi^2}\int_0^{\theta_0} \theta^{-1}\left(\log\theta^{-1}\right)^{\mu}\, \left(|\eta|^2\theta^2 \wedge 1\right)\, \mathrm{d}\theta \\
&\qquad \geq \frac{\kappa}{2}\frac{|\mathbb{S}^{d-2}|}{4\pi^2}\int_{\frac{1}{|\eta|}}^{\theta_0} \theta^{-1}\left(\log\theta^{-1}\right)^{\mu}\, \mathrm{d}\theta \\
&\qquad = \frac{\kappa}{2}\frac{|\mathbb{S}^{d-2}|}{4\pi^2}\frac{1}{\mu+1}\left[\left(\log|\eta|\right)^{\mu+1} - \left(\log\frac{1}{\theta_0}\right)^{\mu+1}\right] \\
&\qquad \geq C\left(\log|\eta|\right)^{\mu+1}
\end{aligned}$$

for some constant $C > 0$ depending only on the dimension and the collision kernel b. We conclude by noting that for all $|\eta| \geq \sqrt{e}$ one has

$$\log|\eta| = \frac{1}{2}\log|\eta|^2 \geq \frac{\log\langle\eta\rangle_\alpha}{\log(\mathrm{e}+\alpha)},$$

since for any $\alpha \geq 0$ the function $[\mathrm{e},\infty) \ni s \mapsto H(s) := \log s - \frac{\log(\alpha+s)}{\log(\alpha+\mathrm{e})}$ is non-decreasing with $H(e) = 0$. ∎

Proof of Proposition 2.12. We have $\langle Q(g,f),f\rangle = \mathrm{Re}\langle Q(g,f),f\rangle$ and by Bobylev's identity,

$$
\begin{aligned}
-\mathrm{Re}\langle Q(g,f),f\rangle &= \mathrm{Re}\int_{\mathbb{R}^d\times\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left[\overline{\hat{g}(0)\hat{f}(\eta)}-\overline{\hat{g}(\eta^-)\hat{f}(\eta^+)}\right]\hat{f}(\eta)\,\mathrm{d}\sigma\,\mathrm{d}\eta\\
&= \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left\langle\begin{pmatrix}\hat{f}(\eta)\\ \hat{f}(\eta^+)\end{pmatrix},\begin{pmatrix}2\overline{\hat{g}(0)} & -\hat{g}(\eta^-)\\ -\overline{\hat{g}(\eta^-)} & 0\end{pmatrix}\begin{pmatrix}\hat{f}(\eta)\\ \hat{f}(\eta^+)\end{pmatrix}\right\rangle_{\mathbb{C}^2}\mathrm{d}\sigma\,\mathrm{d}\eta\\
&= \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left\langle\begin{pmatrix}\hat{f}(\eta)\\ \hat{f}(\eta^+)\end{pmatrix},\begin{pmatrix}\overline{\hat{g}(0)} & -\hat{g}(\eta^-)\\ -\overline{\hat{g}(\eta^-)} & \hat{g}(0)\end{pmatrix}\begin{pmatrix}\hat{f}(\eta)\\ \hat{f}(\eta^+)\end{pmatrix}\right\rangle_{\mathbb{C}^2}\mathrm{d}\sigma\,\mathrm{d}\eta\\
&\quad - \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left\langle\begin{pmatrix}\hat{f}(\eta)\\ \hat{f}(\eta^+)\end{pmatrix},\begin{pmatrix}-\hat{g}(0) & 0\\ 0 & \hat{g}(0)\end{pmatrix}\begin{pmatrix}\hat{f}(\eta)\\ \hat{f}(\eta^+)\end{pmatrix}\right\rangle_{\mathbb{C}^2}\mathrm{d}\sigma\,\mathrm{d}\eta\\
&=: I_1 - I_2.
\end{aligned}
$$

To estimate $I_2 = \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\hat{g}(0)\left(|\hat{f}(\eta^+)|^2-|\hat{f}(\eta)|^2\right)\mathrm{d}\sigma\,\mathrm{d}\eta$, we do a change of variables $\eta^+\longrightarrow\eta$ as in [ADVW00] in the first part, treating b as if it were integrable, and using a limiting argument to make the calculation rigorous (this is a version of the cancellation lemma of [ADVW00] on the Fourier side). We then obtain with $\hat{g}(0)=\|g\|_{L^1}$

$$
I_2 = |\mathbb{S}^{d-2}|\int_0^{\pi/2}\sin^{d-2}\theta\, b(\cos\theta)\left[\frac{1}{\cos^d\frac{\theta}{2}}-1\right]\mathrm{d}\theta\,\|g\|_{L^1(\mathbb{R}^d)}\|f\|^2_{L^2(\mathbb{R}^d)}.
$$

In particular, since $\frac{1}{\cos^d\frac{\theta}{2}}-1=\frac{d}{8}\theta^2+\mathcal{O}(\theta^3)$, the θ-integral is finite and it follows that

$$
|I_2|\le C\|g\|_{L^1(\mathbb{R}^d)}\|f\|^2_{L^2(\mathbb{R}^d)}.
$$

For the integral I_1, we note that since $g\ge 0$, the matrix in I_1 is positive definite by Bochner's theorem and has the lowest eigenvalue $\hat{g}(0)-|\hat{g}(\eta^-)|$, hence

$$
\begin{aligned}
I_1 &\ge \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(\hat{g}(0)-|\hat{g}(\eta^-)|\right)\left(|\hat{f}(\eta)|^2+|\hat{f}(\eta^+)|^2\right)\mathrm{d}\sigma\,\mathrm{d}\eta\\
&\ge \frac{1}{2}\int_{\mathbb{R}^d}|\hat{f}(\eta)|^2\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(\hat{g}(0)-|\hat{g}(\eta^-)|\right)\mathrm{d}\sigma\,\mathrm{d}\eta.
\end{aligned}
$$

By Lemma 2.13, we have for the power-law case

$$
\begin{aligned}
I_1 &\ge \frac{C_g'}{2}\int_{\{|\eta|\ge 1\}}|\hat{f}(\eta)|^2|\eta|^{2\nu}\,\mathrm{d}\eta \ge \frac{C_g'}{2^{1+\nu}}\int_{\{|\eta|\ge 1\}}|\hat{f}(\eta)|^2|\eta|^{2\nu}\,\mathrm{d}\eta\\
&\ge \frac{C_g'}{2^{1+\nu}}\|f\|^2_{H^1}-\frac{C_g'}{2}\|f\|^2_{L^2},
\end{aligned}
$$

where we used $|\eta|^2 \geq \frac{\langle \eta \rangle^2}{2}$ for $|\eta| \geq 1$. Similarly, we obtain in the Debye-Yukawa case

$$\begin{aligned}
I_1 &\geq \frac{C_g'}{2} \int_{\{|\eta| \geq R\}} |\hat{f}(\eta)|^2 \left(\frac{\log\langle \eta \rangle_\alpha}{\log(\alpha + \mathrm{e})} \right)^{\mu+1} \mathrm{d}\eta \\
&\geq \frac{C_g'}{2(\log(\alpha + \mathrm{e}))^{\mu+1}} \left\| \left(\log\langle D_v \rangle_\alpha\right)^{\frac{\mu+1}{2}} f \right\|_{L^2}^2 - \frac{C_g'}{2} \left(\frac{\log(\alpha + R^2)}{2\log(\alpha + \mathrm{e})} \right)^{\mu+1} \|f\|_{L^2} \\
&\geq \frac{C_g'}{2(\log(\alpha + \mathrm{e}))^{\mu+1}} \left\| \left(\log\langle D_v \rangle_\alpha\right)^{\frac{\mu+1}{2}} f \right\|_{L^2}^2 - \frac{C_g'}{2} \left(\log R \right)^{\mu+1} \|f\|_{L^2}.
\end{aligned}$$

In the last inequality we used the fact that for $R \geq \sqrt{\mathrm{e}}$ the function $\alpha \mapsto \frac{\log(\alpha+R^2)}{2\log(\alpha+\mathrm{e})}$ is decreasing.

Combining the estimates of I_1 and I_2, we arrive at the claimed sub-elliptic estimates for the Boltzmann operator. ∎

CHAPTER 3

Gevrey smoothing for Maxwellian Molecules with Power-law Interaction

3.1 Gevrey regularity and (ultra-)analyticity of weak solutions with L^2 initial data

In this section, we will prove the Gevrey smoothing of weak solutions with initial datum $0 \le f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$ and, *additionally*, $f_0 \in L^2(\mathbb{R}^d)$.

The following is our starting point for the proof of the regularising properties of the homogenous Boltzmann equation:

Proposition 3.1. *Let f be a weak solution of the Cauchy problem (1.2) with initial datum $0 \le f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$, and let $T_0 > 0$. Then for all $t \in (0, T_0]$, $\beta > 0$, $\alpha \in (0,1)$, and $\Lambda > 0$ we have $G_\Lambda f \in \mathscr{C}\left([0,T_0]; L^2(\mathbb{R}^d)\right)$ and*

$$\begin{aligned}\frac{1}{2}\|G_\Lambda(t,D_v)f(t,\cdot)\|^2_{L^2} - \frac{1}{2}\int_0^t \left\langle f(\tau,\cdot), \left(\partial_\tau G^2_\Lambda(\tau,D_v)\right) f(\tau,\cdot)\right\rangle \,\mathrm{d}\tau \\ = \frac{1}{2}\|\mathbb{1}_\Lambda(D_v)f_0\|^2_{L^2} + \int_0^t \left\langle Q(f,f)(\tau,\cdot), G^2_\Lambda(\tau,D_v)f(\tau,\cdot)\right\rangle \,\mathrm{d}\tau.\end{aligned} \tag{3.1}$$

Informally, equation (3.1) follows from using $\varphi(t,\cdot) := G^2_\Lambda(t,D_v)f(t,\cdot)$ in the weak formulation of the homogenous Boltzmann equation.

Recall that $G^2_\Lambda f \in L^\infty([0,T_0]; H^\infty(\mathbb{R}^d))$ for any finite $T_0 > 0$, so it misses the required regularity in time needed to be used as a test function. The proof of Proposition 3.1 is analogous to Morimoto *et al.* [MUXY09], for the sake of completeness and the convenience of the reader, we prove it in Appendix A.

Together with Proposition 3.1 the coercivity estimate Lemma 2.13 implies

Corollary 3.2 (A priori bound for weak solutions)**.** *Let f be a weak solution of the Cauchy problem (1.2) with initial datum $0 \le f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$, and let $T_0 > 0$. Then there exist constants $\widetilde{C}_{f_0}, C_{f_0} > 0$ (depending only on the dimension d, the collision kernel b, $\|f_0\|_{L^1_2}$ and $\|f_0\|_{L\log L}$) such that for all $t \in (0, T_0]$, $\beta > 0$, $\alpha \in (0,1)$, and $\Lambda > 0$ we have*

$$\begin{aligned}\|G_\Lambda f\|_{L^2}^2 \le \|\mathbb{1}_\Lambda(D_v)f_0\|_{L^2}^2 &+ \int_0^t 2\left(-\widetilde{C}_{f_0}\|G_\Lambda f\|_{H^\nu}^2 + C_{f_0}\|G_\Lambda f\|_{L^2}^2\right) \mathrm{d}\tau \\ &+ \int_0^t 2\left|\left\langle Q(f, G_\Lambda f) - G_\Lambda Q(f,f), G_\Lambda f\right\rangle\right| \mathrm{d}\tau \qquad (3.2) \\ &+ \int_0^t 2\beta\|G_\Lambda f\|_{H^\alpha}^2 \,\mathrm{d}\tau.\end{aligned}$$

Proof. We want to apply the coercivity result from Lemma 2.13 to the second integral on the right hand side of Proposition 3.1. Therefore, we write

$$\begin{aligned}\langle Q(f,f), G_\Lambda^2 f\rangle &= \langle G_\Lambda Q(f,f), G_\Lambda f\rangle \\ &= \langle Q(f, G_\Lambda f), G_\Lambda f\rangle + \langle G_\Lambda Q(f,f) - Q(f, G_\Lambda f), G_\Lambda f\rangle \\ &\le -\widetilde{C}_{f_0}\|G_\Lambda f\|_{H^\nu}^2 + \underbrace{C\|f_0\|_{L^1_2}}_{=:C_{f_0}}\|G_\Lambda f\|_{L^2}^2 \\ &\quad + \langle G_\Lambda Q(f,f) - Q(f, G_\Lambda f), G_\Lambda f\rangle.\end{aligned}$$

Moreover,

$$\partial_\tau G_\Lambda^2(\tau,\eta) = 2\beta\langle\eta\rangle^{2\alpha} G_\Lambda^2(t,\eta).$$

Inserting those two results into (3.1), we obtain

$$\begin{aligned}\|G_\Lambda f\|_{L^2}^2 \le \|\mathbb{1}_\Lambda(D_v)f_0\|_{L^2}^2 + 2\beta\int_0^t \|G_\Lambda f(\tau,\cdot)\|_{H^\alpha}^2 \,\mathrm{d}\tau \\ + 2\int_0^t \left(-\widetilde{C}_{f_0}\|G_\Lambda f\|_{H^\nu}^2 + C_{f_0}\|G_\Lambda f\|_{L^2}^2\right) \mathrm{d}\tau \\ + 2\int_0^t \langle G_\Lambda Q(f,f) - Q(f, G_\Lambda f), G_\Lambda f\rangle \,\mathrm{d}\tau.\end{aligned}$$ ∎

The term $\langle G_\Lambda Q(f,f) - Q(f, G_\Lambda f), G_\Lambda f\rangle$ is called *commutation error*.

3.2 Bound on the commutation error

Next, we prove a new bound on the commutation error. An important ingredient is the following elementary observation:

Lemma 3.3 (Strict concavity bound). *Let $\alpha \in (0,1]$ be fixed. The map*

$$0 \le u \mapsto \epsilon(\alpha, u) := (1+u)^\alpha - u^\alpha$$

has the following properties:

(i) If $\alpha \in (0,1)$, then $\epsilon(\alpha,\cdot)$ is strictly *decreasing on $[0,\infty)$ with $\lim_{u\to\infty} \epsilon(\alpha,u) = 0$. In particular, for any $\gamma \ge 1$ and $0 \le \gamma s^- \le s^+$ one has*

$$\epsilon\left(\alpha, \tfrac{s^+}{s^-}\right) \le \epsilon(\alpha,\gamma) \le \epsilon(\alpha,1) = 2^\alpha - 1 < 1. \tag{3.3}$$

Moreover, for all $\alpha \in (0,1)$ and all $u > 0$

$$\epsilon(\alpha,u) \le u^{\alpha-1}.$$

(ii) If $u > 0$, then $\epsilon(\cdot,u)$ is strictly *increasing on $[0,1]$.*

(iii) For all $s^-, s^+ \ge 0$

$$(1+s^- + s^+)^\alpha \le \epsilon\left(\alpha, \tfrac{s^+}{s^-}\right)(1+s^-)^\alpha + (1+s^+)^\alpha.$$

Proof. Since

$$\frac{\partial}{\partial u}\epsilon(\alpha,u) = \alpha\left((1+u)^{\alpha-1} - u^{\alpha-1}\right) < 0 \quad \text{for } \alpha \in (0,1),$$

$\epsilon(\alpha,\cdot)$ is strictly decreasing. Furthermore, for fixed $u > 0$ we have

$$\frac{\partial}{\partial \alpha}\epsilon(\alpha,u) = \log(1+u)\,(1+u)^\alpha - \log u\, u^\alpha > 0,$$

which shows that $\epsilon(\cdot,u)$ is strictly increasing.

For $\alpha \in (0,1)$ and $u \ge 0$ we estimate

$$\epsilon(u,\alpha) = \alpha \int_u^{1+u} r^{\alpha-1}\,\mathrm{d}r \le \alpha u^{\alpha-1} \le u^{\alpha-1}.$$

In particular, $\lim_{u\to\infty}\epsilon(\alpha,u) = 0$. By monotonicity, the chain of inequalities (3.3) follows.

Let $s^-, s^+ \ge 0$. Then

$$\begin{aligned}(1+s^-+s^+)^\alpha &= (s^-)^\alpha\left[\left(1+\tfrac{1+s^+}{s^-}\right)^\alpha - \left(\tfrac{1+s^+}{s^-}\right)^\alpha\right] + (1+s^+)^\alpha \\ &\le \epsilon\left(\alpha, \tfrac{1+s^+}{s^-}\right)(1+s^-)^\alpha + (1+s^+)^\alpha \\ &\le \epsilon\left(\alpha, \tfrac{s^+}{s^-}\right)(1+s^-)^\alpha + (1+s^+)^\alpha\end{aligned}$$

where we made use of the monotonicity of $\epsilon(\alpha,\cdot)$ in the last inequality. ∎

Remark. The proof of Lemma 3.3 is so simple that one might wonder whether it could be of any use. In fact, it is crucial. Its usefulness is hidden in the fact that it enables us to gain a small exponent in the commutator estimates, see Proposition 3.5 and Lemma 3.6 below. Furthermore, $\epsilon(\alpha, \gamma)$ can be made as small as we like if γ can be chosen large enough, which will be important in the proof of Theorem 2.5. ♦

Corollary 3.4. *Let $\widetilde{G}(s) := e^{\beta t(1+s)^\alpha}$ for $s \geq 0$, $\alpha \in (0, 1]$. Then, for all $s^- + s^+ = s$ with $0 \leq s^- \leq s^+$,*

$$|\widetilde{G}(s) - \widetilde{G}(s^+)| \leq 2\alpha\beta t(1 + s^+)^\alpha \big(1 - \tfrac{s^+}{s}\big)\, \widetilde{G}(s^-)^{\epsilon\left(\alpha, \frac{s^+}{s^-}\right)}\, \widetilde{G}(s^+)$$

with $\epsilon(\alpha, u)$ from Lemma 3.3.

Proof. Since $s^+ \leq s$ and $\alpha \in (0, 1]$,

$$\begin{aligned} |\widetilde{G}(s) - \widetilde{G}(s^+)| &\leq \int_{s^+}^{s} \left|\frac{\mathrm{d}}{\mathrm{d}r}\widetilde{G}(r)\right| \mathrm{d}r = \alpha\beta t \int_{s^+}^{s} (1 + r)^{\alpha-1}\widetilde{G}(r)\, \mathrm{d}r \\ &\leq \alpha\beta t(1 + s^+)^{\alpha-1}(s - s^+)\widetilde{G}(s). \end{aligned}$$

In addition, since $s \leq 2s^+$,

$$\frac{s - s^+}{1 + s^+} = \left(1 - \frac{s^+}{s}\right) \frac{s}{1 + s^+} \leq 2\left(1 - \frac{s^+}{s}\right).$$

Moreover, since $s = s^+ + s^-$, the strict concavity Lemma 3.3 gives

$$\widetilde{G}(s) \leq \widetilde{G}(s^-)^{\epsilon\left(\alpha, \frac{s^+}{s^-}\right)}\, \widetilde{G}(s^+),$$

which completes the proof. ■

Proposition 3.5 (Bound on Commutation Error). *Let f be a weak solution of the Cauchy problem (1.2) with initial datum $0 \leq f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$. Recall $\epsilon(\alpha, u) = (1 + u)^\alpha - u^\alpha$. Then for all $t \in (0, T_0]$, $\beta > 0$, $\alpha \in (0, 1)$, and $\Lambda > 0$ we have*

$$\begin{aligned} &\big|\big\langle Q(f, G_\Lambda f) - G_\Lambda Q(f, f), G_\Lambda f\big\rangle\big| \\ &\quad \leq 2\alpha\beta t \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1 - \frac{|\eta^+|^2}{|\eta|^2}\right) G(\eta^-)^{\epsilon(\alpha, |\eta^+|^2/|\eta^-|^2)} |\hat{f}(\eta^-)| \\ &\qquad\qquad \times G_\Lambda(\eta^+)|\hat{f}(\eta^+)|\, G_\Lambda(\eta)|\hat{f}(\eta)|\, \langle\eta^+\rangle^{2\alpha}\, \mathrm{d}\sigma \mathrm{d}\eta, \end{aligned} \tag{3.4}$$

for $d \geq 2$, and

$$\begin{aligned} &\big|\big\langle Q(f, G_\Lambda f) - G_\Lambda Q(f, f), G_\Lambda f\big\rangle\big| \\ &\quad \leq 2\alpha\beta t \int_{\mathbb{R}}\int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} b_1(\theta)\sin^2\theta\, G(\eta^-)^{\epsilon(\alpha, |\eta^+|^2/|\eta^-|^2)} |\hat{f}(\eta^-)| \\ &\qquad\qquad \times G_\Lambda(\eta^+)|\hat{f}(\eta^+)|\, G_\Lambda(\eta)|\hat{f}(\eta)|\, \langle\eta^+\rangle^{2\alpha}\, \mathrm{d}\theta \mathrm{d}\eta, \end{aligned} \tag{3.5}$$

in the one-dimensional case.

Remark. If the weight G was growing *polynomially*, the term $G(\eta^-)$ in the integral (3.4), respectively (3.5), would be replaced by 1. In this case, the "bad terms" which contain η^- can simply be bounded by $\|\hat{f}\|_{L^\infty} \le \|f\|_{L^1} = \|f_0\|_{L^1}$ and the rest can be bounded nicely in terms of $\|G_\Lambda f\|_{L^2}$ and $\|G_\Lambda f\|_{H^\alpha}$, see the discussion in Appendix B.

If the weight G is exponential, the estimate of the terms containing η^- in (3.4), respectively (3.5), is an additional challenge and the methods we devised in order to control this term in the commutation error is probably the most important new contribution of this work. ♦

Proof of Proposition 3.5. We start with $d \ge 2$. By Bobylev's identity, one has

$$\begin{aligned}
&\left|\langle Q(f, G_\Lambda f) - G_\Lambda Q(f, f), G_\Lambda f\rangle\right| = \left|\langle \mathcal{F}\left[Q(f, G_\Lambda f) - G_\Lambda Q(f, f)\right], \mathcal{F}\left[G_\Lambda f\right]\rangle_{L^2}\right| \\
&\quad \le \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right) G_\Lambda(\eta)|\hat{f}(\eta)|\,|\hat{f}(\eta^-)|\,|\hat{f}(\eta^+)||G_\Lambda(\eta^+) - G_\Lambda(\eta)|\,\mathrm{d}\sigma\,\mathrm{d}\eta \\
&\quad = \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right) G_\Lambda(\eta)|\hat{f}(\eta)|\,|\hat{f}(\eta^-)|\,|\hat{f}(\eta^+)||G(\eta^+) - G(\eta)|\,\mathrm{d}\sigma\,\mathrm{d}\eta,
\end{aligned}$$

where the latter equality follows from the fact that G_Λ is supported on the ball $\{|\eta| \le \Lambda\}$ and $|\eta^+| \le |\eta|$.

To estimate $|G(\eta^+) - G(\eta)|$, we use Corollary 3.4 with $s := |\eta|^2$ and, accordingly, $s^\pm = |\eta^\pm|^2$. Notice that

$$|\eta^\pm|^2 = \frac{|\eta|^2}{2}\left(1 \pm \frac{\eta}{|\eta|}\cdot\sigma\right), \quad |\eta|^2 = |\eta^+|^2 + |\eta^-|^2,$$

and, writing $\cos\theta = \frac{\eta\cdot\sigma}{|\eta|}$, we also have

$$|\eta^+|^2 = |\eta|^2\cos^2\tfrac{\theta}{2}, \quad |\eta^-|^2 = |\eta|^2\sin^2\tfrac{\theta}{2}.$$

Since b is supported on angles in $[0, \pi/2]$, one sees $0 \le |\eta^-|^2 \le \frac{1}{2}|\eta|^2$ and $\frac{1}{2}|\eta|^2 \le |\eta^+|^2 \le |\eta|^2$. Therefore, $s^- \le \frac{s}{2} \le s^+ \le s$ and $s = s^+ + s^-$.

It follows that for all $\eta \in \mathbb{R}^d$ with $|\eta| \le \Lambda$, noting that $|\eta^+| \le |\eta| \le \Lambda$,

$$|G(\eta) - G(\eta^+)| \le 2\alpha\beta t\langle\eta^+\rangle^{2\alpha}\left(1 - \tfrac{|\eta^+|^2}{|\eta|^2}\right) G(\eta^-)^{\epsilon(\alpha, |\eta^+|^2/|\eta^-|^2)} G_\Lambda(\eta^+), \qquad (3.6)$$

which finishes the proof in dimension $d \ge 2$.

For the Kac model we remark that the above proof depends only on $|\eta^-| \le |\eta^+| \le |\eta|$ and $|\eta^-|^2 + |\eta^+|^2 = |\eta|^2$, hence $|\eta^-|^2 \le |\eta|^2/2$, and the strict concavity Lemma 3.3 and the Corollary 3.4. Since, by symmetry, we assume that b_1 is supported in $[-\pi/4, \pi/4]$, the same bounds for η^- and η^+ hold in dimension one and the above proof can be literally translated, with obvious changes in notation, to the Kac equation. ■

The bound on the commutation error in Proposition 3.5 is a trilinear expression in the weak solution f. In order to close the a priori bound from Corollary 3.2 in

L^2, one of the terms has to be controlled *uniformly* in η. Seemingly impossible with the growing weights, it is exactly at this place where the gain of the small exponent $\epsilon(\alpha, |\eta^+|^2/|\eta^-|^2) \le \epsilon(\alpha, 1) < 1$ in the $G(\eta^-)$ term in (3.4) and (3.5) allows us to proceed with this strategy. This gain of the small exponent is new and enabled by the strict concavity bound of Lemma 3.3 and its Corollary 3.4 and it is crucial for our inductive approach for controlling the commutation error.

The change of variables is a standard computation used earlier, for instance in [ADVW00, MUXY09]. We repeat it for the convenience of the reader and, more importantly, since some care has to be exercised in view of the strategy of our inductive setup for controlling the commutation error.

Lemma 3.6. *The inequality*

$$\left|\left\langle Q(f, G_\Lambda f) - G_\Lambda Q(f, f), G_\Lambda f\right\rangle\right| \le I_{d,\Lambda} + I^+_{d,\Lambda}$$

holds, where, for $d \ge 2$

$$\begin{aligned} I_{d,\Lambda} = \alpha\beta t \int_{\mathbb{R}^d} \bigg(\int_0^{\frac{\pi}{2}} \int_{\mathbb{S}^{d-2}(\eta)} & \sin^d\theta\, b(\cos\theta)\, G(\eta^-)^{\epsilon\left(\alpha, \cot^2\frac{\theta}{2}\right)}\, |\hat{f}(\eta^-)| \\ & \times \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-|)\, \mathrm{d}\omega\, \mathrm{d}\theta \bigg) |G_\Lambda(\eta)\hat{f}(\eta)|^2 \langle\eta\rangle^{2\alpha}\, \mathrm{d}\eta. \end{aligned} \tag{3.7}$$

Here the vector η^- is expressed as a function of η and σ, that is,

$$\eta^- = \eta^-(\eta, \sigma) = \frac{1}{2}(\eta - |\eta|\sigma) = |\eta| \sin^2(\tfrac{\theta}{2}) \frac{\eta}{|\eta|} - |\eta| \sin(\tfrac{\theta}{2}) \cos(\tfrac{\theta}{2})\, \omega \tag{3.8}$$

and σ is a vector on the unit sphere given by

$$\sigma = \sigma(\theta, \omega) = \cos(\theta) \frac{\eta}{|\eta|} + \sin(\theta)\, \omega \tag{3.9}$$

with polar angle $\theta \in [0, \pi/2]$ with respect to the north pole in the η direction, $\omega \in \mathbb{S}^{d-2}(\eta) := \{\widetilde{\omega} \in \mathbb{R}^d : \widetilde{\omega} \perp \eta, |\widetilde{\omega}| = 1\}$, the $(d-2)$-sphere in $\mathbb{R}^d$ orthogonal to the η direction, and $\mathrm{d}\omega$ the canonical measure on $\mathbb{S}^{d-2}$.

$$\begin{aligned} I^+_{d,\Lambda} = 2^d \alpha\beta t \int_{\mathbb{R}^d} \bigg(\int_0^{\frac{\pi}{4}} \int_{\mathbb{S}^{d-2}(\eta^+)} & \sin^d\vartheta\, b(\cos 2\vartheta)\; G(\eta^-)^{\epsilon(\alpha, \cot^2\vartheta)} |\hat{f}(\eta^-)| \\ & \times \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-|)\, \mathrm{d}\omega\, \mathrm{d}\vartheta \bigg) |G_\Lambda(\eta^+)\hat{f}(\eta^+)|^2 \langle\eta^+\rangle^{2\alpha}\, \mathrm{d}\eta^+ \end{aligned} \tag{3.10}$$

where now the vector η^- is expressed as a function of η^+ and σ, that is,

$$\eta^- = \eta^-(\eta^+, \sigma) = \eta^+ - |\eta^+| \left(\frac{\eta^+ \cdot \sigma}{|\eta^+|} \right)^{-1} \sigma = -|\eta^+| \tan(\vartheta)\, \omega \tag{3.11}$$

where σ is now a vector on the unit sphere with north pole in the η^+ direction given by

$$\sigma = \sigma(\vartheta, \omega) = \cos(\vartheta)\frac{\eta^+}{|\eta^+|} + \sin(\vartheta)\,\omega \tag{3.12}$$

with polar angle $\vartheta \in [0, \pi/4]$ and $\omega \in \mathbb{S}^{d-2}(\eta^+)$, the $(d-2)$-sphere in $\mathbb{R}^d$ orthogonal to the η^+ direction. If $d = 2$ we set $\mathbb{S}^0 := \emptyset$ in this context.

For $d = 1$ we have

$$I_{1,\Lambda} = \alpha\beta t \int_{\mathbb{R}} \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \sin^2\theta\, b_1(\theta)\, G(\eta^-)^{\epsilon\left(\alpha, \cot^2\frac{\theta}{2}\right)} |\hat{f}(\eta^-)|\, \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-|)\, \mathrm{d}\theta$$
$$\times |G_\Lambda(\eta)\hat{f}(\eta)|^2 \langle\eta\rangle^{2\alpha}\, \mathrm{d}\eta,$$
$$I_{1,\Lambda}^+ = \sqrt{2}\alpha\beta t \int_{\mathbb{R}} \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \sin^2\theta\, b_1(\theta)\, G(\eta^-)^{\epsilon\left(\alpha, \cot^2\frac{\theta}{2}\right)} |\hat{f}(\eta^-)|\, \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-|)\, \mathrm{d}\theta$$
$$\times |G_\Lambda(\eta^+)\hat{f}(\eta^+)|^2 \langle\eta^+\rangle^{2\alpha}\, \mathrm{d}\eta^+,$$

where in the first case $\eta^- = \eta^-(\eta, \theta) = \eta \sin\theta$ and in the second case

$$\eta^- = \eta^-(\eta^+, \theta) = \eta^+ \tan\theta$$

and there is no need to distinguish between the θ and ϑ parametrisation.

Remark. In the η, respectively η^+, integrals above η^- and σ are always the same vectors expressed in different parametrisations. We therefore have the relation $\vartheta = \theta/2$, see Figure 3.1 for the geometry of the collision process in Fourier space. ♦

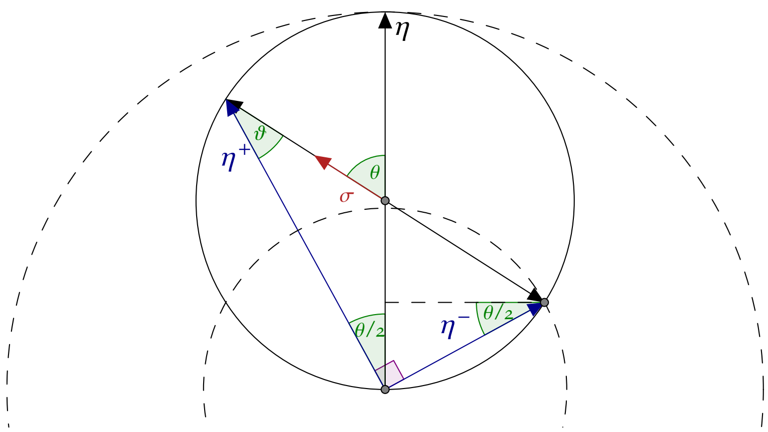

Figure 3.1: Geometry of the collision process in Fourier space.

Remark. From the bounds given in Lemma 3.6 one might already see that, in order to bound the commutation error by some multiple of $\|G_\Lambda f\|^2_{H^\alpha(\mathbb{R}^d)}$, one has to control integrals of the form

$$\sup_{|\eta|\le\Lambda}\int_0^{\frac{\pi}{2}}\int_{\mathbb{S}^{d-2}(\eta)} \sin^d\theta\, b(\cos\theta)\, G^{\epsilon\left(\alpha,\cot^2\frac{\theta}{2}\right)}(\eta^-)\,|\hat f(\eta^-)|\,\mathbb{1}_{\frac{\Lambda}{\sqrt2}}(|\eta^-|)\,\mathrm{d}\omega\,\mathrm{d}\theta,$$

with the parametrisation (3.8) for η^-, and similarly for (3.10) and the corresponding integrals in the one dimensional case. Due to the characteristic function in η^-, this uniform control is not needed on the full ball of radius Λ, but only on a strictly smaller one, giving rise to an *induction-over-length-scales* type of argument. ♦

Proof of Lemma 3.6. Let $d \ge 2$. Using the elementary estimate

$$|G_\Lambda(\eta)\hat f(\eta)|\,|G_\Lambda(\eta^+)\hat f(\eta^+)| \le \frac{1}{2}\left(|G_\Lambda(\eta)\hat f(\eta)|^2 + |G_\Lambda(\eta^+)\hat f(\eta^+)|^2\right)$$

in the bound (3.4) gives

$$\left|\left\langle Q(f, G_\Lambda f) - G_\Lambda Q(f,f), G_\Lambda f\right\rangle\right| \le \widetilde{I}_{d,\Lambda} + \widetilde{I}^+_{d,\Lambda}$$

with

$$\widetilde{I}_{d,\Lambda} = \alpha\beta t\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1-\frac{|\eta^+|^2}{|\eta|^2}\right)G(\eta^-)^{\epsilon(\alpha,|\eta^+|^2/|\eta^-|^2)}|\hat f(\eta^-)| \\ \times \mathbb{1}_{\frac{\Lambda}{\sqrt2}}(|\eta^-|)\,|G_\Lambda(\eta)\hat f(\eta)|^2\,\langle\eta^+\rangle^{2\alpha}\,\mathrm{d}\sigma\mathrm{d}\eta,$$

and

$$\widetilde{I}^+_{d,\Lambda} = \alpha\beta t\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1-\frac{|\eta^+|^2}{|\eta|^2}\right)G(\eta^-)^{\epsilon(\alpha,|\eta^+|^2/|\eta^-|^2)}|\hat f(\eta^-)| \\ \times \mathbb{1}_{\frac{\Lambda}{\sqrt2}}(|\eta^-|)\,|G_\Lambda(\eta^+)\hat f(\eta^+)|^2\,\langle\eta^+\rangle^{2\alpha}\,\mathrm{d}\sigma\mathrm{d}\eta.$$

First we consider $\widetilde{I}_{d,\Lambda}$. Writing σ in a parametrisation where the north pole is in the η direction, one has

$$\sigma = \cos\theta\frac{\eta}{|\eta|} + \sin\theta\,\omega$$

where $\cos\theta = \frac{\eta\cdot\sigma}{|\eta|} \ge 0$ and ω is a unit vector orthogonal to η, that is, $\omega \in \mathbb{S}^{d-2}(\eta)$. Due to the support condition on b one has $\cos\theta \ge 0$, that is, σ is restricted to the northern hemisphere $\theta \in [0, \pi/2]$. In this parametisation one has $\mathrm{d}\sigma = \sin^{d-2}\theta\mathrm{d}\theta\mathrm{d}\omega$. From the definition of $\eta^\pm$ one sees

$$\eta^\pm = \frac{1}{2}(\eta \pm |\eta|\sigma) = \frac{|\eta|}{2}(1\pm\cos\theta)\frac{\eta}{|\eta|} \pm \frac{|\eta|}{2}\sin(\theta)\,\omega$$

so

$$\eta^+ = |\eta|\cos^2(\tfrac{\theta}{2})\frac{\eta}{|\eta|} + |\eta|\sin(\tfrac{\theta}{2})\cos(\tfrac{\theta}{2})\,\omega.$$

In particular,

$$|\eta^+| = |\eta|\cos\frac{\theta}{2}, \quad \text{and} \quad 1 - \frac{|\eta^+|^2}{|\eta|^2} = 1 - \cos^2\frac{\theta}{2} = \sin^2\frac{\theta}{2}.$$

Moreover,

$$\eta^- = |\eta|\sin^2\frac{\theta}{2}\frac{\eta}{|\eta|} - |\eta|\sin\frac{\theta}{2}\cos\frac{\theta}{2}\,\omega, \quad \text{and} \quad |\eta^-| = |\eta|\sin\frac{\theta}{2},$$

so

$$\frac{|\eta^+|^2}{|\eta^-|^2} = \frac{\cos^2\frac{\theta}{2}}{\sin^2\frac{\theta}{2}} = \cot^2\frac{\theta}{2}.$$

After this preparation, using also $\langle\eta^+\rangle^{2\alpha} \le \langle\eta\rangle^{2\alpha}$ and $\sin\frac{\theta}{2} \le \sin\theta$ for $\theta \in [0, \frac{\pi}{2}]$, the inequality $\widetilde{I}_{d,\Lambda} \le I_{d,\Lambda}$ is immediate. The inclusion of the additional factor $\mathbb{1}_\Lambda(|\eta|) = \mathbb{1}_{\sin\frac{\theta}{2}\Lambda}(|\eta^-|) \le \mathbb{1}_{\Lambda/\sqrt{2}}(|\eta^-|)$ seems artificial for the moment, but will be convenient to keep track of the fact that η^- is always restricted to a ball of radius $\frac{\Lambda}{\sqrt{2}}$.

Concerning $\widetilde{I}_{d,\Lambda}$, we want to implement a change of variables from η to η^+. As a function of η and σ, $\eta^+ = \frac{1}{2}(\eta - |\eta|\sigma)$. Thus

$$\left|\frac{\partial\eta^+}{\partial\eta}\right| = \left|\frac{1}{2}\left(\mathbb{1} + \frac{\eta}{|\eta|}\otimes\sigma\right)\right| = \frac{1}{2^d}\left(1 + \frac{\eta}{|\eta|}\cdot\sigma\right) \ge \frac{1}{2^d},$$

since $\eta\cdot\sigma \ge 0$ and the second equality is an application of Sylvester's determinant theorem. Therefore, the Jacobian of the transformation from η to η^+ can be bounded by

$$\left|\frac{\partial\eta}{\partial\eta^+}\right| = \left|\frac{\partial\eta^+}{\partial\eta}\right|^{-1} \le 2^d.$$

In addition,

$$|\eta^+|^2 = \frac{|\eta|^2}{2}\left(1 + \frac{\eta\cdot\sigma}{|\eta|}\right) \quad \text{and} \quad \eta^+\cdot\sigma = \frac{|\eta|}{2}\left(1 + \frac{\eta\cdot\sigma}{|\eta|}\right) = \frac{|\eta^+|^2}{|\eta|},$$

which implies

$$\frac{\eta^+\cdot\sigma}{|\eta^+|} = \frac{|\eta^+|}{|\eta|} \quad \text{and} \quad \frac{\eta\cdot\sigma}{|\eta|} = 2\frac{|\eta^+|^2}{|\eta|^2} - 1 = 2\left(\frac{\eta^+\cdot\sigma}{|\eta^+|}\right)^2 - 1.$$

Moreover, from the definition of $\eta^\pm$ one sees

$$\eta = 2\eta^+ - |\eta|\sigma$$

so

$$\eta^- = \eta^+ - |\eta|\sigma = \eta^+ - |\eta^+| \left(\frac{\eta^+ \cdot \sigma}{|\eta^+|}\right)^{-1} \sigma.$$

Therefore, taking care of the domain of integration,

$$\begin{aligned}\widetilde{I}_d^+ \le 2^d \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} b\left(2\left(\frac{\eta^+ \cdot \sigma}{|\eta^+|}\right)^2 - 1\right)\left(1 - \left(\frac{\eta^+ \cdot \sigma}{|\eta^+|}\right)^2\right) \mathbb{1}_{\frac{\eta^+\cdot\sigma}{|\eta^+|}\Lambda}(|\eta^+|) \\ \times G^{\epsilon\left(\alpha, |\eta^+|^2/|\eta^-|^2\right)}(\eta^-)|\hat{f}(\eta^-)|\, |G_\Lambda(\eta^+)\hat{f}(\eta^+)|^2 \langle\eta^+\rangle^{2\alpha}\, \mathrm{d}\sigma\, \mathrm{d}\eta^+.\end{aligned}$$

Introducing spherical coordinates with north pole in the η^+ direction, one has

$$\sigma = \sigma(\vartheta, \omega) = \cos(\vartheta)\frac{\eta^+}{|\eta^+|} + \sin(\vartheta)\,\omega$$

where now $\cos\vartheta = \frac{\eta^+\cdot\sigma}{|\eta^+|}$. From figure 3.1 one sees $\vartheta = \frac{\theta}{2} \in [0, \pi/4]$. In this parametrisation one has

$$\eta^- = \eta^+ - \frac{|\eta^+|}{\cos\vartheta}\sigma = -|\eta^+|\tan(\vartheta)\,\omega$$

and again $\mathrm{d}\sigma = \sin^{d-2}\vartheta\, \mathrm{d}\vartheta \mathrm{d}\omega$. Thus

$$\begin{aligned}\widetilde{I}_d^+ \le 2^d \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-2}} \int_0^{\frac{\pi}{4}} b\,(\cos 2\vartheta)\, \sin^d\vartheta\, G^{\epsilon\left(\alpha, \cot^2\vartheta\right)}(\eta^-)|\hat{f}(\eta^-)|\, \mathbb{1}_{(\cos\vartheta)\Lambda}(|\eta^+|) \\ \times |G(\eta^+)\hat{f}(\eta^+)|^2 \langle\eta^+\rangle^{2\alpha}\, \mathrm{d}\vartheta\, \mathrm{d}\omega\, \mathrm{d}\eta^+.\end{aligned}$$

Since $|\eta^-| = |\eta^+|\tan\vartheta$ we obtain $\mathbb{1}_{(\cos\vartheta)\Lambda}(|\eta^+|) = \mathbb{1}_{(\sin\vartheta)\Lambda}(|\eta^-|) \le \mathbb{1}_{\Lambda/\sqrt{2}}(|\eta^-|)$, because $\vartheta \in [0, \pi/4]$. Hence $\widetilde{I}_{d,\Lambda}^+ \le I_{d,\Lambda}^+$.

The proof in the $d = 1$ case is completely analogous. ∎

3.3 Extracting pointwise information from local L^2 bounds

Lemma 3.7. *Let $m \ge 2$ and $h \in W^{m,\infty}(\mathbb{R})$ and $q \ge \frac{1}{m}$. Then there exists a constant $L_m < \infty$ depending only on $q, m, \|h\|_{L^\infty(\mathbb{R})}$ and $\|h^{(m)}\|_{L^\infty(\mathbb{R})}$ such that*

$$|h(r)|^q \le L_m \int_{\Omega_r} |h(\xi)|^{q-\frac{1}{m}}\, \mathrm{d}\xi \quad \textit{for all } r \in \mathbb{R},$$

where $\Omega_r = [r, r+2]$ if $r \ge 0$ and $\Omega_r = [r-2, r]$ if $r < 0$.

Looking into the proof of Lemma 3.7, it is clear that its $m = 1$ version also holds, even with a much simpler proof. Before actually going into the proof, we state an important consequence of it, which will enable us to get pointwise decay estimates on a function once suitable L^2 norms are bounded.

For $m \in \mathbb{N}$ define $\|D^m f\|_{L^\infty(\mathbb{R}^d)} := \sup_{\omega\in\mathbb{S}^{d-1}} \|(\omega\cdot\nabla)^m f\|_{L^\infty(\mathbb{R}^d)}$. Notice that this norm is invariant under rotations of the function f.

Corollary 3.8. *Let $H \in \mathscr{C}^m(\mathbb{R}^n)$. Then there exists a constant $L_{m,n} < \infty$ (depending only on $m, n, \|H\|_{L^\infty(\mathbb{R}^n)}$ and, $\|D^m H\|_{L^\infty(\mathbb{R}^n)}$) such that*

$$|H(x)| \leq L_{m,n} \left(\int_{Q_x} |H(\xi)|^2 \, \mathrm{d}\xi \right)^{\frac{m}{2m+n}},$$

where Q_x is a cube in $\mathbb{R}^n$ of side length 2, with x being one of the corners, such that it is oriented away from x in the sense that $x \cdot (\xi - x) \geq 0$ for all $\xi \in Q_x$.

Remark. The constant $L_{m,n}$ in Corollary 3.8 is invariant under rotations of the function H. This will be convenient for its application in Sections 3.5 and 3.6. ♦

Proof. We apply Lemma 3.7 iteratively in each coordinate direction to obtain

$$\begin{aligned}
|H(x_1, x_2, \dots, x_n)|^{2+\frac{n}{m}} &\leq L_m^{(1)} \int_{\Omega_{x_1}} |H(\xi_1, x_2, \dots, x_d)|^{2+\frac{n-1}{m}} \, \mathrm{d}\xi_1 \\
&\leq L_m^{(1)} L_m^{(2)} \int_{\Omega_{x_1}} \int_{\Omega_{x_2}} |H(\xi_1, \xi_2, x_3 \dots, x_d)|^{2+\frac{n-2}{m}} \, \mathrm{d}\xi_1 \, \mathrm{d}\xi_2 \\
&\leq L_m^{(1)} \cdots L_m^{(n)} \int_{\Omega_{x_1}} \cdots \int_{\Omega_{x_d}} |H(\xi_1, \dots, \xi_d)|^2 \, \mathrm{d}\xi_1 \cdots \mathrm{d}\xi_n.
\end{aligned}$$

The constants $L_m^{(i)}$, $i = 1, \dots, n$, only depend on m,

$$\|H(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_n)\|_{L^\infty(\mathbb{R})} \leq \|H\|_{L^\infty(\mathbb{R}^n)}$$

and

$$\|\partial_i^m H(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_n)\|_{L^\infty(\mathbb{R})} \leq \|D^m H\|_{L^\infty(\mathbb{R}^n)}.$$

Setting $L_{m,n} = \prod_{i=1}^n L_m^{(i)}$ yields the stated inequality with $Q_x = \Omega_{x_1} \times \cdots \times \Omega_{x_n}$. ■

Remark. It is worth noticing that the exponent in Corollary 3.8 is decreasing in the dimension and increasing in m. ♦

For the proof of Lemma 3.7 we need the following interpolation result between L^∞ norms of derivatives of a function.

Lemma 3.9 (Kolmogorov-Landau inequality on the unit interval)**.** *Let $m \geq 2$ be an integer. There exists a constant $C_m > 0$ such that for all $w \in W^{m,\infty}([0,1])$,*

$$\|w^{(k)}\|_{L^\infty([0,1])} \leq C_m \left(\frac{\|w\|_{L^\infty([0,1])}}{u^k} + u^{m-k} \|w^{(m)}\|_{L^\infty([0,1])} \right), \quad k = 1, \dots, m-1,$$

for all $0 < u \leq 1$.

Proof. The result dates back to E. Landau and A. N. Kolmogorov who proved it on $\mathbb{R}$ and $\mathbb{R}^+$. A proof of the inequality on a finite interval can be found in the book by R. A. DeVore and G. G. Lorentz [DL93] (pp.37–39), but for the reader's convenience we also give a short proof in Appendix C. ■

For us, the important consequence we are going to make use of is

Corollary 3.10. *Let $C_m > 0$ be the constant from Lemma 3.9. Then for all $w \in W^{m,\infty}([0,1])$,*

$$\|w^{(k)}\|_{L^\infty([0,1])} \leq 2C_m \|w\|^{1-k/m}_{L^\infty([0,1])} \max\left\{\|w\|^{k/m}_{L^\infty([0,1])}, \|w^{(m)}\|^{k/m}_{L^\infty([0,1])}\right\}, \tag{3.13}$$

$k = 1, \ldots, m-1$.

Proof. If $\|w^{(m)}\|_{L^\infty([0,1])} \leq \|w\|_{L^\infty([0,1])}$, we choose $u = 1$ in the bound from Lemma 3.9, which gives

$$\|w^{(k)}\|_{L^\infty([0,1])} \leq 2C_m \|w\|_{L^\infty([0,1])}.$$

In this case, and if $\|w^{(m)}\|_{L^\infty([0,1])} \geq \|w\|_{L^\infty([0,1])}$, we can choose

$$u = \|w\|^{1/m}_{L^\infty([0,1])} \|w^{(m)}\|^{-1/m}_{L^\infty([0,1])} \leq 1$$

to obtain

$$\|w^{(k)}\|_{L^\infty([0,1])} \leq 2C_m \|w\|^{1-k/m}_{L^\infty([0,1])} \|w^{(m)}\|^{k/m}_{L^\infty([0,1])}.$$

Together this proves (3.13). ■

We can now turn to the

Proof of Lemma 3.7. Assume without loss of generality that $r \geq 0$, so that $\Omega_r = [r, r+2]$. By the Sobolev embedding theorem h is continuous and we let r^* be a point in Ω_r where $|h|$ attains its maximum. We can assume that $r^* \in [r, r+1]$ and set $\langle h\rangle_{r^*} := \int_{r^*}^{r^*+1} h(\xi)\,\mathrm{d}\xi$ (otherwise we use $\langle h\rangle_{r^*} := \int_{r^*-1}^{r^*} h(\xi)\,\mathrm{d}\xi$). Then for some $p \geq 1$ we have

$$|h(r^*)|^p - \left|\langle h^p\rangle_{r^*}\right| \leq \int_{r^*}^{r^*+1} |h^p(r^*) - h^p(\xi)|\,\mathrm{d}\xi = \int_0^1 |h^p(r^*) - h^p(r^* + \zeta)|\,\mathrm{d}\zeta.$$

By the fundamental theorem of calculus, for any $\zeta \in [0,1]$, the integrand can be bounded by

$$\begin{aligned}|h^p(r^*) - h^p(r^*+\zeta)| &\leq p \int_0^1 |h(r^*+s\zeta)|^{p-1} |h'(r^*+s\zeta)| \zeta\,\mathrm{d}s \\ &\leq p \sup_{s\in[0,1]} |h'(r^*+s\zeta)| \int_0^1 |h(r^*+s\zeta)|^{p-1} \zeta\,\mathrm{d}s.\end{aligned}$$

We now use that

$$\sup_{s\in[0,1]} |h'(r^*+s\zeta)| = \sup_{x\in[0,\zeta]} |h'(r^*+x)| \le \sup_{x\in[0,1]} |h'(r^*+x)| = \|h'(r^*+\cdot)\|_{L^\infty([0,1])}$$

and apply the Kolmogorov-Landau inequality for the first derivative in its multiplicative form (Corollary 3.10) to the function $[0,1] \ni x \mapsto h(r^*+x) \in W^{m,\infty}([0,1])$ to obtain

$$\begin{aligned}
&\|h'(r^*+\cdot)\|_{L^\infty([0,1])}\\
&\quad\le 2C_m \|h(r^*+\cdot)\|_{L^\infty([0,1])}^{1-1/m} \max\left\{\|h(r^*+\cdot)\|_{L^\infty([0,1])}^{1/m}, \|h^{(m)}(r^*+\cdot)\|_{L^\infty([0,1])}^{1/m}\right\}\\
&\quad\le 2C_m |h(r^*)|^{1-1/m} \max\left\{\|h\|_{L^\infty(\mathbb{R})}^{1/m}, \|h^{(m)}\|_{L^\infty(\mathbb{R})}^{1/m}\right\}.
\end{aligned}$$

It follows that

$$\begin{aligned}
|h(r^*)|^p - \left|\langle h^p\rangle_{r^*}\right| \le 2pC_m |h(r^*)|^{1-1/m} \max\left\{\|h\|_{L^\infty(\mathbb{R})}^{1/m}, \|h^{(m)}\|_{L^\infty(\mathbb{R})}^{1/m}\right\}&\\
\times \int_0^1\int_0^1 |h(r^*+s\zeta)|^{p-1}\zeta \,\mathrm{d}s\,\mathrm{d}\zeta.&
\end{aligned}$$

The latter integral can be further estimated by

$$\begin{aligned}
\int_0^1\int_0^1 |h(r^*+s\zeta)|^{p-1}\zeta\,\mathrm{d}s\,\mathrm{d}\zeta &= \int_0^1\int_0^\zeta |h(r^*+x)|^{p-1}\,\mathrm{d}x\,\mathrm{d}\zeta\\
&\le \int_0^1\int_0^1 |h(r^*+x)|^{p-1}\,\mathrm{d}\zeta\,\mathrm{d}x = \int_0^1 |h(r^*+x)|^{p-1}\,\mathrm{d}x\\
&= \int_{r^*}^{r^*+1} |h(\xi)|^{p-1}\,\mathrm{d}\xi \le \int_{\Omega_r} |h(\xi)|^{p-1}\,\mathrm{d}\xi.
\end{aligned}$$

Using

$$\begin{aligned}
\left|\langle h^p\rangle_{r^*}\right| \le \int_{r^*}^{r^*+1} |h(\xi)|^p\,\mathrm{d}\xi \le \|h\|_{L^\infty(\Omega_r)} \int_{\Omega_r} |h(\xi)|^{p-1}\,\mathrm{d}\xi&\\
\le |h(r^*)|^{1-1/m}\|h\|_{L^\infty(\mathbb{R})}^{1/m} \int_{\Omega_r} |h(\xi)|^{p-1}\,\mathrm{d}\xi&
\end{aligned}$$

we get

$$|h(r^*)|^p \le L_m |h(r^*)|^{1-1/m} \int_{\Omega_r} |h(\xi)|^{p-1}\,\mathrm{d}\xi$$

with $L_m = 2pC_m \max\left\{\|h\|_{L^\infty(\mathbb{R})}^{1/m}, \|h^{(m)}\|_{L^\infty(\mathbb{R})}^{1/m}\right\} + \|h\|_{L^\infty(\mathbb{R})}^{1/m}$, and therefore

$$|h(r^*)|^{p-1+1/m} \le L_m \int_{\Omega_r} |h(\xi)|^{p-1}\,\mathrm{d}\xi.$$

Choosing $q := p-1+1/m \ge 1/m$ then yields

$$|h(r)|^q \le |h(r^*)|^q \le L_m \int_{\Omega_r} |h(\xi)|^{q-1/m}\,\mathrm{d}\xi,$$

which is the claimed inequality. ■

3.4 Gevrey smoothing of weak solutions for L^2 initial data: Part 1

Equipped with Corollary 3.8 we can construct an inductive scheme based upon a uniform bound on $G(\eta^-)^{\epsilon(\alpha,1)}|\hat{f}(\eta^-)|$. As already remarked, this result will depend on the dimension, and will actually deteriorate quickly as dimension increases. Nevertheless it leads to strong regularity properties of weak solutions in the physically relevant cases.

Theorem 3.11. *Assume that the initial datum f_0 satisfies $f_0 \geq 0$, $f_0 \in L\log L(\mathbb{R}^d) \cap L^1_m(\mathbb{R}^d)$ for some $m \geq 2$, and, in addition, $f_0 \in L^2(\mathbb{R}^d)$. Let $0 < \nu < 1$ and assume that the cross-section b satisfies the conditions* **(B1)–(B3)** *with power-law type singularity* (1.18) *for $d \geq 2$. For $d = 1$, assume that b_1 satisfies the conditions* **(K1)–(K3)**.

Let f be a weak solution of the Cauchy problem (1.2) *with initial datum f_0. Set $\alpha_{m,d} := \log\left(\frac{4m+d}{2m+d}\right)/\log 2$. Then, for all $0 < \alpha \leq \min\left\{\alpha_{m,d}, \nu\right\}$ and $T_0 > 0$, there exists $\beta > 0$, such that for all $t \in [0, T_0]$*

$$e^{\beta t \langle D_v \rangle^{2\alpha}} f(t, \cdot) \in L^2(\mathbb{R}^d), \tag{3.14}$$

that is, $f \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d)$ for all $t \in (0, T_0]$.

By decreasing β, if necessary, one even has a uniform bound:

Corollary 3.12. *Let $T_0 > 0$. Under the same conditions as in Theorem 3.11 there exist $\beta > 0$ and $M_1 < \infty$ such that*

$$\sup_{0 \leq t \leq T_0} \sup_{\eta \in \mathbb{R}^d} e^{\beta t \langle \eta \rangle^{2\alpha}} |\hat{f}(t, \eta)| \leq M_1. \tag{3.15}$$

Remark. (i) For strong singularities, the restriction on the Gevrey class originates in the bound on the commutation error, with the best value in $d = 1$ dimension. The aim of Part 2 below will be to recover the two-dimensional result in *any dimension* $d \geq 2$. Under slightly stronger assumptions on the angular cross-section, which still covers all physically relevant cases, we can get the one-dimensional result in any dimension $d \geq 1$, see Part 3.

(ii) In dimensions $d = 1, 2, 3$ and $m = 2$, corresponding to initial data with finite energy, we have $\alpha_{2,d} = \log\left(\frac{8+d}{4+d}\right)/\log 2 \geq \log\left(\frac{11}{7}\right)/\log 2 \simeq 0.652077$. This means that for $\nu = \frac{1}{2}$ the weak solution gets analytic and even ultra-analytic for $\nu > \frac{1}{2}$.

(iii) In the case of physical Maxwellian molecules, where $\nu = \frac{1}{4}$, in three dimensions and with initial datum having finite mass, energy and entropy, we obtain Gevrey $G^2(\mathbb{R}^3)$ regularity.

(iv) Even though the range of α in Theorem 3.11 above deteriorates as the dimension increases, it only fails to cover (ultra-)analyticity results in dimensions $d \geq 6$. Theorems 3.16 and 3.21 below yield results uniformly in the dimension.

♦

We will prove Theorem 3.11 inductively over suitable length scales $\Lambda_N \to \infty$ as $N \to \infty$ in Fourier space. To prepare for this, we fix some $M < \infty$, $0 < T_0 < \infty$ and introduce

Definition 3.13 (Hypothesis $\mathrm{Hyp1}_\Lambda(M)$)**.** Let $M \geq 0$. Then for all $0 \leq t \leq T_0$

$$\sup_{|\zeta|\leq\Lambda} G(t,\zeta)^{\epsilon(\alpha,1)}|\hat{f}(t,\zeta)| \leq M. \tag{3.16}$$

Remark. Recall that $G(t,\zeta) = e^{\beta t \langle\zeta\rangle^\alpha}$, that is, it depends on α, β, and t, and also f is a time dependent function, even though we suppress this dependence in our notation. Thus $\mathrm{Hyp1}_\Lambda(M)$ also depends on the parameters in $G(t,\zeta)$ and on M and T_0, which, for simplicity, we do not emphasise in our notation. We will later fix some $T_0 > 0$ and a suitable large enough M. The main reason why this is possible is that, since $\|\hat{f}\|_{L^\infty} \leq \|f\|_{L^1} = \|f_0\|_{L^1} < \infty$, for any $\Lambda, \beta, T_0 > 0$, the hypothesis $\mathrm{Hyp1}_\Lambda(M)$ is true for large enough M and even any $M > \|f_0\|_{L^1}$ is possible by choosing $\beta > 0$ small enough. ♦

A first step into the inductive proof is the following

Lemma 3.14. *Let $\alpha \leq \nu$ and define $c_{b,d} := |\mathbb{S}^{d-2}| \int_0^{\frac{\pi}{2}} \sin^d\theta\, b(\cos\theta)\,\mathrm{d}\theta$ for $d \geq 3$, $c_{b,2} := \int_0^{\frac{\pi}{2}} \sin^2\theta\, b(\cos\theta)\,\mathrm{d}\theta$, $c_{b,1} := \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \sin^2\theta\, b_1(\theta)\,\mathrm{d}\theta$, which are finite by the integrability assumptions (1.20) and (1.22), and let $\beta \leq \frac{\widetilde{C}_{f_0}}{(1+2^{d-1})\,c_{b,d}\alpha T_0 M+1}$. Then, for any weak solution of the homogenous Boltzmann equation,*

$$\mathrm{Hyp1}_\Lambda(M) \quad\Rightarrow\quad \|G_{\sqrt{2}\Lambda} f\|_{L^2(\mathbb{R}^d)} \leq \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2(\mathbb{R}^d)}\, e^{C_{f_0}T_0} \tag{3.17}$$

for all $0 \leq t \leq T_0$.

Remark. The main point of this lemma is that the right hand side of (3.17) does not depend on M. This is crucial for our analysis and might seem a bit surprising, at first. It is achieved by making β small enough. ♦

Proof. Let $d \geq 2$. Since $\cot^2\frac{\theta}{2} \geq 1$ for $\theta \in [0, \frac{\pi}{2}]$ and $\cot^2\vartheta \geq 1$ for $\vartheta \in [0, \frac{\pi}{4}]$, we can bound $\epsilon(\alpha, \cot^2\frac{\theta}{2})$ and $\epsilon(\alpha, \cot^2\vartheta)$ by $\epsilon(\alpha, 1)$ in the integrals $I_{d,\sqrt{2}\Lambda}$ and $I^+_{d,\sqrt{2}\Lambda}$ from Lemma 4.7.

Assume $\mathrm{Hyp1}_\Lambda(M)$ holds. Then

$$G(t,\zeta)^{\epsilon(\alpha,1)}|\hat{f}(t,\zeta)| \leq M \quad \text{for all} \quad |\zeta| \leq \Lambda.$$

In particular, the terms containing η^- in $I_{d,\sqrt{2}\Lambda}$ and $I^+_{d,\sqrt{2}\Lambda}$ can be bounded by M. Thus, these integrals can now be further estimated by

$$\begin{aligned} I_{d,\sqrt{2}\Lambda} &\leq \alpha\beta t\, M\, |\mathbb{S}^{d-2}| \int_0^{\frac{\pi}{2}} \sin^d\theta\, b(\cos\theta)\,\mathrm{d}\theta \int_{\mathbb{R}^d} |G_{\sqrt{2}\Lambda}(\eta)\hat{f}(\eta)|^2\, \langle\eta\rangle^{2\alpha}\,\mathrm{d}\eta \\ &= \alpha\beta t\, M\, c_{b,d} \|G_{\sqrt{2}\Lambda} f\|^2_{H^\alpha(\mathbb{R}^d)} \end{aligned}$$

and

$$I^+_{d,\sqrt{2}\Lambda} \leq 2^d \alpha\beta t\, M\, |\mathbb{S}^{d-2}| \int_0^{\frac{\pi}{4}} \sin^d\vartheta\, b\,(\cos 2\vartheta)\;\mathrm{d}\vartheta \int_{\mathbb{R}^d} |G(\eta^+)\hat{f}(\eta^+)|^2 \langle\eta^+\rangle^{2\alpha}\,\mathrm{d}\eta^+.$$

In the ϑ integral, we bound $\sin\vartheta \leq \sin(2\vartheta)$ to obtain

$$I^+_{d,\sqrt{2}\Lambda} \leq 2^{d-1}\alpha\beta t\, M\, c_{b,d}\|G_{\sqrt{2}\Lambda} f\|^2_{H^\alpha(\mathbb{R}^d)}.$$

By Lemma 4.7, the commutation error corresponding to the weight $G_{\sqrt{2}\Lambda}$ is thus bounded by

$$\begin{aligned} \left|\left\langle Q(f, G_{\sqrt{2}\Lambda} f) - G_{\sqrt{2}\Lambda} Q(f,f), G_{\sqrt{2}\Lambda} f\right\rangle\right| &\leq I_{d,\sqrt{2}\Lambda} + I^+_{d,\sqrt{2}\Lambda} \\ &\leq (1 + 2^{d-1})\,\alpha\beta t\, M\, c_{b,d}\|G_{\sqrt{2}\Lambda} f\|^2_{H^\alpha(\mathbb{R}^d)}. \end{aligned} \tag{3.18}$$

With Corollary 3.2 we then have

$$\begin{aligned} \|G_{\sqrt{2}\Lambda} f\|^2_{L^2(\mathbb{R}^d)} \leq& \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|^2_{L^2} + \int_0^t 2C_{f_0}\|G_{\sqrt{2}\Lambda} f\|^2_{L^2(\mathbb{R}^d)}\,\mathrm{d}\tau \\ &+ \int_0^t 2\Big(- \widetilde{C}_{f_0}\|G_{\sqrt{2}\Lambda} f\|^2_{H^\nu(\mathbb{R}^d)} \\ &\qquad + \Big((1 + 2^{d-1})\,\alpha\beta t\, M\, c_{b,d} + \beta\Big)\, \|G_{\sqrt{2}\Lambda} f\|^2_{H^\alpha(\mathbb{R}^d)}\Big)\,\mathrm{d}\tau. \end{aligned}$$

Since $\alpha \leq \nu$ and $\beta \leq \frac{\widetilde{C}_{f_0}}{(1+2^{d-1})c_{b,d}\,\alpha T_0 M + 1}$, this implies

$$\|G_{\sqrt{2}\Lambda} f\|^2_{L^2(\mathbb{R}^d)} \leq \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|^2_{L^2(\mathbb{R}^d)} + \int_0^t 2C_{f_0}\|G_{\sqrt{2}\Lambda} f\|^2_{L^2(\mathbb{R}^d)}\,\mathrm{d}\tau$$

and with Gronwall's inequality

$$\|G_{\sqrt{2}\Lambda} f\|^2_{L^2(\mathbb{R}^d)} \leq \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|^2_{L^2(\mathbb{R}^d)}\, \mathrm{e}^{2C_{f_0}T_0} \tag{3.19}$$

follows.

For $d = 1$, we note that, with the obvious change in notation, the above proof literally translates to the Kac equation. ∎

The second ingredient gives a uniform bound in terms of a weighted L^2 norm and some a priori uniform bound on some higher derivative of $\hat{f}$.

Lemma 3.15. *Assume that there exist finite constants A_m and B, such that*

$$\|f(t,\cdot)\|_{L^1_m} \le A_m, \quad \text{and} \quad \|(G_{\sqrt{2}\Lambda} f)(t,\cdot)\|_{L^2(\mathbb{R}^d)} \le B \tag{3.20}$$

for some integer $m \ge 2$ and for all $0 \le t \le T_0$. Set

$$\widetilde{\Lambda} := \frac{1+\sqrt{2}}{2}\Lambda \tag{3.21}$$

and assume furthermore that

$$\Lambda \ge \Lambda_0 := \frac{4\sqrt{d}}{\sqrt{2}-1}. \tag{3.22}$$

Then for all $|\eta| \le \widetilde{\Lambda}$

$$|\hat{f}(t,\eta)| \le K_1\, G(t,\eta)^{-\frac{2m}{2m+d}} \quad \text{for all} \quad 0 \le t \le T_0, \tag{3.23}$$

with a constant K_1 depending only on the dimension d, m, A_m, and B.

Remark. The exponent $\frac{2m}{2m+d}$ in equation (3.23) comes from Corollary 3.8, choosing $n = d$. This is responsible for our definition of $\alpha_{m,d}$, since then $\epsilon\left(\alpha_{m,d}, 1\right) = \frac{2m}{2m+d}$. ♦

Remark. The assumptions of Lemma 3.15 are quite natural: since the Boltzmann equation conserves mass and kinetic energy does not increase, we have the a priori estimate

$$\|f(t,\cdot)\|_{L^1_2(\mathbb{R}^d)} \le \|f_0\|_{L^1_2(\mathbb{R}^d)} =: A_2,$$

and due to the known results on moment propagation for the homogeneous Boltzmann equation in the Maxwellian molecules case[1], we have

$$f_0 \in L^1_m(\mathbb{R}^d) \quad \Longrightarrow \quad f(t,\cdot) \in L^1_m(\mathbb{R}^d) \text{ uniformly in } t \ge 0$$

for any $m > 2$ in addition to assumptions (1.23). ♦

The importance of Lemma 3.15 is that it effectively converts a local L^2 bound on suitable balls into a *pointwise bound* on slightly smaller balls.

Proof of Lemma 3.15. By the Riemann-Lebesgue lemma, the function $\hat{f}$ has continuous and bounded derivatives of order up to m. Since for any multi-index $\alpha \in \mathbb{N}_0^d$ one has

[1] see, for instance, Villani's review [Vil02] pp. 73ff for references.

$\partial^\alpha \hat{f} = (-2\pi \mathrm{i})^{|\alpha|}\, \widehat{v^\alpha f}$, we obtain the bound

$$\begin{aligned}
\|D^m \hat{f}(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} &= \sup_{\omega\in\mathbb{S}^{d-1}} \|(\omega\cdot\nabla)^m \hat{f}(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} \\
&\le \sup_{\omega\in\mathbb{S}^{d-1}} \sup_{\eta\in\mathbb{R}^d} \sum_{|\alpha|=m} \binom{m}{\alpha} |\omega^\alpha|\, |\partial^\alpha \hat{f}(\eta)| \\
&\le (2\pi)^m \sup_{\omega\in\mathbb{S}^{d-1}} \int_{\mathbb{R}^d} \sum_{|\alpha|=m} \binom{m}{\alpha} |\omega^\alpha v^\alpha|\, f(v)\,\mathrm{d}v \\
&\le (2\pi)^m \sup_{\omega\in\mathbb{S}^{d-1}} \int_{\mathbb{R}^d} (\omega\cdot v)^m f(v)\,\mathrm{d}v \\
&\le (2\pi)^m \int_{\mathbb{R}^d} |v|^m f(v)\,\mathrm{d}v \le (2\pi)^m \|f(t,\cdot)\|_{L^1_m(\mathbb{R}^d)} \le (2\pi)^m A_m.
\end{aligned}$$

Of course, also $\|\hat{f}\|_{L^\infty(\mathbb{R}^d)} \le \|f\|_{L^1(\mathbb{R}^d)} \le A_m$.

Let $\eta\in\mathbb{R}^d$ such that $|\eta| \le \widetilde{\Lambda}$. By Corollary 3.8 applied to the function $\hat{f}$, there is a constant $L_{m,d}$ that depends only on d, m, and A_m such that

$$|\hat{f}(\eta)| \le L_{m,d} \left(\int_{Q_\eta} |\hat{f}(\zeta)|^2 \,\mathrm{d}\zeta \right)^{\frac{m}{2m+d}},$$

where Q_η is the cube of side length 2 at η, such that all sides are oriented away from the origin. The definitions of $\widetilde{\Lambda}$ and Λ_0 guarantee by Pythagoras' theorem, that, for $|\eta| \le \widetilde{\Lambda}$, Q_η always stays inside the ball around the origin with radius $\sqrt{2}\Lambda$. Since the orientation of Q_η is such that η is the point closest to the origin and the weight G is radial and increasing, we have

$$\begin{aligned}
|\hat{f}(\eta)| &\le L_{m,d} \left(G(\eta)^{-2} \int_{Q_\eta} G(\zeta)^2 |\hat{f}(\zeta)|^2 \,\mathrm{d}\zeta \right)^{\frac{m}{2m+d}} \\
&\le L_{m,d}\, G(\eta)^{-\frac{2m}{2m+d}} \left(\int_{\{|\eta|\le\sqrt{2}\Lambda\}} G(\zeta)^2 |\hat{f}(\zeta)|^2 \,\mathrm{d}\zeta \right)^{\frac{m}{2m+d}} \\
&\le L_{m,d} B^{\frac{2m}{2m+d}}\, G(\eta)^{-\frac{2m}{2m+d}}.
\end{aligned}$$

Setting $K_1 := L_{m,d} B^{\frac{2m}{2m+d}}$ yields the claimed inequality. ∎

Proof of Theorem 3.11. By Lemma 3.14, 3.15, and Remark 3.4, a suitable choice for A_m, B, and the length scales Λ_N is

$$\begin{aligned}
B &:= \|f_0\|_{L^2(\mathbb{R}^d)} e^{C_{f_0} T_0}, \\
A_m &:= \sup_{t\ge 0} \|f(t,\cdot)\|_{L^1_m(\mathbb{R}^d)} < \infty,
\end{aligned}$$

and

$$\Lambda_N := \frac{\Lambda_{N-1} + \sqrt{2}\Lambda_{N-1}}{2} = \frac{1+\sqrt{2}}{2}\Lambda_{N-1} = \left(\frac{1+\sqrt{2}}{2}\right)^N \Lambda_0$$

with Λ_0 from (3.22).

Furthermore, we set

$$M_1 := \max\{2A_m + 1, K_1\},$$

with the constant K_1 from equation (3.23).

For the start of the induction, we need $\mathrm{Hyp1}_{\Lambda_0}(M_1)$ to be true. Since

$$\sup_{0\le t\le T_0} \sup_{|\eta|\le\Lambda_0} G(\eta)^{\epsilon(\alpha,1)}|\hat{f}(\eta)| \le \mathrm{e}^{\epsilon(\alpha,1)\beta T_0(1+\Lambda_0^2)^\alpha} A_m,$$

and from our choice of M_1, there exists $\beta_0 > 0$ such that $\mathrm{Hyp1}_{\Lambda_0}(M_1)$ is true for all $0 \le \beta \le \beta_0$.

Now, we choose

$$\beta = \min\left(\beta_0, \frac{\tilde{C}_{f_0}}{(1+2^{d-1})c_{b,d}\,\alpha T_0 M_1 + 1}\right).$$

With this choice the conditions of Lemma 3.14 and 3.15 are fulfilled and $\mathrm{Hyp1}_{\Lambda_0}(M_1)$ is true.

For the induction step assume that $\mathrm{Hyp1}_{\Lambda_N}(M_1)$ is true. Then Lemma 3.14 gives

$$\|G_{\sqrt{2}\Lambda_N} f\|_{L^2(\mathbb{R}^d)} \le \|\mathbb{1}_{\sqrt{2}\Lambda_N}(D_v) f_0\|_{L^2(\mathbb{R}^d)}\, \mathrm{e}^{C_{f_0} T_0} \le B.$$

Note that $\epsilon(\alpha,1) \le \frac{2m}{2m+d}$, since $\alpha \le \min\{\alpha_{m,d}, \nu\}$, see Remark 3.4. In addition, $\Lambda_{N+1} = \widetilde{\Lambda}_N$, so Lemma 3.15 shows

$$\sup_{|\eta|\le\Lambda_{N+1}} G(\eta)^{\epsilon(\alpha,1)}|\hat{f}(\eta)| \le K_1 \le M_1,$$

that is, $\mathrm{Hyp1}_{\Lambda_{N+1}}(M_1)$ is true. By induction, it is true for all $N \in \mathbb{N}$. Invoking Lemma 3.14 again, we also have

$$\|G_{\sqrt{2}\Lambda_N} f\|_{L^2(\mathbb{R}^d)} \le B$$

for all $N \in \mathbb{N}$, and passing to the limit $N \to \infty$, we see $\|Gf\|_{L^2(\mathbb{R}^d)} \le B$, which concludes the proof of the theorem. ∎

Proof of Corollary 3.12. The proof of Theorem 3.11 showed that given $T_0 > 0$ there exists $M_1 > 0$ and $\beta > 0$ such that $\mathrm{Hyp1}_{\Lambda_N}(M_1)$ is true for all $N \in \mathbb{N}$. This clearly implies (3.15). ∎

3.5 Gevrey smoothing of weak solutions for L^2 initial data: Part 2

The results of Part 1 are best in one dimension and give the correct smoothing in terms of the Gevrey class for ν not too close to one, more precisely $\nu \leq \alpha_{m,d}$. In order to improve this in higher dimensions $d \geq 2$ and for a larger range of singularities $0 < \nu < 1$, the commutator estimates have to be refined. We have

Theorem 3.16. *Let $d \geq 3$. Assume that the initial datum f_0 satisfies $f_0 \geq 0$, $f_0 \in L\log L(\mathbb{R}^d) \cap L^1_m(\mathbb{R}^d)$ for some $m \geq 2$, and, in addition, $f_0 \in L^2(\mathbb{R}^d)$. Further assume that the cross-section b satisfies conditions* **(B1)–(B3)** *with power-law type singulartity (1.18) for some $0 < \nu < 1$.*

Let f be a weak solution of the Cauchy problem (1.2) with initial datum f_0, then for all $0 < \alpha \leq \min\{\alpha_{m,2}, \nu\}$ and $T_0 > 0$, there exists $\beta > 0$, such that for all $t \in [0, T_0]$

$$e^{\beta t \langle D_v \rangle^{2\alpha}} f(t, \cdot) \in L^2(\mathbb{R}^d), \tag{3.24}$$

that is, $f \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d)$ for all $t \in (0, T_0]$.

In particular, the weak solution is real analytic if $\nu = \frac{1}{2}$ and ultra-analytic if $\nu > \frac{1}{2}$.

The beauty of this theorem is that, in contrast to Theorem 3.11, its result does not deteriorate as dimension increases. We also have a corollary similar to Corollary 3.12, however with a weaker conclusion. Moreover, it is *not* uniform in the time $t \geq 0$ but only holds on finite, but arbitrary, time intervals $[0, T_0]$.

Corollary 3.17. *Under the same assumptions as in Theorem 3.16, for any weak solution f of the Cauchy problem (1.2) and any $0 < T_0 < \infty$ there exists $\widetilde{\beta} > 0$ and $M < \infty$ such that*

$$\sup_{0 \leq t \leq T_0} \sup_{\eta \in \mathbb{R}^d} e^{\widetilde{\beta} t \langle \eta \rangle^{2\alpha}} |\hat{f}(t, \eta)| \leq M. \tag{3.25}$$

The proof of Theorem 3.16 is again based on an induction over length scales in Fourier space. Having a close look at the integrals $I_{d,\Lambda}$ and $I^+_{d,\Lambda}$ from Lemma 3.6 and using that $\epsilon(\alpha, \gamma)$ is decreasing in γ, one sees that it should be enough to bound expressions of the form

$$\int_{\mathbb{S}^{d-2}(\eta)} G(\eta^-)^{\epsilon(\alpha,1)} |\hat{f}(\eta^-)| \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-|) \, d\omega$$

and

$$\int_{\mathbb{S}^{d-2}(\eta^+)} G(\eta^-)^{\epsilon(\alpha,1)} |\hat{f}(\eta^-)| \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-|) \, d\omega$$

uniformly in η and θ, respectively η^+ and ϑ, with the parametrisation (3.8), respectively (3.11), that is, instead of having to use the purely pointwise estimates expressed in the hypothesis $\mathrm{Hyp1}_\Lambda$ from the previous section, one can take advantage of averaging over codimension 2 spheres first. This motivates

Definition 3.18 (Hypothesis $\mathrm{Hyp2}_\Lambda$(M))**.** Let $M \geq 0$ be finite. Then for all $0 \leq t \leq T_0$,

$$\sup_{\zeta\in\mathbb{R}^d\setminus\{0\}} \sup_{(z,\rho)\in A_\Lambda} \int_{\mathbb{S}^{d-2}(\zeta)} G\left(t, z\tfrac{\zeta}{|\zeta|} - \rho\omega\right)^{\epsilon(\alpha,1)} \left|\hat{f}\left(t, z\tfrac{\zeta}{|\zeta|} - \rho\omega\right)\right| \mathrm{d}\omega \leq M, \quad (3.26)$$

where $A_\Lambda = \{(z,\rho) \in \mathbb{R}^2 : 0 \leq z \leq \rho, z^2 + \rho^2 \leq \Lambda^2\}$ and $\mathbb{S}^{d-2}(\zeta) = \{\omega \in \mathbb{R}^d : \omega \perp \zeta, |\omega| = 1\}$.

Again, we have

Lemma 3.19. *Let $\alpha \leq \nu$, define $c_{b,d,2} = \int_0^{\frac{\pi}{2}} \sin^d\theta b(\cos\theta)\,\mathrm{d}\theta$ (which is finite by the integrability assumption (1.20)), and let $\beta \leq \frac{\check{C}_{f_0}}{(1+2^{d-1})c_{b,d,2}\alpha T_0 M+1}$. Then, for any weak solution of the homogenous Boltzmann equation,*

$$\mathrm{Hyp2}_\Lambda(M) \quad \Rightarrow \quad \|G_{\sqrt{2}\Lambda} f\|_{L^2(\mathbb{R}^d)} \leq \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2(\mathbb{R}^d)}\, \mathrm{e}^{C_{f_0}T_0} \qquad (3.27)$$

for all $0 \leq t \leq T_0$.

Proof. Using the monotonicity of $\epsilon(\alpha,\gamma)$ in γ and (3.7) one sees

$$\begin{aligned} I_{d,\sqrt{2}\Lambda} \leq \alpha\beta t \int_{\mathbb{R}^d} \Bigg(\int_0^{\frac{\pi}{2}} \Bigg(\int_{\mathbb{S}^{d-2}(\eta)} & G(\eta^-)^{\epsilon(\alpha,1)}\, |\hat{f}(\eta^-)|\, \mathbb{1}_\Lambda(|\eta^-|)\,\mathrm{d}\omega \Bigg) \\ & \times \sin^d\theta\, b(\cos\theta)\,\mathrm{d}\theta \Bigg) |G_{\sqrt{2}\Lambda}(\eta)\hat{f}(\eta)|^2 \langle\eta\rangle^{2\alpha}\,\mathrm{d}\eta \end{aligned}$$

where $\eta^- = \eta^-(\eta,\theta,\omega)$ is expressed via the parametrisation (3.8). For $\sigma = (\theta,\omega) \in [0,\frac{\pi}{2}] \times \mathbb{S}^{d-2}$, one has $\eta^- = |\eta|\sin^2\frac{\theta}{2}\frac{\eta}{|\eta|} + |\eta|\sin\frac{\theta}{2}\cos\frac{\theta}{2}\,\omega$ and if $|\eta| \leq \sqrt{2}\Lambda$, then $|\eta^-| \leq \Lambda$. Identifying $z = |\eta|\sin^2\frac{\theta}{2}$ and $\rho = |\eta|\sin\frac{\theta}{2}\cos\frac{\theta}{2}$, and the direction of ζ with the direction of η, hypothesis ($\mathrm{Hyp2}_\Lambda$) clearly implies

$$\sup_{|\eta|\leq\sqrt{2}\Lambda} \sup_{\theta\in[0,\pi/2]} \int_{\mathbb{S}^{d-2}(\eta)} G(\eta^-)^{\epsilon(\alpha,1)}\, |\hat{f}(\eta^-)|\, \mathbb{1}_\Lambda(|\eta^-|)\,\mathrm{d}\omega \leq M.$$

It follows that

$$\begin{aligned} I_{d,\sqrt{2}\Lambda} &\leq \alpha\beta t\, M \int_{\mathbb{R}^d} \int_0^{\frac{\pi}{2}} \sin^d\theta b(\cos\theta)\,\mathrm{d}\theta\, |G_{\sqrt{2}\Lambda}(\eta)\hat{f}(\eta)|^2 \langle\eta\rangle\,\mathrm{d}\eta \\ &= \alpha\beta t\, M\, c_{b,d,2} \|G_{\sqrt{2}\Lambda} f\|^2_{H^\alpha(\mathbb{R}^d)}. \end{aligned}$$

Similarly one has

$$\begin{aligned} I^+_{d,\sqrt{2}\Lambda} \leq 2^d\alpha\beta t \int_{\mathbb{R}^d} \Bigg(\int_0^{\frac{\pi}{4}} \Bigg(\int_{\mathbb{S}^{d-2}(\eta^+)} & G(\eta^-)^{\epsilon(\alpha,1)}\, |\hat{f}(\eta^-)|\, \mathbb{1}_\Lambda(|\eta^-|)\,\mathrm{d}\omega \Bigg) \\ & \times \sin^d\vartheta\, b(\cos 2\vartheta)\,\mathrm{d}\vartheta \Bigg) |G_{\sqrt{2}\Lambda}(\eta^+)\hat{f}(\eta^+)|^2 \langle\eta^+\rangle^{2\alpha}\,\mathrm{d}\eta^+, \end{aligned}$$

where $\eta^- = \eta^-(\eta, \vartheta, \omega)$ is expressed via the parametrisation (3.11). The vectors η^- and η^+ are orthogonal and we have $\eta^- = -|\eta^+| \tan\vartheta\, \omega$ for $(\vartheta, \omega) \in [0, \frac{\pi}{4}] \times \mathbb{S}^{d-2}(\eta^+)$.

Setting $z = 0$ and $\rho = |\eta^+| \tan\vartheta$ we have $\rho = |\eta^-| \le \Lambda$ in the ϑ and η^+ integrals above. Thus (Hyp2$_\Lambda$) again implies

$$\sup_{|\eta^+|\le\sqrt{2}\Lambda} \sup_{\vartheta\in[0,\pi/4]} \int_{\mathbb{S}^{d-2}(\eta^+)} G(\eta^-)^{\epsilon(\alpha,1)} |\hat{f}(\eta^-)| \mathbb{1}_\Lambda(|\eta^-|) \,\mathrm{d}\omega \le M.$$

Hence,

$$\begin{aligned} I^+_{d,\sqrt{2}\Lambda} &\le 2^d \alpha\beta t\, M \int_0^{\frac{\pi}{2}} \sin^d \theta b(\cos\theta)\,\mathrm{d}\theta \int_{\mathbb{R}^d} |G_{\sqrt{2}\Lambda}(\eta^+)\hat{f}(\eta^+)|^2 \langle\eta^+\rangle \,\mathrm{d}\eta^+ \\ &\le 2^{d-1}\alpha\beta t\, M\, c_{b,d,2} \|G_{\sqrt{2}\Lambda} f\|^2_{H^\alpha(\mathbb{R}^d)}. \end{aligned}$$

The rest of the proof is the same as in the proof of Lemma 3.14. ■

To close the induction process, we next show

Lemma 3.20. *Let $\beta \le \frac{1}{T_0}$. Assume that there exist finite constants A_m and B, such that*

$$\|f(t,\cdot)\|_{L^1_m} \le A_m, \quad \textit{and} \quad \|(G_{\sqrt{2}\Lambda} f)(t,\cdot)\|_{L^2(\mathbb{R}^d)} \le B \tag{3.28}$$

for some integer $m \ge 2$ and for all $0 \le t \le T_0$.

Set $\widetilde{\Lambda} := \frac{1+\sqrt{2}}{2}\Lambda$ and assume that

$$\Lambda \ge \Lambda_0 := \frac{4\sqrt{2}}{\sqrt{2}-1}. \tag{3.29}$$

Then for all $\zeta \in \mathbb{R}^d \setminus \{0\}$ and $0 \le z \le \rho$ with $\rho^2 + z^2 \le \widetilde{\Lambda}^2$ one has

$$\int_{\mathbb{S}^{d-2}(\zeta)} \left|\hat{f}\left(t, z\tfrac{\zeta}{|\zeta|} + \rho\omega\right)\right| \mathrm{d}\omega \le K_2\, \widetilde{G}(t, z^2+\rho^2)^{-\frac{2m}{2m+2}} \quad \textit{for all } 0 \le t \le T_0$$

with a constant K_2 depending only on d, m, A_m, and B. Recall that $\widetilde{G}(t,s) = \mathrm{e}^{\beta t(1+s)^\alpha}$.

Proof. Fix $0 < t \le T_0$, $\zeta \in \mathbb{R}^d \setminus \{0\}$, and set $F(\rho, z) := \hat{f}(t, z\frac{\zeta}{|\zeta|} + \rho\omega)$, where we drop, for simplicity, the dependence on the time t in our notation for F. Then, since $\|f(t,\cdot)\|_{L^1_m} \le A_m$ one has $\hat{f}(t,\cdot) \in \mathscr{C}^m(\mathbb{R}^d)$ and thus also $F \in \mathscr{C}^m(\mathbb{R}^2)$ with $\|F\|_{L^\infty} \le A_m$, $\|\partial^m_\rho F\|_{L^\infty} \le (2\pi)^m A_m$, and $\|\partial^m_z F\|_{L^\infty} \le (2\pi)^m A_m$ and Corollary 3.8 applied to F yields

$$\left|\hat{f}\left(z\tfrac{\zeta}{|\zeta|} + \rho\omega\right)\right| \le L_{m,2} \left(\int_\rho^{\rho+2} \int_z^{z+2} \left|\hat{f}\left(x\tfrac{\zeta}{|\zeta|} + y\omega\right)\right|^2 \mathrm{d}x\mathrm{d}y\right)^{\frac{m}{2m+2}}. \tag{3.30}$$

where we also dropped the dependence of $\hat{f}$ on the time variable t. Furthermore, we will drop the time dependence of G and $\widetilde{G}$ in the following, that is, $G(\xi)$ and $\widetilde{G}(s)$ will stand for $G(t,\xi)$, respectively $\widetilde{G}(t,s)$.

To recover the L^2 norm of $G_{\sqrt{2}\Lambda} f$ in the right hand side of (3.30) we now need to take care of three things:

(i) Multiply with a suitable power of the radially increasing weight G;

(ii) Integrate over the missing $d-2$ directions, which will be taken care of by integrating over $\mathbb{S}^{d-2}(\zeta)$ and taking into account additional factors to get the d-dimensional Lebesgue measure;

(iii) Ensure that the region of integration $[\rho, \rho+2] \times [z, z+2] \times \mathbb{S}^{d-2}(\zeta)$ stays inside a ball of radius $\sqrt{2}\Lambda$ uniformly in the direction of ζ. This we control by choosing Λ_0 large enough (a simple geometric consideration shows that Λ_0 from the statement of Lemma 3.20 works) and restricting ρ and z by $\rho^2 + z^2 \leq \widetilde{\Lambda}^2$.

Let $z, \rho \geq 0$. In the region of integration in (3.30), the point $\rho\omega + z\frac{\eta}{|\eta|}$ is closest to the origin in $\mathbb{R}^d$, and since the weight G is radially increasing, we get

$$\left|\hat{f}\left(z\tfrac{\zeta}{|\zeta|} + \rho\omega\right)\right| \leq L_{m,2}\widetilde{G}\left(z^2+\rho^2\right)^{-\frac{2m}{2m+2}} \left(\int_{\rho}^{\rho+2}\int_{z}^{z+2} G\left(x\tfrac{\zeta}{|\zeta|} + y\omega\right)^2 \left|\hat{f}\left(x\tfrac{\zeta}{|\zeta|} + y\omega\right)\right|^2 \mathrm{d}x\mathrm{d}y\right)^{\frac{m}{2m+2}}. \tag{3.31}$$

Assume that $z^2 + \rho^2 \leq \widetilde{\Lambda}^2$. Then the integration of inequality (3.31) over $\mathbb{S}^{d-2}(\zeta)$ yields with an application of Jensen's inequality ($t \mapsto t^{\frac{m}{2m+2}}$ is concave!)

$$\int_{\mathbb{S}^{d-2}(\zeta)} \left|\hat{f}\left(z\tfrac{\zeta}{|\zeta|} + \rho\omega\right)\right| \mathrm{d}\omega \leq L_{m,2}|\mathbb{S}^{d-2}|^{\frac{m+2}{2m+2}}\, \widetilde{G}\left(z^2+\rho^2\right)^{-\frac{2m}{2m+2}}$$
$$\times \left(\int_{\mathbb{S}^{d-2}(\zeta)}\int_{\rho}^{\rho+2}\int_{z}^{z+2} G_{\sqrt{2}\Lambda}\left(x\tfrac{\zeta}{|\zeta|} + y\omega\right)^2 \left|\hat{f}\left(x\tfrac{\eta}{|\eta|} + y\omega\right)\right|^2 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\omega\right)^{\frac{m}{2m+2}}.$$

Now assume additionally $0 \leq z \leq \rho$ and $\Lambda_0^2 \leq \rho^2 + z^2 \leq \widetilde{\Lambda}^2$. Since $0 \leq z \leq \rho$ we have $\Lambda_0^2 \leq z^2 + \rho^2 \leq 2\rho^2$ and therefore

$$\int_{\mathbb{S}^{d-2}(\zeta)}\int_{\rho}^{\rho+2}\int_{z}^{z+2} G_{\sqrt{2}\Lambda}\left(x\tfrac{\zeta}{|\zeta|} + y\omega\right)^2 \left|\hat{f}\left(x\tfrac{\zeta}{|\zeta|} + y\omega\right)\right|^2 \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\omega$$
$$\leq 2^{\frac{d-2}{2}}\Lambda_0^{2-d}\int_{\mathbb{S}^{d-2}(\zeta)}\int_{\rho}^{\rho+2}\int_{z}^{z+2} G_{\sqrt{2}\Lambda}\left(x\tfrac{\zeta}{|\zeta|} + y\omega\right)^2 \left|\hat{f}\left(x\tfrac{\zeta}{|\zeta|} + y\omega\right)\right|^2 y^{d-2}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\omega$$
$$\leq 2^{\frac{d-2}{2}}\Lambda_0^{2-d}\|G_{\sqrt{2}\Lambda}f\|^2_{L^2(\mathbb{R}^d)},$$

since $y^{d-2}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}\omega$ is the d-dimensional Lebesgue measure in the cylindrical coordinates $(x, y\omega)$ with $x \in \mathbb{R}$, $y > 0$, $\omega \in \mathbb{S}^{d-2}(\zeta)$ along the cylinder with axis ζ. So with the assumption $\|G_{\sqrt{2}\Lambda}f\|_{L^2(\mathbb{R}^d)} \leq B$ we obtain

$$\int_{\mathbb{S}^{d-2}(\zeta)} \left|\hat{f}\left(t, z\tfrac{\zeta}{|\zeta|} + \rho\omega\right)\right| \mathrm{d}\omega$$
$$\leq L_{m,2}|\mathbb{S}^{d-2}|^{\frac{m+2}{2m+2}}\left(2^{\frac{d-2}{2}}\Lambda_0^{2-d}B^2\right)^{\frac{m}{2m+2}}\widetilde{G}\left(t, z^2+\rho^2\right)^{-\frac{2m}{2m+2}}.$$

In the case $z^2+\rho^2 \le \Lambda_0^2$ we have $\widetilde{G}(t,z^2+\rho^2)^{-1}\, e^{\beta t(1+\Lambda_0^2)^\alpha} \ge 1$ and we can simply bound

$$\begin{aligned}\int_{\mathbb{S}^{d-2}(\zeta)} &\left|\hat{f}\left(t, z\tfrac{\zeta}{|\zeta|}+\rho\omega\right)\right| \mathrm{d}\omega \\ &\le \widetilde{G}\left(t,z^2+\rho^2\right)^{-\frac{2m}{2m+2}} \mathrm{e}^{\frac{2m}{2m+2}\beta t(1+\Lambda_0^2)^\alpha} |\mathbb{S}^{d-2}|\, \|\hat{f}(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} \\ &\le A_m |\mathbb{S}^{d-2}| \mathrm{e}^{1+\Lambda_0^2} \widetilde{G}\left(t,z^2+\rho^2\right)^{-\frac{2m}{2m+2}}\end{aligned}$$

since $\beta \le 1/T_0$, by assumption. So choosing

$$K_2 := \max\left(L_{m,2}|\mathbb{S}^{d-2}|^{\frac{m+2}{2m+2}} \left(2^{\frac{d-2}{2}}\Lambda_0^{2-d}B^2\right)^{\frac{m}{2m+2}}, A_m|\mathbb{S}^{d-2}|e^{1+\Lambda_0^2}\right)$$

finishes the proof of the lemma. ∎

Now we have all the ingredients for the inductive

Proof of Theorem 3.16. By Lemmata 3.19 and 3.20 a suitable choice for A_m and B is

$$\begin{aligned} B &:= \|f_0\|_{L^2(\mathbb{R}^d)} \mathrm{e}^{C_{f_0}T_0}, \\ A_m &:= \sup_{t\ge 0} \|f(t,\cdot)\|_{L^1_m(\mathbb{R}^d)} < \infty. \end{aligned}$$

Note that the finiteness of A_m is guaranteed since $f_0 \in L^1_m(\mathbb{R}^d)$, see Remark 3.4. We further choose the length scales Λ_N to be

$$\Lambda_N := \frac{\Lambda_{N-1}+\sqrt{2}\Lambda_{N-1}}{2} = \frac{1+\sqrt{2}}{2}\Lambda_{N-1} = \left(\frac{1+\sqrt{2}}{2}\right)^N \Lambda_0$$

with Λ_0 now from (3.29), and we set

$$M_2 := \max\left\{2|\mathbb{S}^{d-2}|A_m+1, K_2\right\}$$

with the constant K_2 from Lemma 3.20.

For the start of the induction, we need $\mathrm{Hyp2}_{\Lambda_0}(M_2)$ to be true. Since

$$\begin{aligned}\sup_{0\le t\le T_0}\ \sup_{\zeta\in\mathbb{R}^d\setminus\{0\}}\ \sup_{(z,\rho)\in A_{\Lambda_0}} \int_{\mathbb{S}^{d-2}(\zeta)} G\left(t, z\tfrac{\zeta}{|\zeta|}-\rho\omega\right)^{\epsilon(\alpha,1)} \left|\hat{f}\left(t, z\tfrac{\zeta}{|\zeta|}-\rho\omega\right)\right| \mathrm{d}\omega \\ \le |\mathbb{S}^{d-2}| \mathrm{e}^{\beta T_0(1+\Lambda_0^2)} A_m\end{aligned}$$

and from our choice of M_2 there exists $\beta_0 > 0$ such that $\mathrm{Hyp2}_{\Lambda_0}(M_2)$ is true for all $0 \le \beta \le \beta_0$.

Now, we choose

$$\beta = \min\left(\beta_0, T_0^{-1}, \frac{\tilde{C}_{f_0}}{(1+2^{d-1})c_{b,d,2}\,\alpha T_0 M_2 + 1}\right).$$

With this choice the conditions of Lemma 3.19 and 3.20 are fulfilled and $\mathrm{Hyp2}_{\Lambda_0}(M_2)$ is true.

For the induction step assume that $\mathrm{Hyp2}_{\Lambda_N}(M_2)$ is true. Then Lemma 3.19 gives

$$\|G_{\sqrt{2}\Lambda_N} f\|_{L^2(\mathbb{R}^d)} \leq \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2(\mathbb{R}^d)} \, \mathrm{e}^{C_{f_0} T_0} = B$$

and then, since $\epsilon(\alpha, 1) \leq \frac{2m}{2m+2}$ by our choice of α, and $\Lambda_{N+1} = \widetilde{\Lambda}_N$, Lemma 3.20 shows that $\mathrm{Hyp2}_{\Lambda_{N+1}}(M_2)$ is true, so by induction, it is true for all $N \in \mathbb{N}$. Invoking Lemma 3.19 again, we also have

$$\|G_{\sqrt{2}\Lambda_N} f\|_{L^2(\mathbb{R}^d)} \leq B$$

for all $N \in \mathbb{N}$ and letting $N \to \infty$, we see $\|Gf\|_{L^2(\mathbb{R}^d)} \leq B$, which concludes the proof of Theorem 3.16. ∎

Proof of Corollary 3.17. Theorem 3.16 shows that $Gf \in L^2(\mathbb{R}^d)$ for all $0 \leq t \leq T_0$. Applying Corollary 3.8 with $n = d$ to $\hat{f}$ yields

$$\begin{aligned} |\hat{f}(\eta)| &\leq L_{m,d} G(\eta)^{-\frac{2m}{2m+d}} \left(\int_{Q_\eta} G(\zeta)^2 |\hat{f}(\zeta)|^2 \, \mathrm{d}\zeta \right)^{\frac{m}{2m+d}} \\ &\leq L_{m,d} \|Gf\|_{L^2(\mathbb{R}^d)}^{\frac{2m}{2m+d}} G(\eta)^{-\frac{2m}{2m+d}}, \end{aligned}$$

where we also used that the Fourier multiplier is radially increasing. This proves the uniform bound (3.25) with $\widetilde{\beta} = \beta \frac{2m}{2m+d}$. ∎

3.6 Gevrey smoothing of weak solutions for L^2 initial data: Part 3

Under the slightly stronger assumption that the angular collision cross-section b is bounded away from the singularity, we can state our theorem about Gevrey regularisation in its strongest form.

Theorem 3.21. *Assume that the initial datum f_0 satisfies $f_0 \geq 0$, $f_0 \in L \log L(\mathbb{R}^d) \cap L^1_m(\mathbb{R}^d)$ for some $m \geq 2$, and, in addition, $f_0 \in L^2(\mathbb{R}^d)$. Further assume that the cross-section b in dimensions $d \geq 2$ satisfies the* singularity condition (1.18) *for some $0 < \nu < 1$ and the* boundedness condition (2.5).

Let f be a weak solution of the Cauchy problem (1.2) with initial datum f_0, then for all $0 < \alpha \leq \min\{\alpha^{m,1}, \nu\}$ and all $T_0 > 0$, there exists $\beta > 0$, such that for all $t \in [0, T_0]$

$$\mathrm{e}^{\beta t \langle D_v \rangle^{2\alpha}} f(t, \cdot) \in L^2(\mathbb{R}^d), \tag{3.32}$$

that is, $f \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d)$ for all $t \in (0, T_0]$.

In particular, the weak solution is real analytic if $\nu = \frac{1}{2}$ and ultra-analytic if $\nu > \frac{1}{2}$.

Remark. Thus, under slightly stronger assumptions on b than in Theorem 3.11, which we stress are nevertheless fulfilled in any physically reasonable cases, we can prove the same regularity in *any dimension* as can be obtained for radially symmetric solutions of the homogenous Boltzmann equation. ♦

Corollary 3.22. *Under the same assumptions as in Theorem 3.21, for any weak solution f of the Cauchy problem (1.2) and any $0 < T_0 < \infty$ there exist $\beta > 0$ and $M < \infty$ such that*

$$\sup_{0\le t\le T_0}\sup_{\eta\in\mathbb{R}^d} \mathrm{e}^{\beta t\langle\eta\rangle^{2\alpha}}\,|\hat{f}(t,\eta)| \le M. \tag{3.33}$$

Proof. Given Theorem 3.21, the proof of Corollary 3.22 is the same as the proof of Corollary 3.17. ■

The proof of Theorem 3.21 shows the delicate interplay between the angular singularity of the collision kernel, the strict concavity of the Gevrey weights, and the use of averages of the weak solution in Fourier space, together with our inductive procedure, which has proved to be successful in Theorems 3.11 and 3.16. Again, the main work is to bound the expressions $I_{d,\Lambda}$ and $I^+_{d,\Lambda}$ from Lemma 3.6. Before we start the proof of Theorem 3.21, we start with some preparations. It is clear that we only have to prove Theorem 3.21 in dimension $d \ge 2$ and for singularities $\nu > \alpha_{2,m}$, since otherwise the result is already contained in Theorems 3.11 and 3.16.

Looking at the integral $I_{d,\Lambda}$ from Lemma 3.6, one has

$$\begin{aligned} I_{d,\Lambda} = \alpha\beta t \int_{\mathbb{R}^d}\Bigg(\int_0^{\frac{\pi}{2}}\int_{\mathbb{S}^{d-2}(\eta)} &\sin^d\theta b(\cos\theta)\, G(\eta^-)^{\epsilon\left(\alpha,\cot^2\frac{\theta}{2}\right)}\,|\hat{f}(\eta^-)| \\ &\times \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-|)\,\mathrm{d}\omega\,\mathrm{d}\theta\Bigg)\,|G_\Lambda(\eta)\hat{f}(\eta)|^2\,\langle\eta\rangle^{2\alpha}\,\mathrm{d}\eta. \end{aligned}$$

where we use the parametrization (3.8) for $\eta^- = \eta^-(\eta,\theta,\omega)$. Splitting the θ integral above at a point $\theta_0 \in (0,\frac{\pi}{2})$ and using the monotonicity of the cotangent on $[0,\frac{\pi}{2}]$ and of $\epsilon(\alpha,\gamma)$ in γ one sees

$$I_{d,\Lambda} \le I_{d,\Lambda,1} + I_{d,\Lambda,2}$$

with

$$\begin{aligned} I_{d,\Lambda,1} :=& \alpha\beta T_0\,\|G_\Lambda f\|^2_{H^\alpha(\mathbb{R}^d)}\int_0^{\theta_0}\sin^d\theta\, b(\cos\theta)\,\mathrm{d}\theta \\ &\times\Bigg(\sup_{0<\theta\le\frac{\pi}{2}}\sup_{0<|\eta|\le\Lambda}\int_{\mathbb{S}^{d-2}(\eta)} G(\eta^-(\eta,\theta,\omega))^{\epsilon\left(\alpha,\cot^2\frac{\theta_0}{2}\right)} \\ &\qquad\times|\hat{f}(\eta^-(\eta,\theta,\omega))|\mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-(\eta,\theta,\omega)|)\,\mathrm{d}\omega\Bigg) \end{aligned} \tag{3.34}$$

and

$$\begin{aligned} I_{d,\Lambda,2} :=& C_{\theta_0}\alpha\beta T_0 \|G_\Lambda f\|^2_{H^\alpha(\mathbb{R}^d)} \\ &\times \Bigg(\sup_{0<|\eta|\le\Lambda} \int_{\theta_0}^{\frac{\pi}{2}} \int_{\mathbb{S}^{d-2}(\eta)} G(\eta^-(\eta,\theta,\omega))^{\epsilon(\alpha,1)} |\hat{f}(\eta^-(\eta,\theta,\omega))| \\ &\qquad\qquad \times \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-(\eta,\theta,\omega)|)\,\mathrm{d}\omega\,\mathrm{d}\theta\Bigg), \end{aligned} \tag{3.35}$$

where C_{θ_0} is an upper bound for $b(\cos\theta)$ on $[\theta_0, \frac{\pi}{2}]$. Now we choose $\theta_0 > 0$ so small that

$$\epsilon(\alpha, \cot^2\frac{\theta_0}{2}) \le \epsilon(\alpha_{2,m}, 1) = \frac{2m}{2m+2}$$

and note that from Corollary 3.17, since $\nu > \alpha_{2,m}$, there exists a finite M_2 such that

$$\begin{aligned} \sup_{0<\theta\le\frac{\pi}{2}} \sup_{0<|\eta|\le\Lambda} \int_{\mathbb{S}^{d-2}(\eta)} & G(\eta^-(\eta,\theta,\omega))^{\epsilon(\alpha_{2,m},1)} |\hat{f}(\eta^-(\eta,\theta,\omega))| \\ & \times \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-(\eta,\theta,\omega)|)\,\mathrm{d}\omega \le M_2 < \infty. \end{aligned}$$

So from (3.34) we get the bound

$$I_{d,\Lambda,1} \le \alpha\beta T_0 M_2 c_{b,d,2} \|G_\Lambda f\|^2_{H^\alpha(\mathbb{R}^d)} \tag{3.36}$$

where the finiteness of $c_{b,d,2}$ follows from the singularity condition and the boundedness of $b(\cos\theta)$ away from $\theta = 0$.

For the integral $I^+_{d,\Lambda}$ from Lemma 3.6, a completely analogous reasoning as above shows for small enough ϑ_0 such that $\epsilon(\alpha\cot\vartheta) \le \epsilon(\alpha_{2,m}, 1)$ we also have

$$I^+_{d,\Lambda} \le I^+_{d,\Lambda,1} + I^+_{d,\Lambda,2}$$

with

$$I^+_{d,\Lambda,1} \le 2^{d-1}\alpha\beta T_0 M_2 c_{b,d,2} \|G_\Lambda f\|^2_{H^\alpha(\mathbb{R}^d)} \tag{3.37}$$

and

$$\begin{aligned} I^+_{d,\Lambda,2} :=& 2^d C_{\vartheta_0}\alpha\beta T_0 \|G_\Lambda f\|^2_{H^\alpha(\mathbb{R}^d)} \\ &\times \Bigg(\sup_{0<|\eta^+|\le\Lambda} \int_{\vartheta_0}^{\frac{\pi}{4}} \int_{\mathbb{S}^{d-2}(\eta^+)} G(\eta^-(\eta^+,\vartheta,\omega))^{\epsilon(\alpha,1)} \\ &\qquad \times |\hat{f}(\eta^-(\eta^+,\vartheta,\omega))| \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-(\eta^+,\vartheta,\omega)|)\,\mathrm{d}\omega\,\mathrm{d}\vartheta\Bigg) \end{aligned} \tag{3.38}$$

where we use the parametrisation (3.11) for $\eta^- = \eta^-(\eta^+,\vartheta,\omega)$ and where C_{ϑ_0} is an upper bound for $b(\cos(2\vartheta))$ on $[\vartheta_0, \frac{\pi}{4}]$.

Recall that we always assume $\alpha \le \alpha_{1,m}$, so $\epsilon(\alpha, 1) \le \epsilon(\alpha_{1,m}, 1) = \frac{2m}{2m+1}$. Thus we see that in order to set up our inductive procedure for controlling $I_{d\Lambda}$ and $I^+_{d,\Lambda}$ it is natural to introduce

Definition 3.23 (Hypothesis $\mathrm{Hyp3}_\Lambda(M)$)**.** Let $M \geq 0$ be finite, $0 < \theta_0, \vartheta_0 < \frac{\pi}{4}$, $T_0 > 0$, and $m \geq 2$ an integer. Then for all $0 \leq t \leq T_0$ one has

$$\sup_{|\eta| \leq \sqrt{2}\Lambda} \int_{\theta_0}^{\frac{\pi}{2}} \int_{\mathbb{S}^{d-2}(\eta)} G\left(t, \eta^-(\eta,\theta,\omega)\right)^{\frac{2m}{2m+1}} \left|\hat{f}\left(\eta^-(\eta,\theta,\omega)\right)\right| \\ \times \mathbb{1}_\Lambda(|\eta^-(\eta,\theta,\omega)|)\, \mathrm{d}\omega\, \mathrm{d}\theta \leq M, \tag{3.39}$$

where we use the parametrisation given in (3.8) for η^-, and

$$\sup_{|\eta^+| \leq \sqrt{2}\Lambda} \int_{\vartheta_0}^{\frac{\pi}{4}} \int_{\mathbb{S}^{d-2}(\eta^+)} G\left(t, \eta^-(\eta^+,\vartheta,\omega)\right)^{\frac{2m}{2m+1}} \left|\hat{f}\left(\eta^-(\eta^+,\vartheta,\omega)\right)\right| \\ \times \mathbb{1}_\Lambda(|\eta^-(\eta^+,\vartheta,\omega)|)\, \mathrm{d}\omega\, \mathrm{d}\vartheta \leq M, \tag{3.40}$$

where we use the parametrisation given in (3.11) for η^-.

For the induction proof of Theorem 3.21, we again start with

Lemma 3.24. *Let $M \geq 0$, $T_0 > 0$, $m \geq 2$ an integer, $\alpha_{m,2} < \nu < 1$, $0 < \alpha \leq \nu$ and recall $c_{b,d,2} = \int_0^{\frac{\pi}{2}} \sin^d \theta b(\cos\theta)\, \mathrm{d}\theta$ (which is finite by the singularity assumption (1.20) and the boundedness assumption (2.5)). Let M_2 be from Corollary 3.17 and $\beta \leq \frac{\tilde{C}_{f_0}}{\alpha T_0[(1+2^{d-1})c_{b,d,2}M_2+(C_{\theta_0}+2^d C_{\vartheta_0})M]+1}$.*

Then for any weak solution of the homogenous Boltzmann equation,

$$\mathrm{Hyp3}_\Lambda(M) \quad \Rightarrow \quad \|G_{\sqrt{2}\Lambda} f\|_{L^2(\mathbb{R}^d)} \leq \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2(\mathbb{R}^d)}\, e^{C_{f_0} T_0} \tag{3.41}$$

for all $0 \leq t \leq T_0$.

Proof. Given Lemma 3.6 and the above discussion with the bounds in (3.36), (3.37), and using the hypothesis $\mathrm{Hyp3}_\Lambda$ for the terms in (3.35) and (3.38), one sees that the commutation error on the level $\sqrt{2}\Lambda$ is bounded by

$$\begin{aligned} &\left|\left\langle Q(f, G_{\sqrt{2}\Lambda} f) - G_{\sqrt{2}\Lambda} Q(f,f), G_{\sqrt{2}\Lambda} f\right\rangle\right| \leq I_{d,\sqrt{2}\Lambda} + I^+_{d,\sqrt{2}\Lambda} \\ &\leq (1+2^{d-1})\alpha\beta T_0 M_2 c_{b,d,2} \|G_\Lambda f\|^2_{H^\alpha(\mathbb{R}^d)} + (C_{\theta_0} + 2^d C_{\vartheta_0})\alpha\beta T_0 M \|G_\Lambda f\|^2_{H^\alpha(\mathbb{R}^d)}. \end{aligned}$$

Given this bound on the commutation error, the rest of the proof is the same as in the proof of Lemma 3.14. ∎

To close the induction step we also need a suitable version of Lemma 3.20, but before we prove this we need a preparatory Lemma.

Lemma 3.25. *Let $H : \mathbb{R}^d \to \mathbb{R}_+$ be a locally integrable function and let $\eta, \eta_+ \in \mathbb{R}^d$ with $|\eta|, |\eta^+| \geq \Lambda_0 > 0$, $0 < \theta_0 \leq \frac{\pi}{2}$, and $0 < \vartheta_0 \leq \frac{\pi}{8}$. Then with the parametrisation*

$\eta^- = \eta^-(\eta, \theta, \omega)$ *given in* (3.8) *one has*

$$\int_{\theta_0}^{\frac{\pi}{2}} \int_0^2 H\left(\eta^-(\eta,\theta,\omega) + z\tfrac{\eta}{|\eta|}\right) \mathrm{d}z\, \mathrm{d}\theta \\ \leq \frac{2}{\Lambda_0 \cos\theta_0} \int_{\Lambda_0 \sin^2 \frac{\theta_0}{2}}^{\frac{|\eta|}{2}+2} \int_{\Lambda_0 \sin\theta_0}^{\frac{|\eta|}{2}} H\left(x\tfrac{\eta}{|\eta|} - y\omega\right) \mathrm{d}y\, \mathrm{d}x$$

for any unit vector ω orthogonal to η.

Moreover, with the parametrisation $\eta^- = \eta^-(\eta^+, \theta, \omega)$ given in (3.11) *one has, for any* $\widetilde{\Lambda} \geq \frac{1+\sqrt{2}}{2}\Lambda_0$,

$$\int_{\vartheta_0}^{\frac{\pi}{4}} \int_0^2 H\left(\eta^-(\eta^+,\vartheta,\omega) + z\tfrac{\eta}{|\eta|}\right) \mathbb{1}_{\frac{\widetilde{\Lambda}}{\sqrt{2}}}(|\eta^-(\eta^+,\vartheta,\omega)|)\, \mathrm{d}z\, \mathrm{d}\vartheta \\ \leq \frac{1}{2\Lambda_0} \int_0^2 \int_{\Lambda_0 \tan\vartheta_0}^{\frac{\widetilde{\Lambda}}{\sqrt{2}}} H\left(x\tfrac{\eta}{|\eta|} - y\omega\right) \mathrm{d}y\, \mathrm{d}x$$

Remark. The restriction $\vartheta_0 \leq \frac{\pi}{8}$ is only for convenience, to ensure that $\Lambda_0 \tan\vartheta_0 \leq \frac{\widetilde{\Lambda}}{\sqrt{2}}$. ♦

Proof. Fix η as required and ω orthogonal to it. We want to have a map $\Phi_1 : (\theta, z) \mapsto \Phi_1(\theta, z) = (x, y)$ such that

$$\eta^-(\eta,\theta,\omega) + z\tfrac{\eta}{|\eta|} = x\tfrac{\eta}{|\eta|} - y\omega.$$

From the parametrisation (3.8) we read off

$$x = |\eta| \sin^2\frac{\theta}{2} + z \quad \text{and } y = \frac{|\eta|}{2}\sin\theta$$

and we can compute the Jacobian going from the (θ, z) variables to (x, y) as

$$\left|\frac{\partial(x,y)}{\partial(\theta,z)}\right| = |\det D\Phi_1| = \frac{|\eta|}{2}\cos\theta \geq \frac{|\eta|}{2}\cos\theta_0.$$

Since $|\eta| \geq \Lambda_0$, $\theta \in [\theta_0, \frac{\pi}{2}]$, and $0 \leq z \leq 2$, we have $\Lambda_0 \sin^2\frac{\theta_0}{2} \leq x \leq |\eta|\sin^2\frac{\pi}{4} = \frac{\eta}{2}$ and $\frac{\Lambda_0}{2}\sin\theta_0 \leq y \leq \frac{\eta}{2}$. So doing a change of variables $(\theta, z) = \Phi_1^{-1}(x, y)$ in the integral we can bound

$$\int_{\theta_0}^{\frac{\pi}{2}} \int_0^2 H\left(\eta^-(\eta,\theta,\omega) + z\tfrac{\eta}{|\eta|}\right) \mathrm{d}z\, \mathrm{d}\theta \\ \leq \frac{2}{\Lambda_0 \cos\theta_0} \int_{\Lambda_0 \sin^2 \frac{\theta_0}{2}}^{\frac{|\eta|}{2}+2} \int_{\Lambda_0 \sin\theta_0}^{\frac{|\eta|}{2}} H\left(x\tfrac{\eta}{|\eta|} + y\omega\right) \mathrm{d}y\, \mathrm{d}x$$

since the map Φ_1 is a nice diffeomorphism.

For the second bound the calculation is, in fact, a bit easier, one just has to take care that $|\eta^-|$ cannot be too large, which is taken into account by the factor $\mathbb{1}_\Lambda(|\eta^-|)$. We now want a map $\Phi_2 : (\theta, z) \mapsto \Phi_2(\theta, z) = (x, y)$ such that

$$\eta^-(\eta^+, \vartheta, \omega) + z\tfrac{\eta^+}{|\eta^+|} = x\tfrac{\eta^+}{|\eta^+|} - y\omega.$$

From the parametrisation (3.8) we read off

$$x = z \quad \text{and } y = |\eta^-| = |\eta^+| \tan\vartheta$$

and the Jacobian going from the (ϑ, z) variables to (x, y) is simply

$$\left|\frac{\partial(x,y)}{\partial(\vartheta,z)}\right| = |\det D\Phi_2| = 2|\eta^+| \geq 2\Lambda_0.$$

We certainly have $0 \leq x \leq 2$ and also $\Lambda_0 \tan\vartheta_0 \leq y$. Since $y = |\eta^-|$, we also have the restriction $y \leq \Lambda$. So the proof of the second inequality follows similar to the proof of first one. ■

Finally, we can state and prove the second step in our inductive procedure.

Lemma 3.26. *Let $\beta \leq \frac{1}{T_0}$. Assume that there exist finite constants A_m and B, such that*

$$\|f(t,\cdot)\|_{L^1_m} \leq A_m, \quad \text{and} \quad \|(G_{\sqrt{2}\Lambda} f)(t,\cdot)\|_{L^2(\mathbb{R}^d)} \leq B \tag{3.42}$$

for some integer $m \geq 2$ and for all $0 \leq t \leq T_0$.

Set $\widetilde{\Lambda} := \frac{1+\sqrt{2}}{2}\Lambda$ and assume that

$$\Lambda \geq \Lambda_0 := 3. \tag{3.43}$$

Then there exists a finite K_3, depending only on d, m, A_m, and B such that $\mathrm{Hyp3}_{\widetilde{\Lambda}}(K_3)$ *is true.*

Proof. Fix $0 < t \leq T_0$, a direction $\eta \in \mathbb{R}^d \setminus \{0\}$, and define the function

$$z \mapsto F(z) := \hat{f}(t, \eta^- + z\tfrac{\eta}{|\eta|})$$

of the single real variable z, where we think of η^- as given in the η-parametrisation (3.8) for some θ and $\omega \in \mathbb{S}^{d-2}(\eta)$, and where we drop, for simplicity, the dependence on the time t in our notation for F and f. Then, since $\|f(t,\cdot)\|_{L^1_m} \leq A_m$ one has $\hat{f}(t,\cdot) \in \mathscr{C}^m(\mathbb{R}^d)$ and thus also $F \in \mathscr{C}^m(\mathbb{R})$ with $\|F\|_{L^\infty} \leq A_m$, $\|\partial_z^m F\|_{L^\infty} \leq (2\pi)^m A_m$, and Corollary 3.8 applied to F now gives

$$|\hat{f}(\eta^-)| \leq L_{m,1}\left(\int_0^2 |\hat{f}(\eta^- + z\tfrac{\eta}{|\eta|})|^2\,\mathrm{d}z\right)^{\frac{m}{2m+2}}.$$

We multiply this with the radially increasing weight G to get

$$G(\eta^-)^{\frac{2m}{2m+1}}|\hat{f}(\eta^-)| \le L_{m,1}\left(\int_0^2 |G(\eta^- + z\tfrac{\eta}{|\eta|})\hat{f}(\eta^- + z\tfrac{\eta}{|\eta|})|^2\,\mathrm{d}z\right)^{\frac{m}{2m+2}}.$$

Integrating this with respect to ω and θ, where we think of $\eta^- = \eta^-(\eta,\theta,\omega)$ in the parametrisation (3.8), and using Jensen's inequality for concave functions, one gets

$$\begin{aligned}&\int_{\theta_0}^{\frac{\pi}{2}}\int_{\mathbb{S}^{d-2}(\eta)} G(\eta^-)^{\frac{2m}{2m+1}}|\hat{f}(\eta^-)|\,\mathrm{d}\theta\,\mathrm{d}\omega\\ &\le L_{m,1}(\tfrac{\pi}{2})^{\frac{m+1}{2m+1}}|\mathbb{S}^{d-2}|^{\frac{m+1}{2m+1}}\\ &\qquad\times\left(\int_{\theta_0}^{\frac{\pi}{2}}\int_{\mathbb{S}^{d-2}(\eta)}\int_0^2 |G(\eta^- + z\tfrac{\eta}{|\eta|})\hat{f}(\eta^- + z\tfrac{\eta}{|\eta|})|^2\,\mathrm{d}z\,\mathrm{d}\theta\,\mathrm{d}\omega\right)^{\frac{m}{2m+1}}.\end{aligned}\tag{3.44}$$

Now assume that $|\eta| \ge \Lambda_0$. Because of the first part of Lemma 3.25, we can further bound

$$\begin{aligned}(3.44) &\le L_{m,1}(\tfrac{\pi}{2})^{\frac{m+1}{2m+1}}|\mathbb{S}^{d-2}|^{\frac{m+1}{2m+1}}\left(\frac{2}{\Lambda_0\cos\theta_0}\right)^{\frac{m}{2m+1}}\\ &\quad\times\left(\int_{\mathbb{S}^{d-2}(\eta)}\int_{\Lambda_0\sin^2\frac{\theta_0}{2}}^{\frac{|\eta|}{2}+2}\int_{\Lambda_0\sin\theta_0}^{\frac{|\eta|}{2}} |G(x\tfrac{\eta}{|\eta|} - y\omega)\hat{f}(x\tfrac{\eta}{|\eta|} - y\omega)|^2\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}\omega\right)^{\frac{m}{2m+1}}\\ &\le L_{m,1}(\tfrac{\pi}{2})^{\frac{m+1}{2m+1}}|\mathbb{S}^{d-2}|^{\frac{m+1}{2m+1}}\left(\frac{2}{\Lambda_0\cos\theta_0}\right)^{\frac{m}{2m+1}}(\Lambda_0\sin\theta_0)^{2-d}\\ &\quad\times\left(\int_{\mathbb{S}^{d-2}(\eta)}\int_{\Lambda_0\sin^2\frac{\theta_0}{2}}^{\frac{|\eta|}{2}+2}\int_{\Lambda_0\sin\theta_0}^{\frac{|\eta|}{2}} |G(x\tfrac{\eta}{|\eta|} - y\omega)\hat{f}(x\tfrac{\eta}{|\eta|} - y\omega)|^2\,y^{d-2}\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}\omega\right)^{\frac{m}{2m+1}}.\end{aligned}$$

Again, the integration measure $y^{d-2}\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}\omega$ is d-dimensional Lebesgue measure in the cylindrical coordinates $(x, y\omega)$ with respect to the cylinder in the η direction. One checks that the condition $\Lambda \ge \Lambda_0 \ge 3$ ensures that

$$(\widetilde{\Lambda}/2 + 2)^2 + (\widetilde{\Lambda}/2) \le (\sqrt{2}\Lambda)^2$$

so since $|\eta| \le \widetilde{\Lambda}$, we can extend the integration above to a ball of radius $\sqrt{2}\Lambda$ to get

$$\begin{aligned}(3.44) &\le L_{m,1}(\tfrac{\pi}{2})^{\frac{m+1}{2m+1}}|\mathbb{S}^{d-2}|^{\frac{m+1}{2m+1}}\left(\frac{2}{\Lambda_0\cos\theta_0}\right)^{\frac{m}{2m+1}}(\Lambda_0\sin\theta_0)^{2-d}\|G_{\sqrt{2}\Lambda}f\|_{L^2(\mathbb{R}^d)}^{\frac{2m}{2m+1}}\\ &\le L_{m,1}(\tfrac{\pi}{2})^{\frac{m+1}{2m+1}}|\mathbb{S}^{d-2}|^{\frac{m+1}{2m+1}}\left(\frac{2}{\Lambda_0\cos\theta_0}\right)^{\frac{m}{2m+1}}(\Lambda_0\sin\theta_0)^{2-d}B^{\frac{2m}{2m+1}}.\end{aligned}\tag{3.45}$$

If $|\eta| \le \Lambda_0$ we simply bound

$$\begin{aligned}\int_{\theta_0}^{\frac{\pi}{2}}\int_{\mathbb{S}^{d-2}(\eta)} G(\eta^-)^{\frac{2m}{2m+1}}|\hat{f}(\eta^-)|\,\mathrm{d}\theta\,\mathrm{d}\omega &\le \|\hat{f}\|_{L^\infty}\frac{\pi}{2}|\mathbb{S}^{d-2}|\mathrm{e}^{\beta T_0(1+\Lambda_0^2/2)}\\ &\le A_m\frac{\pi}{2}|\mathbb{S}^{d-2}|\mathrm{e}^{1+\Lambda_0^2/2}.\end{aligned}\tag{3.46}$$

Concerning the bound in the second half of Hyp3$_{\widetilde{\Lambda}}$, a completely analogous calculation as the one above, using the second half of Lemma 3.25 gives for $\Lambda_0 \le |\eta^+| \le \widetilde{\Lambda}$,

$$\begin{aligned}
&\int_{\vartheta_0}^{\frac{\pi}{2}} \int_{\mathbb{S}^{d-2}(\eta^+)} G\left(t, \eta^-(\eta^+, \vartheta, \omega)\right)^{\frac{2m}{2m+1}} \left|\hat{f}\left(\eta^-(\eta^+, \vartheta, \omega)\right)\right| \\
&\qquad\qquad\qquad\qquad \times \mathbb{1}_{\frac{\Lambda}{\sqrt{2}}}(|\eta^-(\eta^+, \vartheta, \omega)|)\, \mathrm{d}\omega\, \mathrm{d}\vartheta \\
&\le L_{m,1}(\tfrac{\pi}{2})^{\frac{m+1}{2m+1}} |\mathbb{S}^{d-2}|^{\frac{m+1}{2m+1}} \left(\frac{1}{2\Lambda_0}\right)^{\frac{m}{2m+1}} (\Lambda_0 \tan\vartheta_0)^{2-d} \qquad (3.47)\\
&\quad \times \left(\int_{\mathbb{S}^{d-2}(\eta^+)} \int_0^2 \int_0^{\frac{\widetilde{\Lambda}}{\sqrt{2}}} |G(x\tfrac{\eta}{|\eta|} - y\omega)\hat{f}(x\tfrac{\eta}{|\eta|} - y\omega)|^2\, y^{d-2} \mathrm{d}y\, \mathrm{d}x\, \mathrm{d}\omega \right)^{\frac{m}{2m+1}} .
\end{aligned}$$

By our choice of $\widetilde{\Lambda}$ and Λ_0, we always have $2^2 + (\widetilde{\Lambda}/2)^2 \le (\sqrt{2}\Lambda)^2$, so we can extend the integration above to the whole ball $|\eta^+| \le \sqrt{2}\Lambda$ to see

$$\begin{aligned}
(3.47) &\le L_{m,1}(\tfrac{\pi}{2})^{\frac{m+1}{2m+1}} |\mathbb{S}^{d-2}|^{\frac{m+1}{2m+1}} \left(\frac{1}{2\Lambda_0}\right)^{\frac{m}{2m+1}} (\Lambda_0 \tan\vartheta_0)^{2-d} \|G_{\sqrt{2}\Lambda} f\|_{L^2(\mathbb{R}^d)}^{\frac{2m}{2m+1}} \\
&\le L_{m,1}(\tfrac{\pi}{2})^{\frac{m+1}{2m+1}} |\mathbb{S}^{d-2}|^{\frac{m+1}{2m+1}} \left(\frac{1}{2\Lambda_0}\right)^{\frac{m}{2m+1}} (\Lambda_0 \tan\vartheta_0)^{2-d} B^{\frac{2m}{2m+1}} . \qquad (3.48)
\end{aligned}$$

If $|\eta^+| \le \Lambda_0$ we simply bound as above

$$\int_{\vartheta_0}^{\frac{\pi}{4}} \int_{\mathbb{S}^{d-2}(\eta^+)} G(\eta^-)^{\frac{2m}{2m+1}} |\hat{f}(\eta^-)|\, \mathrm{d}\vartheta\, \mathrm{d}\omega \le A_m \frac{\pi}{4} |\mathbb{S}^{d-2}| e^{1+\Lambda_0^2}. \qquad (3.49)$$

Now we set K_3 equal to the maximum of the constants in (3.45), (3.46), (3.48), (3.49). With this choice, K_3 depends only on d, m, A_m, and B and Hyp3$_{\widetilde{\Lambda}}(K_3)$ is true. ∎

Proof of Theorem 3.21. In view of Lemmata 3.24 and 3.26, a suitable choice for A_m and B is

$$B := \|f_0\|_{L^2(\mathbb{R}^d)} e^{C_{f_0} T_0}, \qquad A_m := \sup_{t \ge 0} \|f(t, \cdot)\|_{L^1_m(\mathbb{R}^d)}.$$

The finiteness of A_m is guaranteed since $f_0 \in L^1_m(\mathbb{R}^d)$, see Remark 3.4. We again choose the length scales Λ_N to be

$$\Lambda_N := \frac{\Lambda_{N-1} + \sqrt{2}\Lambda_{N-1}}{2} = \frac{1+\sqrt{2}}{2} \Lambda_{N-1} = \left(\frac{1+\sqrt{2}}{2}\right)^N \Lambda_0$$

with $\Lambda_0 = 3$, see (3.43), and we set

$$M_3 := \max\left\{ 2|\mathbb{S}^{d-2}| A_m + 1, K_3 \right\},$$

with the constant K_3 from Lemma 3.26. Since

$$\sup_{0\le t\le T_0}\sup_{|\eta|\le\sqrt{2}\Lambda}\int_{\theta_0}^{\frac{\pi}{2}}\int_{\mathbb{S}^{d-2}(\eta)} G\left(t,\eta^-(\eta,\theta,\omega)\right)^{\frac{2m}{2m+1}}\left|\hat{f}\left(\eta^-(\eta,\theta,\omega)\right)\right| \,\mathrm{d}\omega\,\mathrm{d}\theta$$
$$\le \frac{\pi}{2}\,|\mathbb{S}^{d-2}|\,e^{\frac{2m}{2m+1}\beta T_0(1+\Lambda_0^2)^{\alpha}}\,A_m,$$

and similarly for the η^+ term, it follows from our choice of M_3 that there exists $\beta_0 > 0$ such that $\mathrm{Hyp3}_{\Lambda_0}(M_3)$ is true for all $0 \le \beta \le \beta_0$.

Now, we pick

$$\beta = \min\left(\beta_0, T_0^{-1}, \frac{\tilde{C}_{f_0}}{\alpha T_0[(1+2^{d-1})c_{b,d,2}M_2 + (C_{\theta_0} + 2^d C_{\vartheta_0})M] + 1}\right)$$

with the constant M_2 from Corollary 3.17, so that the conditions of Lemma 3.24 and 3.26 are fulfilled.

For the induction step assume that $\mathrm{Hyp3}_{\Lambda_N}(M_3)$ is true. Lemma 3.24 then implies

$$\|G_{\sqrt{2}\Lambda_N} f\|_{L^2(\mathbb{R}^d)} \le \|\mathbb{1}_{\sqrt{2}\Lambda_N}(D_v) f_0\|_{L^2(\mathbb{R}^d)}\, e^{C_{f_0}T_0} \le B,$$

and Lemma 3.26 shows that $\mathrm{Hyp3}_{\Lambda_{N+1}}(M_3)$ is true.

It follows that $\mathrm{Hyp3}_{\Lambda_N}(M_3)$ is true for all $N \in \mathbb{N}$, and therefore also

$$\|G_{\sqrt{2}\Lambda_N} f\|_{L^2(\mathbb{R}^d)} \le B$$

for all $N \in \mathbb{N}$. In particular, letting $N \to \infty$, we see that $\|Gf\|_{L^2(\mathbb{R}^d)} \le B$, which concludes the proof of Theorem 3.21. ■

3.7 Removing the L^2 constraint: Gevrey regularity and (ultra-)analyticity of weak solutions

In this section we will give the proofs of Theorem 2.2, 2.4, and 2.5 in a slightly more general form. More precisely, we will prove

Theorem 3.27 (Gevrey smoothing I). *Let $0 < \nu < 1$. Assume that the cross-section b satisfies the conditions* **(B1)–(B3)** *with power-law type singularity (1.18) for $d \ge 2$. For $d = 1$, assume that b_1 satisfies conditions* **(K1)–(K3)**. *Let f be a weak solution of the Cauchy problem (1.2) with initial datum $f_0 \ge 0$ and $f_0 \in L^1_m(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$ for some integer $m \ge 2$. Then, for all $0 < \alpha \le \min\{\alpha_{m,d}, \nu\}$,*

$$f(t,\cdot) \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d) \tag{3.50}$$

for all $t > 0$, where $\alpha_{m,d} = \frac{\log[(4m+d)/(2m+d)]}{\log 2}$.

Theorem 3.28 (Gevrey smoothing II). *Let $d \geq 2$. Assume that the cross-section b satisfies the conditions of Theorem 3.27. Let f be a weak solution of the Cauchy problem (1.2) with initial datum $f_0 \geq 0$ and $f_0 \in L^1_m(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$ for some integer $m \geq 2$. Then, for all $0 < \alpha \leq \min\{\alpha_{m,2}, \nu\}$,*

$$f(t,\cdot) \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d) \tag{3.51}$$

for all $t > 0$, where $\alpha_{m,2} = \frac{\log[(4m+2)/(2m+2)]}{\log 2}$. In particular, the weak solution is real analytic if $\nu = \frac{1}{2}$ and ultra-analytic if $\nu > \frac{1}{2}$ in any dimension.

If the integrability condition (1.20) is replaced by the slightly stronger condition (2.5), which is true in all physically relevant cases, we can prove the stronger result

Theorem 3.29 (Gevrey smoothing III). *Let $d \geq 2$. Assume that the cross-section b satisfies the conditions of Theorem 3.27 and the condition (2.5), that is, b is bounded away from the singularity. Let f be a weak solution of the Cauchy problem (1.2) with initial datum $f_0 \geq 0$ and $f_0 \in L^1_m(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$ for some integer $m \geq 2$. Then, for all $0 < \alpha \leq \min\{\alpha_{m,1}, \nu\}$,*

$$f(t,\cdot) \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d) \tag{3.52}$$

for all $t > 0$, where $\alpha_{m,1} = \frac{\log[(4m+1)/(2m+1)]}{\log 2}$.

We even have the uniform bound

Corollary 3.30. *Under the same assumptions as in Theorem 3.27 (or 3.28, respectively 3.29), for any weak solution f of the Cauchy problem (1.2) initial datum $f_0 \geq 0$ and $f_0 \in L^1_m(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$ for some integer $m \geq 2$ and for any $0 < \alpha \leq \min\{\alpha_{d,m}, \nu\}$ (or any $0 < \alpha \leq \min\{\alpha_{m,2}, \nu\}$, respectively $0 < \alpha \leq \min\{\alpha_{m,1}, \nu\}$) there exist constants $0 < K, C < \infty$ such that*

$$\sup_{0 \leq t < \infty} \sup_{\eta \in \mathbb{R}^d} \mathrm{e}^{K \min(t,1) \langle \eta \rangle^{2\alpha}} |\hat{f}(t,\eta)| \leq C. \tag{3.53}$$

Proof of Theorems 3.27 through 3.29. In the case where the initial condition f_0 obeys $f_0 \geq 0$ and $f_0 \in L^1_m(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$ for some integer $m \geq 2$, but is not necessarily in $L^2(\mathbb{R}^d)$, we use the known H^∞ smoothing of the Boltzmann [DW04, AE05, MUXY09] and Kac equation[2] [LX09] in a mild way (see also Appendix B): for $\tau > 0$ one has $f(\tau,\cdot) \in L^2(\mathbb{R}^d)$ and using this as a new initial condition in Theorems 2.2 through 2.5, and noting that T_0 in those theorems is arbitrary, this implies that $f(t,\cdot) \in G^{\frac{1}{2\alpha}}(\mathbb{R}^d)$ for $t > 0$. ■

[2] A H^∞ smoothing effect for the homogeneous non-cutoff Kac equation was first proved by L. Desvillettes [Des95], but under the stronger assumption that all polynomial moments of the initial datum f_0 are bounded, i.e. $f_0 \in L^1_k(\mathbb{R}) \cap L\log L(\mathbb{R})$ for all $k \in \mathbb{N}$.

Proof of Corollary 3.30. Using known results about propagation of Gevrey regularity by Desvillettes, Furioli, and Terraneo [DFT09] for the non-cutoff homogeneous Boltzmann and Kac equation for Maxwellian molecules, the bounds from Corollary 3.12 through 3.22 extend to all times. ∎

CHAPTER 4

Strong smoothing for Maxwellian molecules with Debye-Yukawa type interaction

In this chapter we prove the strong smoothing property of the homogeneous Boltzmann equation for Debye-Yukawa type Maxwellian molecules. The main differences compared with Chapter 3 are:

(1) For the weights needed in the proof of Theorem 2.10 we have a much stronger enhanced subadditivity bound, see Lemma 4.1. The proof is more involved than the one in the proof of Gevrey smoothing, Lemma 3.3, though.

(2) Because of the stronger form of the subadditivity bound, we can allow for a bigger loss in the induction step. We can therefore work with a more straightforward version of the 'impossible' L^2-to-L^∞ bound, see Lemma 4.3.

(3) Due to the special form of the weights we use in the Debye-Yukawa case, which are in some sense in between the power type weights used in [MUXY09] and the sub-Gaussian weight described in Section 2.2, we do not have to do much of the additional songs and dances from Chapter 3.

4.1 Enhanced subadditivity and properties of the Fourier weights

Lemma 4.1. *Let $\mu > 0$ and $h : [0, \infty) \to [0, \infty)$, $s \mapsto h(s) = \big(\log(\alpha + s)\big)^{\mu+1}$ for some $\alpha \geq e^{\mu}$. Then h is increasing, concave and for any $0 \leq s_- \leq s_+$,*

$$h(s_- + s_+) \leq \frac{\mu + 1}{1 + \log \alpha} h(s_-) + h(s_+). \tag{4.1}$$

Remark. For $\alpha \geq e^{\mu}$, one has $h(0) \geq \mu^{\mu+1} > 0$, and from the concavity of h one concludes the subadditivity estimate

$$h(s_-) + h(s_+) \geq h(s_- + s_+) + h(0) > h(s_- + s_+)$$

for all $s_-, s_+ \geq 0$. Note that this is the best possible bound for general $s_-, s_+ \geq 0$. For $0 \leq s_- \leq s_+$ Lemma 4.1 shows that the subadditivity bound can be improved to gain the small factor $\frac{\mu+1}{1+\log \alpha}$, which is strictly less than one for $\alpha > e^{\mu}$, in front of $h(s_-)$. So this is indeed an *enhanced* subadditivity property of the function h.

Lemma 4.1 plays a similar role in the proof of Theorem 2.10, as Lemma 3.3 in Chapter 3. Here the situation is a bit simpler than in Chapter 3, since by choosing α large enough, we can make the term $\frac{\mu+1}{1+\log \alpha}$ as small as we like. ♦

Proof. Since

$$h'(s) = \frac{\mu + 1}{\alpha + s} \big(\log(\alpha + s)\big)^{\mu} \geq 0 \quad \text{if } \alpha \geq 1,$$

the function h is increasing. Further,

$$\begin{aligned} h''(s) &= \frac{\mu + 1}{(\alpha + s)^2} \big(\log(\alpha + s)\big)^{\mu-1} \big(\mu - \log(\alpha + s)\big) \\ &\leq \frac{\mu + 1}{(\alpha + s)^2} \big(\log(\alpha + s)\big)^{\mu-1} \big(\mu - \log(\alpha)\big) \leq 0 \end{aligned}$$

for $\alpha \geq e^{\mu}$, so h is concave.

For all $s_-, s_+ \geq 0$,

$$h(s_- + s_+) = h(s_-) \frac{h(s_- + s_+) - h(s_+)}{h(s_-)} + h(s_+),$$

and by concavity, $s_+ \mapsto h(s_- + s_+) - h(s_+)$ is decreasing, so using $0 \leq s_- \leq s_+$ one has

$$h(s_- + s_+) \leq h(s_-) \frac{h(2s_-) - h(s_-)}{h(s_-)} + h(s_+).$$

Since h' is decreasing,

$$h(2s_-) - h(s_-) = \int_{s_-}^{2s_-} h'(r) \, \mathrm{d}r \leq h'(s_-) s_-$$

and we get

$$h(s_- + s_+) \leq h(s_-)\frac{h'(s_-)s_-}{h(s_-)} + h(s_+) = h(s_-)\frac{(\mu+1)s_-}{(\alpha+s_-)\log(\alpha+s_-)} + h(s_+).$$

For $\alpha \geq 1$ the function $F_\alpha : [0,\infty) \to \mathbb{R}$, $F_\alpha(s) := (\alpha+s)\log(\alpha+s)$, is strictly convex and thus

$$F_\alpha(s) \geq F_\alpha(0) + F'_\alpha(0)s = \alpha\log\alpha + (1+\log\alpha)s \geq (1+\log\alpha)s.$$

It follows that $\frac{s_-}{(\alpha+s_-)\log(\alpha+s_-)} \leq \frac{1}{1+\log\alpha}$ and therefore

$$h(s_- + s_+) \leq h(s_-)\frac{\mu+1}{1+\log\alpha} + h(s_+).$$

■

Proposition 4.2. *Let $\beta, t, \mu > 0$, $\alpha \geq \mathrm{e}^\mu$ and define the function $\widetilde{G} : [0,\infty) \to \mathbb{R}$ by*

$$\widetilde{G}(r) := \mathrm{e}^{\beta t 2^{-\mu-1}(\log(\alpha+r))^{\mu+1}}.$$

Then for all $0 \leq s_- \leq s_+$ with $s_- + s_+ = s$ one has

$$\left|\widetilde{G}(s) - \widetilde{G}(s_+)\right| \leq 2^{-\mu}\beta t(\mu+1)\left(1 - \frac{s_+}{s}\right)(\log(\alpha+s))^\mu\,\widetilde{G}(s_-)^{\frac{\mu+1}{1+\log\alpha}}\,\widetilde{G}(s_+).$$

Proof. Using

$$\widetilde{G}'(s) = 2^{-\mu-1}\beta t(\mu+1)\frac{1}{\alpha+s}\,(\log(\alpha+s))^\mu\,\widetilde{G}(s)$$

one has

$$\widetilde{G}(s) - \widetilde{G}(s_+) = \int_{s_+}^{s} \widetilde{G}'(r)\,\mathrm{d}r \leq 2^{-\mu-1}\beta t(\mu+1)\frac{s-s_+}{\alpha+s_+}\,(\log(\alpha+s))^\mu\,\widetilde{G}(s),$$

where we used that $s_+ \leq s$ and the fact that $\widetilde{G}$ is increasing. Since $s_- + s_+ = s$ and $0 \leq s_- \leq s_+$, in particular $s_+ \geq \frac{s}{2}$, we can further estimate

$$\frac{s-s_+}{\alpha+s_+} = \left(1 - \frac{s_+}{s}\right)\frac{s}{\alpha+s_+} \leq \left(1 - \frac{s_+}{s}\right)\frac{2s_+}{\alpha+s_+} \leq 2\left(1 - \frac{s_+}{s}\right),$$

to obtain

$$\widetilde{G}(s) - \widetilde{G}(s_+) \leq 2^{-\mu}\beta t(\mu+1)\left(1 - \frac{s_+}{s}\right)(\log(\alpha+s))^\mu\,\widetilde{G}(s).$$

The rest now follows from the enhanced subadditivity property (4.1), namely

$$\widetilde{G}(s) = \widetilde{G}(s_- + s_+) \leq \widetilde{G}(s_-)^{\frac{\mu+1}{1+\log\alpha}}\,\widetilde{G}(s_+).$$

■

4.2 Extracting L^∞ bounds from L^2: a simple proof

Following is a simple bound which controls the size of a function h in terms of its local L^2 norm and some global a priori bounds on h and its derivative.

Lemma 4.3. *Let $h \in \mathscr{C}_b^1(\mathbb{R}^d)$, i.e. h is a bounded continuously differentiable function with bounded derivative. Then there exists a constant $L < \infty$ (depending only on d, $\|h\|_{L^\infty(\mathbb{R}^d)}$ and, $\|\nabla h\|_{L^\infty(\mathbb{R}^d)}$) such that for any $x \in \mathbb{R}^d$,*

$$|h(x)| \leq L \left(\int_{Q_x} |h(y)|^2 \, \mathrm{d}y \right)^{\frac{1}{d+2}}, \tag{4.2}$$

where Q_x is a unit cube in $\mathbb{R}^d$ with x being one of the corners, oriented away from the origin in the sense that $x \cdot (y - x) \geq 0$ for all $y \in Q_x$.

Remark. (1) We use the norm $\|\nabla h\|_{L^\infty(\mathbb{R}^d)} = \sup_{\eta \in \mathbb{R}^d} |\nabla h(\eta)|$, where $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$.

(2) The exponent $\frac{1}{d+2}$ can be improved if higher derivatives of the function h are bounded, see Section 3.3. This was important for the results of Chapter 3, but we don't need it here because of the stronger form of the enhanced subadditivity Lemma for the weight we consider in this chapter.

♦

Remark. If $f \in L_1^1(\mathbb{R}^d)$, its Fourier transform satisfies $\hat{f} \in \mathscr{C}_b^1(\mathbb{R}^d)$ by the Riemann-Lebesgue lemma. Since $\nabla_\eta \hat{f}(\eta) = \widehat{2\pi i v f}(\eta)$ one has the a priori bound

$$\|\nabla \hat{f}\|_{L^\infty(\mathbb{R}^d)} \leq 2\pi \|f\|_{L_1^1(\mathbb{R}^d)}.$$

If f is a weak solution of the homogeneous Boltzmann equation, we can also bound $\|\nabla \hat{f}\|_{L^\infty(\mathbb{R}^d)} \leq 2\pi \|f_0\|_{L_2^1(\mathbb{R}^d)}$ uniformly in time due to conservation of energy.

♦

Proof. We first consider the one-dimensional case and prove the d-dimensional result by iteration in each coordinate direction.

Let $u \in \mathscr{C}_b^1(\mathbb{R})$ and $q \geq 1$. Then for any $r \in \mathbb{R}$ we have

$$|u(r)|^q \leq \max\left\{ q\|u'\|_{L^\infty(\mathbb{R})}, \|u\|_{L^\infty(\mathbb{R})} \right\} \int_{I_r} |u(s)|^{q-1} \, \mathrm{d}s, \tag{4.3}$$

where $I_r = [r, r+1]$ if $r \geq 0$ and $I_r = [r-1, r]$ if $r < 0$.

Indeed, assuming for the moment $r \geq 0$,

$$|u(r)|^q - \int_{I_r} |u(s)|^q \, \mathrm{d}s \leq \int_{I_r} |u^q(r) - u^q(s)| \, \mathrm{d}s,$$

and by the fundamental theorem of calculus,

$$|u^q(r) - u^q(s)| \leq q \int_{I_r} |u(t)|^{q-1} |u'(t)| \,\mathrm{d}t \leq q \|u'\|_{L^\infty(\mathbb{R})} \int_{I_r} |u(t)|^{q-1} \,\mathrm{d}t.$$

Combined with the trivial estimate $\int_{I_r} |u(s)|^q \,\mathrm{d}s \leq \|u\|_{L^\infty(\mathbb{R})} \int_{I_r} |u(s)|^{q-1} \,\mathrm{d}s$ one arrives at inequality (4.3) for $r \geq 0$. The case $r < 0$ is analogous.

For the case $d > 1$ we remark that for any $y \in \mathbb{R}^d$,

$$\begin{aligned} \|h(y_1, \dots, y_{j-1}, \cdot, y_{j+1}, \dots, y_d)\|_{L^\infty(\mathbb{R})} &\leq \|h\|_{L^\infty(\mathbb{R}^d)}, \quad \text{and} \\ \|\partial_j h(y_1, \dots, y_{j-1}, \cdot, y_{j+1}, \dots, y_d)\|_{L^\infty(\mathbb{R})} &\leq \|\nabla h\|_{L^\infty(\mathbb{R}^d)} \end{aligned}$$

and setting $q = d + 2$ iterative application of (4.3) in each coordinate direction yields for $x \in \mathbb{R}^d$

$$|h(x)|^{d+2} \leq \left(\max\left\{(d+2)\|\nabla h\|_{L^\infty(\mathbb{R}^d)}, \|h\|_{L^\infty(\mathbb{R}^d)}\right\}\right)^d \int_{I_{x_1} \times \cdots \times I_{x_d}} |h(y)|^{d+2-d} \,\mathrm{d}y,$$

hence

$$\begin{aligned} |h(x)| &\leq \left(\max\left\{(d+2)\|\nabla h\|_{L^\infty(\mathbb{R}^d)}, \|h\|_{L^\infty(\mathbb{R}^d)}\right\}\right)^{\frac{d}{d+2}} \left(\int_{Q_x} |h(y)|^2 \,\mathrm{d}y\right)^{\frac{1}{d+2}} \\ &=: L \|h\|_{L^2(Q_x)}^{\frac{2}{d+2}} \end{aligned}$$

where $Q_x = I_{x_1} \times \cdots \times I_{x_d}$ is a unit cube directed away from the origin with $x \in \mathbb{R}^d$ at one of its corners. ∎

4.3 Smoothing property of the Boltzmann operator

A central step in the proof of Theorem 2.10 is to prove a version for L^2 initial data first. This is the content of Theorem 4.4 below. In the remainder of this chapter we will always assume that the collision kernel satisfies assumptions **(B1)–(B3)** with angular singularity of Debye-Yukawa type (1.19).

Theorem 4.4. *Let f be a weak solution of the Cauchy problem (1.2) with initial datum $f_0 \geq 0$, $f_0 \in L^1_2(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$ and* in addition *$f_0 \in L^2(\mathbb{R}^d)$.*

Set $\alpha = e^{\frac{d}{2} + \frac{d+2}{2}\mu}$. Then for all $T_0 > 0$ there exist $\beta, M > 0$ such that for all $t \in [0, T_0]$

$$\sup_{\eta \in \mathbb{R}^d} e^{\frac{d}{d+2}\beta t (\log\langle \eta \rangle_\alpha)^{\mu+1}} |\hat{f}(t,\eta)| \leq M,$$

and

$$\left\| e^{\beta t (\log\langle D_v \rangle_\alpha)^{\mu+1}} f(t,\cdot) \right\|_{L^2} \leq \|f_0\|_{L^2} e^{C_{f_0} T_0},$$

where the constant C_{f_0} depends only *on $\|f_0\|_{L^1_2}$ and $\|f\|_{L\log L}$.*

We give the proof of Theorem 4.4 in Section 4.4. To prepare for its proof, let $\alpha \geq e^{\mu}$ and $\beta > 0$ and define the Fourier multiplier $G : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$ by

$$G(t,\eta) := e^{\beta t \left(\log\langle\eta\rangle_\alpha\right)^{\mu+1}}, \quad \langle\eta\rangle_\alpha := \left(\alpha + |\eta|^2\right)^{\frac{1}{2}}$$

and for $\Lambda > 0$ the cut-off multiplier $G_\Lambda : \mathbb{R}_+ \times \mathbb{R}^d \to [0,\infty)$ by

$$G_\Lambda(t,\eta) := G(t,\eta)\mathbb{1}_\Lambda(|\eta|)$$

wehre $\mathbb{1}_\Lambda$ is the characteristic function of the interval $[0,\Lambda]$. The associated Fourier multiplication operator is denoted by $G_\Lambda(t,D_v)$,

$$G_\Lambda(t,D_v)f := \mathcal{F}^{-1}\left[G_\Lambda(t,\cdot)\hat{f}(t,\cdot)\right]$$

Note that, due to the cut-off in Fourier space,

$$G_\Lambda f, G_\Lambda^2 f \in L^\infty([0,T_0]; H^\infty(\mathbb{R}^d))$$

for any finite $T_0 > 0$ and $\Lambda > 0$, if $f \in L^\infty([0,T_0]; L^1(\mathbb{R}^d))$, and even analytic in a strip containing $\mathbb{R}^d_v$. In particular, by Sobolev embedding, $G_\Lambda f, G_\Lambda^2 f \in L^\infty([0,T_0]; W^{2,\infty}(\mathbb{R}^d))$, so

$$\left\langle Q(f,f)(t,\cdot), G_\Lambda^2 f(t,\cdot)\right\rangle$$

is well-defined.

As in the power-law case in Chapter 3, we can derive an L^2-reformulation of weak solutions, see Proposition 3.1 and Appendix A. Together with the coercivity estimate from Proposition 2.12 this implies

Corollary 4.5 (A priori bound for weak solutions). *Let f be a weak solution of the Cauchy problem (1.2) with initial datum $f_0 \geq 0$ satisfying $f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$, and let $T_0 > 0$. Then there exist constants $\widetilde{C}_{f_0}, C_{f_0} > 0$ (depending only on the dimension d, the collision kernel b, $\|f_0\|_{L^1_2}$ and $\|f_0\|_{L\log L}$) such that for all $t \in (0,T_0]$, $\beta,\mu > 0$, $\alpha \geq 0$, and $\Lambda > 0$ we have*

$$\begin{aligned}\|G_\Lambda f\|_{L^2}^2 \leq &\|\mathbb{1}_\Lambda(D_v)f_0\|_{L^2}^2 + 2\widetilde{C}_{f_0}\int_0^t \|G_\Lambda f\|_{L^2}^2 \,\mathrm{d}\tau \\ &+ 2\int_0^t \left(\beta - \frac{C_{f_0}}{(\log(e+\alpha))^{\mu+1}}\right)\left\|\left(\log\langle D_v\rangle_\alpha\right)^{\frac{\mu+1}{2}} G_\Lambda f\right\|_{L^2}^2 \mathrm{d}\tau \\ &+ 2\int_0^t \left\langle G_\Lambda Q(f,f) - Q(f,G_\Lambda f), G_\Lambda f\right\rangle \mathrm{d}\tau.\end{aligned} \tag{4.4}$$

Proof. In order to make use of the coercivity property of the Boltzmann collision operator, we write

$$\begin{aligned}\langle Q(f,f), G_\Lambda^2 f\rangle &= \langle G_\Lambda Q(f,f), G_\Lambda f\rangle \\ &= \langle Q(f,G_\Lambda f), G_\Lambda f\rangle + \langle G_\Lambda Q(f,f) - Q(f,G_\Lambda f), G_\Lambda f\rangle\end{aligned}$$

and estimate the first term with Proposition 2.12.

Since $\partial_\tau G_\Lambda^2(\tau,\eta) = 2\beta \left(\log\langle\eta\rangle_\alpha\right)^{\mu+1} G_\Lambda^2(t,\eta)$, we further have

$$\left\langle f, (\partial_\tau G_\Lambda^2)\, f \right\rangle = 2\beta \left\| \left(\log\langle D_v\rangle_\alpha\right)^{\frac{\mu+1}{2}} G_\Lambda f \right\|_{L^2}^2,$$

and inserting those two results into (3.1), one obtains the claimed inequality (4.4). ■

Controlling the commutation error

Proposition 4.6 (Bound on the Commutation Error)**.** *Let f be a weak solution of the Cauchy problem (1.2) with initial datum $f_0 \geq 0$, $f_0 \in L^1_2(\mathbb{R}^d) \cap L\log L(\mathbb{R}^d)$. Then for all $t, \beta, \mu, \Lambda > 0$ and $\alpha \geq e^\mu$ one has the bound*

$$\begin{aligned} &|\langle G_\Lambda Q(f,f) - Q(f, G_\Lambda f), G_\Lambda f\rangle| \\ &\quad \leq \beta t(\mu+1) \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1 - \frac{|\eta^+|^2}{|\eta|^2}\right) G(\eta^-)^{\frac{\mu+1}{1+\log\alpha}} |\hat{f}(\eta^-)| \\ &\qquad\qquad \times \left(\log\langle\eta\rangle_\alpha\right)^\mu\, G_\Lambda(\eta^+)|\hat{f}(\eta^+)|\, G_\Lambda(\eta)|\hat{f}(\eta)|\, \mathrm{d}\sigma\, \mathrm{d}\eta. \end{aligned} \tag{4.5}$$

Remark. The bound (4.5) is very similar to the one we derived in Chapter 3, see Proposition 3.5. In particular, it is a trilinear expression in the weak solution f. The $\hat{f}(\eta^-)$ term is multiplied by a *faster-than-polynomially* growing function. If the Fourier multiplier G were only growing *polynomially*, the factor $G(\eta^-)^{\frac{\mu+1}{1+\log\alpha}}$ would be replaced by 1, making the analysis *much easier*. We will therefore rely on the inductive procedure we developed in Chapter 3 to treat exactly this type of situation. ♦

Proof. Bobylev's identity and a small computation show that

$$\begin{aligned} &\left|\left\langle G_\Lambda Q(f,f) - Q(f, G_\Lambda f), G_\Lambda f\right\rangle\right| = \left|\left\langle \mathcal{F}\left[G_\Lambda Q(f,f) - Q(f, G_\Lambda f)\right], \mathcal{F}\left[G_\Lambda f\right]\right\rangle_{L^2}\right| \\ &\quad \leq \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right) G_\Lambda(\eta)|\hat{f}(\eta)|\, |\hat{f}(\eta^-)|\, |\hat{f}(\eta^+)| |G(\eta) - G(\eta^+)|\, \mathrm{d}\sigma\, \mathrm{d}\eta \end{aligned}$$

since G_Λ is supported on the ball $\{|\eta| \leq \Lambda\}$ and $|\eta^+| \leq |\eta|$. We further have

$$|\eta^\pm|^2 = \frac{|\eta|^2}{2}\left(1 \pm \frac{\eta\cdot\sigma}{|\eta|}\right), \quad |\eta^-|^2 + |\eta^+|^2 = |\eta|^2,$$

in particular by the support assumption on the collision kernel b, $\frac{\eta\cdot\sigma}{|\eta|} \in [0,1]$, and therefore

$$0 \leq |\eta^-|^2 \leq \frac{|\eta|^2}{2} \leq |\eta^+|^2 \leq |\eta|^2.$$

From Proposition 4.2 it now follows that

$$\begin{aligned}\left|G(\eta)-G(\eta^+)\right| &= \left|\widetilde{G}(|\eta|^2)-\widetilde{G}(|\eta^+|^2)\right| \\ &\leq \beta t(\mu+1)\left(1-\frac{|\eta^+|^2}{|\eta|^2}\right)\left(\log\langle\eta\rangle_\alpha\right)^\mu\ G(\eta^-)^{\frac{\mu+1}{1+\log\alpha}}G(\eta^+),\end{aligned}$$

which completes the proof. ∎

Lemma 4.7.

$$\begin{aligned}&\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1-\frac{|\eta^+|^2}{|\eta|^2}\right)\left(\log\langle\eta\rangle_\alpha\right)^\mu\ G_\Lambda(\eta^+)|\hat{f}(\eta^+)|\,G_\Lambda(\eta)|\hat{f}(\eta)|\ \mathrm{d}\sigma\,\mathrm{d}\eta \\ &\qquad\leq c_{b,d}\left(\|G_\Lambda f\|_{L^2}^2+\left\|\left(\log\langle D_v\rangle_\alpha\right)^{\frac{\mu}{2}}G_\Lambda f\right\|_{L^2}^2\right),\end{aligned}\tag{4.6}$$

where

$$c_{b,d}=\frac{1}{2}\max\{1,2^{\mu-1}\}\max\{2^{d-1-\mu}(\log 2)^\mu,1+2^{d-1}\}\,|\mathbb{S}^{d-2}|\int_0^{\frac{\pi}{2}}\sin^d\theta\,b(\cos\theta)\ \mathrm{d}\theta.$$

Proof. Using Cauchy-Schwartz, in the form $ab\leq\frac{a^2}{2}+\frac{b^2}{2}$, one can split the integral into

$$\begin{aligned}&\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1-\frac{|\eta^+|^2}{|\eta|^2}\right)\left(\log\langle\eta\rangle_\alpha\right)^\mu\ G_\Lambda(\eta^+)|\hat{f}(\eta^+)|\,G_\Lambda(\eta)|\hat{f}(\eta)|\ \mathrm{d}\sigma\,\mathrm{d}\eta \\ &\leq\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1-\frac{|\eta^+|^2}{|\eta|^2}\right)\left(\log\langle\eta\rangle_\alpha\right)^\mu\ G_\Lambda(\eta)^2|\hat{f}(\eta)|^2\ \mathrm{d}\sigma\,\mathrm{d}\eta \\ &\quad+\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1-\frac{|\eta^+|^2}{|\eta|^2}\right)\left(\log\langle\eta\rangle_\alpha\right)^\mu\ G_\Lambda(\eta^+)^2|\hat{f}(\eta^+)|^2\ \mathrm{d}\sigma\,\mathrm{d}\eta\end{aligned}$$

and we will treat the two terms separately. To estimate the first integral, one introduces polar coordinates such that $\frac{\eta}{|\eta|}\cdot\sigma=\cos\theta$ and thus, since

$$|\eta^+|^2=|\eta|^2\left(1+\frac{\eta}{|\eta|}\cdot\sigma\right)=|\eta|^2\cos^2\frac{\theta}{2},$$

obtains

$$\begin{aligned}I&:=\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1-\frac{|\eta^+|^2}{|\eta|^2}\right)\left(\log\langle\eta\rangle_\alpha\right)^\mu\ G_\Lambda(\eta)^2|\hat{f}(\eta)|^2\ \mathrm{d}\sigma\,\mathrm{d}\eta \\ &=\frac{1}{2}|\mathbb{S}^{d-2}|\int_0^{\frac{\pi}{2}}\sin^{d-2}\theta\,b(\cos\theta)\,\sin^2\tfrac{\theta}{2}\ \mathrm{d}\theta\int_{\mathbb{R}^d}\left(\log\langle\eta\rangle_\alpha\right)^\mu\ G_\Lambda(\eta)^2|\hat{f}(\eta)|^2\ \mathrm{d}\eta \\ &\leq\frac{1}{2}|\mathbb{S}^{d-2}|\int_0^{\frac{\pi}{2}}\sin^d\theta\,b(\cos\theta)\ \mathrm{d}\theta\left\|\left(\log\langle D_v\rangle_\alpha\right)^{\frac{\mu}{2}}G_\Lambda f\right\|_{L^2}^2.\end{aligned}$$

Notice that the θ integral is finite due to the assumptions on the angular collision kernel. This is another instance where cancellation effects play an important role in controlling the singularity for grazing collisions.

It remains to bound the second integral, and we will do this after a change of variables $\eta \to \eta^+$. This change of variables is well-known to the experts, see, for example, [ADVW00, MUXY09]. We give some details for the convenience of the reader.

Observe that $\frac{\eta^+\cdot\sigma}{|\eta^+|} = \frac{|\eta^+|}{|\eta|}$ and $\frac{\eta\cdot\sigma}{|\eta|} = 2\left(\frac{\eta^+\cdot\sigma}{|\eta^+|}\right)^2 - 1$, and by Sylvester's determinant theorem, one has

$$\left|\frac{\partial\eta^+}{\partial\eta}\right| = \left|\frac{1}{2}\left(1 + \frac{\eta}{|\eta|}\otimes\sigma\right)\right| = \frac{1}{2^d}\left(1 + \frac{\eta}{|\eta|}\cdot\sigma\right) = \frac{1}{2^{d-1}}\left(\frac{\eta^+\cdot\sigma}{|\eta^+|}\right)^2 = \frac{1}{2^{d-1}}\frac{|\eta^+|^2}{|\eta|^2}.$$

Since $0 \le |\eta^-| \le |\eta^+|$ and $|\eta|^2 = |\eta^-|^2 + |\eta^+|^2$, in particular $|\eta|^2 \le 2|\eta^+|^2$, it follows that $\left|\frac{\partial\eta^+}{\partial\eta}\right| \ge 2^{-d}$ and

$$\log\langle\eta\rangle_\alpha = \tfrac{1}{2}\log(\alpha + |\eta|^2) \le \tfrac{1}{2}\log 2 + \tfrac{1}{2}\log(\alpha + |\eta^+|^2) = \tfrac{1}{2}\log 2 + \log\langle\eta^+\rangle_\alpha.$$

For all $x, y \ge 0$ one has

$$\begin{cases} (x+y)^\mu \le 2^{\mu-1}(x^\mu + y^\mu) & \text{for } \mu \ge 1 \text{ by convexity, and} \\ (x+y)^\mu \le x^\mu + y^\mu & \text{for } \mu < 1, \end{cases}$$

where the second statement is a consequence of the fact that for $0 < \mu < 1$ the function $0 \le s \mapsto h(s) = (1+s)^\mu - s^\mu$ is monotone decreasing for all $s > 0$ with $h(0) = 1$. Therefore,

$$\left(\log\langle\eta\rangle_\alpha\right)^\mu \le \max\{1, 2^{\mu-1}\}\left(2^{-\mu}(\log 2)^\mu + \left(\log\langle\eta^+\rangle_\alpha\right)^\mu\right).$$

After those preparatory remarks, we can estimate

$$
\begin{aligned}
I^+ &:= \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1-\frac{|\eta^+|^2}{|\eta|^2}\right)(\log\langle\eta\rangle_\alpha)^\mu\; G_\Lambda(\eta^+)^2|\hat{f}(\eta^+)|^2\,\mathrm{d}\sigma\,\mathrm{d}\eta \\
&= \frac{1}{2}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d} b\left(2\left(\frac{\eta^+\cdot\sigma}{|\eta^+|}\right)^2-1\right)\left(1-\left(\frac{\eta^+\cdot\sigma}{|\eta^+|}\right)^2\right) \\
&\qquad\qquad\qquad \times (\log\langle\eta\rangle_\alpha)^\mu\; G_\Lambda(\eta^+)^2|\hat{f}(\eta^+)|^2\left|\frac{\partial\eta^+}{\partial\eta}\right|^{-1}\,\mathrm{d}\eta^+\,\mathrm{d}\sigma \\
&\le 2^{d-1}\max\{1,2^{\mu-1}\} \\
&\quad\times\Bigg[2^{-\mu}(\log 2)^\mu\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(2\left(\frac{\eta^+\cdot\sigma}{|\eta^+|}\right)^2-1\right)\left(1-\left(\frac{\eta^+\cdot\sigma}{|\eta^+|}\right)^2\right) \\
&\qquad\qquad\qquad \times G_\Lambda(\eta^+)^2|\hat{f}(\eta^+)|^2\,\mathrm{d}\sigma\,\mathrm{d}\eta^+ \\
&\qquad + \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(2\left(\frac{\eta^+\cdot\sigma}{|\eta^+|}\right)^2-1\right)\left(1-\left(\frac{\eta^+\cdot\sigma}{|\eta^+|}\right)^2\right) \\
&\qquad\qquad\qquad \times (\log\langle\eta^+\rangle_\alpha)^\mu\; G_\Lambda(\eta^+)^2|\hat{f}(\eta^+)|^2\,\mathrm{d}\sigma\,\mathrm{d}\eta^+\Bigg].
\end{aligned}
$$

Introducing new polar coordinates with pole $\frac{\eta^+}{|\eta^+|}$, such that $\cos\vartheta = \frac{\eta^+\cdot\sigma}{|\eta^+|} \ge \frac{1}{\sqrt{2}}$, i.e. $\vartheta\in[0,\frac{\pi}{4}]$, one then gets

$$
\begin{aligned}
I^+ \le\; & 2^{d-1}\max\{1,2^{\mu-1}\}\,|\mathbb{S}^{d-2}|\int_0^{\frac{\pi}{4}}\sin^d\vartheta\, b(\cos 2\vartheta)\,\mathrm{d}\vartheta \\
&\times\left[2^{-\mu}(\log 2)^\mu\|G_\Lambda f\|_{L^2}^2 + \left\|(\log\langle D_v\rangle_\alpha)^{\frac{\mu}{2}}\, G_\Lambda f\right\|_{L^2}^2\right].
\end{aligned}
$$

Estimating $\int_0^{\frac{\pi}{4}}\sin^d\vartheta\, b(\cos 2\vartheta)\,\mathrm{d}\vartheta \le \int_0^{\frac{\pi}{2}}\sin^d\theta\, b(\cos\theta)\,\mathrm{d}\theta$ and combining the bounds on I and I^+ proves inequality (4.6). ∎

4.4 Smoothing effect for L^2 initial data: Proof of Theorem 4.4

We now have all the necessary pieces together to start the inductive proof of Theorem 4.4 for initial data that are in addition square integrable.

The proof is based on gradually removing the cut-off Λ in Fourier space, in such a way that the commutation error can be controlled, even though it contains fast growing terms. For fixed $T_0, \mu > 0$ and $\alpha \ge \mathrm{e}^\mu$ we define

Definition 4.8 (Induction Hypothesis $\mathrm{Hyp}_\Lambda(M)$.)**.** Let $M \geq 0$ and $\Lambda > 0$. Then for all $0 \leq t \leq T_0$,

$$\sup_{|\xi|\leq\Lambda} G(t,\xi)^{\frac{\mu+1}{1+\log\alpha}}|\hat{f}(t,\xi)| \leq M.$$

Remark. Recall that the Fourier multiplier G also depends on $\beta > 0$ and $\alpha \geq \mathrm{e}^\mu$ and we suppress this dependence here.

The induction step itself will be divided into two separate steps:

Step 1 $\mathrm{Hyp}_\Lambda(M) \implies \|G_{\sqrt{2}\Lambda} f\|_{L^2} \leq C$ via a *Gronwall argument*.

Step 2 $L^2 \to L^\infty$ *bound:*

$$\|G_{\sqrt{2}\Lambda} f\|_{L^2} \leq C \implies \mathrm{Hyp}_{\widetilde{\Lambda}}(M) \text{ for intermediate } \widetilde{\Lambda} = \frac{1+\sqrt{2}}{2}\Lambda.$$

Here it is essential that M does not increase during the induction procedure. This can be accomplished by choosing β small enough at very beginning.

♦

Lemma 4.9 (Step 1)**.** *Fix T_0, $\mu > 0$ and $\alpha \geq \mathrm{e}^\mu$ and let $M \geq 0$ and $\Lambda > 0$. Let further C_{f_0}, $\widetilde{C}_{f_0}$ and $c_{b,d}$ be the constants from Corollary 4.5 and Lemma 4.7, respectively. If*

$$0 < \beta \leq \beta_0(\alpha) := \frac{C_{f_0}}{(\log(\mathrm{e}+\alpha))^{\mu+1}} \frac{\log\alpha}{\log\alpha + 2T_0(\mu+1)c_{b,d}M}, \tag{4.7}$$

then for any weak solution of the Cauchy problem (1.2) with initial datum $f_0 \geq 0$, $f_0 \in L^1_2 \cap L\log L$,

$$\mathrm{Hyp}_\Lambda(M) \implies \|G_{\sqrt{2}\Lambda} f\|_{L^2(\mathbb{R}^d)} \leq \|1_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2(\mathbb{R}^d)}\, \mathrm{e}^{T_0 A_{f_0}(\alpha)},$$

where $A_{f_0}(\alpha) := \widetilde{C}_{f_0} + \frac{C_{f_0}\log\alpha}{2(\log(\mathrm{e}+\alpha))^{\mu+1}}$ depends on f_0 only through $\|f_0\|_{L^1}$, $\|f_0\|_{L^1_2}$ and $\|f_0\|_{L\log L}$.

Proof. Assume $\mathrm{Hyp}_\Lambda(M)$ is true. Since $|\eta^-| = |\eta|\sin\frac{\theta}{2} \leq \frac{|\eta|}{\sqrt{2}}$ by the assumption on the angular cross-section, the hypothesis implies

$$\sup_{|\eta|\leq\sqrt{2}\Lambda} G(\eta^-)^{\frac{\mu+1}{1+\log\alpha}}|\hat{f}(\eta^-)| \leq M.$$

With this uniform estimate at hand, we can bound the commutation error by

$$\begin{aligned}
&|\langle G_\Lambda Q(f,f) - Q(f, G_\Lambda f), G_\Lambda f\rangle| \\
&\leq 2\beta t(\mu+1)M \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1 - \frac{|\eta^+|^2}{|\eta|^2}\right)(\log\langle\eta\rangle_\alpha)^\mu \\
&\qquad\qquad \times G_{\sqrt{2}\Lambda}(\eta^+)|\hat{f}(\eta^+)|\, G_{\sqrt{2}\Lambda}(\eta)|\hat{f}(\eta)|\, \mathrm{d}\sigma\, \mathrm{d}\eta,
\end{aligned}$$

see equation (4.5). By Lemma 4.7, this can be further bounded by

$$\begin{aligned}&|\langle G_\Lambda Q(f,f) - Q(f, G_\Lambda f), G_\Lambda f\rangle| \\ &\quad\le \beta T_0(\mu+1)c_{b,d}M\left(\|G_{\sqrt{2}\Lambda} f\|_{L^2}^2 + \left\|(\log\langle D_v\rangle_\alpha)^{\frac{\mu}{2}} G_{\sqrt{2}\Lambda} f\right\|_{L^2}^2\right)\end{aligned}$$

for all $0 \le t \le T_0$. Thus, the a priori bound from Corollary 4.5 yields

$$\begin{aligned}\|G_{\sqrt{2}\Lambda} f\|_{L^2}^2 \le{}& \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2}^2 + 2\left(\widetilde{C}_{f_0} + \beta T_0(\mu+1)c_{b,d}M\right)\int_0^t \|G_{\sqrt{2}\Lambda} f\|_{L^2}^2 \,\mathrm{d}\tau \\ &+ 2\int_0^t \Bigg(\beta \left\|(\log\langle D_v\rangle_\alpha)^{\frac{\mu+1}{2}} G_{\sqrt{2}\Lambda} f\right\|_{L^2}^2 \\ &\qquad + \beta T_0(\mu+1)c_{b,d}M \left\|(\log\langle D_v\rangle_\alpha)^{\frac{\mu}{2}} G_{\sqrt{2}\Lambda} f\right\|_{L^2}^2 \\ &\qquad - \frac{C_{f_0}}{(\log(\mathrm{e}+\alpha))^{\mu+1}} \left\|(\log\langle D_v\rangle_\alpha)^{\frac{\mu+1}{2}} G_{\sqrt{2}\Lambda} f\right\|_{L^2}^2\Bigg)\,\mathrm{d}\tau \end{aligned} \tag{4.8}$$

Choosing $\beta \le \beta_0(\alpha)$ as defined in (4.7) ensures that the integrand in the last term on the right hand side of (4.8) is negative. Indeed, setting $B = T_0(\mu+1)c_{b,d}M$ and $C = \frac{C_{f_0}}{(\log(\mathrm{e}+\alpha))^{\mu+1}}$, so that $\beta \le \frac{C\log\alpha}{\log\alpha + 2B}$, one sees that

$$\begin{aligned}\beta\log\langle\eta\rangle_\alpha + \beta B - C\log\langle\eta\rangle_\alpha &\le -\frac{2CB}{\log\alpha+2B}\log\langle\eta\rangle_\alpha + \frac{CB\log\alpha}{\log\alpha+2B} \\ &= \frac{CB\left(\log\alpha - \log(\alpha+|\eta|^2)\right)}{\log\alpha+2B} \le 0,\end{aligned}$$

and further, since $\log\alpha \ge \mu > 0$,

$$\beta B \le \frac{CB\log\alpha}{\log\alpha+2B} = \frac{C\log\alpha}{2}\frac{2B}{\log\alpha+2B} \le \frac{C\log\alpha}{2}.$$

It follows that

$$\begin{aligned}\|G_{\sqrt{2}\Lambda} f\|_{L^2}^2 &\le \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2}^2 + 2\left(\widetilde{C}_{f_0} + \beta T_0(\mu+1)c_{b,d}M\right)\int_0^t \|G_{\sqrt{2}\Lambda} f\|_{L^2}^2 \,\mathrm{d}\tau \\ &\le \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2}^2 + 2A_{f_0}(\alpha)\int_0^t \|G_{\sqrt{2}\Lambda} f\|_{L^2}^2 \,\mathrm{d}\tau.\end{aligned}$$

Now Gronwall's lemma implies

$$\|G_{\sqrt{2}\Lambda} f\|_{L^2}^2 \le \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2}^2\,\mathrm{e}^{2A_{f_0}(\alpha)t} \le \|\mathbb{1}_{\sqrt{2}\Lambda}(D_v) f_0\|_{L^2}^2\,\mathrm{e}^{2A_{f_0}(\alpha)T_0}.$$

■

Lemma 4.10 (Step 2). *Let $\beta, \mu > 0$, $T_0 > 0$, and*

$$\Lambda \geq \Lambda_0 := \frac{2\sqrt{d}}{\sqrt{2}-1}.$$

If there exist finite constants $B_1, B_2 \geq 0$ such that for all $0 \leq t \leq T_0$

$$\|f(t,\cdot)\|_{L^1_1(\mathbb{R}^d)} \leq B_1, \quad \textit{and} \quad \|(G_{\sqrt{2}\Lambda} f)(t,\cdot)\|_{L^2(\mathbb{R}^d)} \leq B_2,$$

then there exists a constant K depending only on the dimension d and the bounds B_1, B_2, such that for all $|\eta| \leq \widetilde{\Lambda} := \frac{1+\sqrt{2}}{2}\Lambda$ and $t \in [0, T_0]$

$$|\hat{f}(t,\eta)| \leq K\, G(t,\eta)^{-\frac{2}{d+2}}.$$

Proof. By Remark 4.2 f satisfies the conditions of Lemma 4.3 with $\|\nabla \hat{f}\|_{L^\infty(\mathbb{R}^d)} \leq 2\pi B_1$, uniformly in $t \in [0, T_0]$. Obviously, also $\|\hat{f}\|_{L^\infty(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)} \leq B_1$. It follows that for any $\eta \in \mathbb{R}^d$

$$|\hat{f}(\eta)| \leq \big(2\pi(d+2)B_1\big)^{\frac{d}{d+2}} \left(\int_{Q_\eta} |\hat{f}|^2 \,\mathrm{d}\eta\right)^{\frac{1}{d+2}}.$$

where Q_η is a unit cube with one corner at η, such that $\eta \cdot (\zeta - \eta) \geq 0$ for all $\zeta \in Q_\eta$. Since its diameter is $\sqrt{d}$, the condition $\Lambda \geq \Lambda_0$ and the choice of $\widetilde{\Lambda}$ guarantee that for $|\eta| \leq \widetilde{\Lambda}$ the cube Q_η always stays inside a ball around the origin with radius $\sqrt{2}\Lambda$. By the orientation of Q_η and since the Fourier weight G is a radial and increasing function in η, we can further estimate

$$\begin{aligned}
|\hat{f}(\eta)| &\leq \big(2\pi(d+2)B_1\big)^{\frac{d}{d+2}}\; G(\eta)^{-\frac{2}{d+2}} \left(\int_{Q_\eta} G(\eta)^2 |\hat{f}|^2 \,\mathrm{d}\eta\right)^{\frac{1}{d+2}} \\
&\leq \big(2\pi(d+2)B_1\big)^{\frac{d}{d+2}}\; G(\eta)^{-\frac{2}{d+2}} \left\| G_{\sqrt{2}\Lambda} f \right\|_{L^2(\mathbb{R}^d)}^{\frac{2}{d+2}} \\
&\leq \left(2\pi(d+2)B_1 B_2^{\frac{2}{d}}\right)^{\frac{d}{d+2}}\; G(\eta)^{-\frac{2}{d+2}}
\end{aligned}$$

which is the claimed inequality with $K = \left(2\pi(d+2)B_1 B_2^{\frac{2}{d}}\right)^{\frac{d}{d+2}}$. ∎

Proof of Theorem 4.4. Let $\mu > 0$ and $T_0 > 0$ be fixed. Set $\alpha_* = \mathrm{e}^{\frac{d}{2}+\frac{d+2}{2}\mu} \geq \mathrm{e}^\mu$, which is chosen in such a way that $\frac{\mu+1}{1+\log\alpha_*} = \frac{2}{d+2}$ and the function $s \mapsto \big(\log(\alpha_* + s)\big)^{\mu+1}$ is concave.

Choosing $\Lambda_0 = \frac{2\sqrt{d}}{\sqrt{2}-1}$ as in Lemma 4.10, we define the length scales for our induction by

$$\Lambda_N := \frac{\Lambda_{N-1} + \sqrt{2}\Lambda_{N-1}}{2} = \frac{1+\sqrt{2}}{2}\Lambda_{N-1} = \left(\frac{1+\sqrt{2}}{2}\right)^N \Lambda_0, \quad N \in \mathbb{N}.$$

By conservation of energy, we have

$$\|f(t,\cdot)\|_{L^1_1} \le \|f(t,\cdot)\|_{L^1_2} = \|f_0\|_{L^1_2} =: B_1$$

in view of Lemma 4.10. By Lemma 4.9 a good (in particular uniform in $N \in \mathbb{N}$) choice for B_2 is

$$B_2 := \|f_0\|_{L^2(\mathbb{R}^d)} \, \mathrm{e}^{T_0 A_{f_0}(\alpha_*)}.$$

Define further

$$M := \max\left\{2B_1 + 1, \left(2\pi(d+2)B_1 B_2^{\frac{2}{d}}\right)^{\frac{d}{d+2}}\right\},$$

where the second expression is just the constant K from Lemma 4.10.

For the start of the induction, we need $\mathrm{Hyp}_{\Lambda_0}(M)$ to hold. Since

$$\sup_{t\in[0,T_0]} \sup_{|\eta|\le\Lambda_0} G(\eta)^{\frac{\mu+1}{1+\log\alpha_*}} |\hat{f}(\eta)| \le \mathrm{e}^{\frac{\mu+1}{1+\log\alpha_*}\beta T_0\left(\frac{1}{2}\log\left(\alpha_*+\Lambda_0^2\right)\right)^{\mu+1}} B_1,$$

there exists $\widetilde{\beta} > 0$ small enough, such that for the the above choice of M, $\mathrm{Hyp}_{\Lambda_0}(M)$ is true for all $0 < \beta \le \widetilde{\beta}$.

For the induction step, assume that $\mathrm{Hyp}_{\Lambda_N}(M)$ is true. Setting

$$\beta = \min\{\beta_0(\alpha_*), \widetilde{\beta}\}$$

with $\beta_0(\alpha)$ from Lemma 4.9, all the assumptions of Lemma 4.9 are fulfilled and it follows that

$$\|G_{\sqrt{2}\Lambda_N} f\|_{L^2(\mathbb{R}^d)} \le \|\mathbb{1}_{\sqrt{2}\Lambda_N}(D_v) f\|_{L^2(\mathbb{R}^d)} \, \mathrm{e}^{T_0 A_{f_0}(\alpha_*)} \le B_2.$$

Notice that the right hand side of this inequality *does not depend on M*. Lemma 4.10 now implies that for all $|\eta| \le \widetilde{\Lambda_N} = \Lambda_{N+1}$

$$G(t,\eta)^{\frac{2}{d+2}} |\hat{f}(t,\eta)| \le K \le M$$

for all $t \in [0, T_0]$. By the choice of α_* this means that $\mathrm{Hyp}_{\Lambda_{N+1}}(M)$ is true.

By induction, it follows that $\mathrm{Hyp}_{\Lambda_N}(M)$ holds for all $N \in \mathbb{N}$, in particular

$$\sup_{t\in[0,T_0]} \sup_{\eta\in\mathbb{R}^d} \mathrm{e}^{\beta t\left(\log\langle\eta\rangle_{\alpha_*}\right)^{\mu+1}} |\hat{f}(\eta)| \le M.$$

Another application of Lemma 4.9 implies

$$\|G_{\sqrt{2}\Lambda_N} f\|_{L^2(\mathbb{R}^d)} \leq \|f_0\|_{L^2(\mathbb{R}^d)}\, \mathrm{e}^{T_0 A_{f_0}(\alpha_*)} \quad \text{for all } N \in \mathbb{N}.$$

Passing to the limit $N \to \infty$, it follows that $\|Gf\|_{L^2(\mathbb{R}^d)} \leq B_2$, that is,

$$\mathrm{e}^{\beta t \left(\log\langle D_v\rangle_{\alpha_*}\right)^{\mu+1}} f(t,\cdot) \in L^2(\mathbb{R}^d).$$

■

4.5 Smoothing effect for arbitrary physical initial data

Proof of Theorem 2.10. Let $T > 0$ be arbitrary (but finite). By the already known H^∞ smoothing property of the homogeneous Boltzmann equation for Maxwellian molecules with Debye-Yukawa type interaction, see Proposition 2.7, for any $0 < t_0 < T$ one has

$$f \in L^\infty([t_0, T]; H^\infty(\mathbb{R}^d)),$$

in particular $f(t,\cdot) \in L^2(\mathbb{R}^d)$ for all $t_0 \leq t \leq T$. Using $f(t_0,\cdot) \in L^1_2 \cap L\log L \cap L^2$ as new initial datum, Theorem 4.4 implies that there exist $\beta, M > 0$ such that

$$\mathrm{e}^{\beta t \left(\log\langle D_v\rangle_{\alpha_*}\right)^{\mu+1}} f(t,\cdot) \in L^2(\mathbb{R}^d)$$

and

$$\|\mathrm{e}^{\beta t \left(\log\langle\cdot\rangle_{\alpha_*}\right)^{\mu+1}} \hat{f}(t,\cdot)\|_{L^\infty(\mathbb{R}^d)} \leq M$$

for all $t \in [t_0, T]$. By the characterisation of the spaces $\mathcal{A}^\mu$ (see Appendix D), and since t_0 and T are arbitrary, it follows that $f(t,\cdot) \in \mathcal{A}^\mu(\mathbb{R}^d)$ for all $t > 0$. ■

Sketch of the Proof of Corollary 2.11. In the notation of [DFT09], basically, the only thing that needs to be checked is that the function $\psi_\alpha : [0,\infty) \to [0,\infty)$, $r \mapsto \psi_\alpha(r) := (\log\sqrt{\alpha + r})^{\mu+1}$ satisfies

(i) $\psi_\alpha(r) \to \infty$ for $r \to \infty$

(ii) $\psi_\alpha(r) \leq r$ for r large enough

(iii) there exists $R \geq 1$ such that for all $0 \leq \lambda \leq 1$

$$\psi_\alpha(\lambda^2|\eta|^2) \geq \lambda^2 \psi_\alpha(|\eta|^2) \quad \text{whenever} \quad \lambda|\eta| \geq R.$$

Property (iii) is fulfilled by any concave function ψ with $\psi(0) \geq 0$. This clearly is the case for ψ_α if $\alpha \geq \mathrm{e}^\mu$, see Lemma 4.1.

So we take the α from Theorem 4.4 and conclude propagation with Theorem 1.2 from [DFT09]. ■

CHAPTER 5

Outlook: Non-Maxwellian Molecules and Inhomogeneous Boltzmann Equation

We close the part on smoothing properties of the Boltzmann equation with a short outlook on the non-Maxwellian and inhomogeneous cases.

5.1 Non-Maxwellian Molecules

Recall that in general the Boltzmann collision kernel B is a function of the relative velocity $|v - v_*|$ and the collision angle θ defined (in the σ-representation) by $\cos\theta = \frac{v-v_*}{|v-v_*|} \cdot \sigma$. Let us assume that this dependence factorises into a *kinetic* factor $\Phi(|v - v_*|)$ and an *angular* contribution $b(\cos\theta)$, such that the Boltzmann kernel takes the form

$$B(|v - v_*|, \cos\theta) = \Phi(|v - v_*|)\, b(\cos\theta).$$

We will mainly have the case of inverse power law interaction potentials in mind, where $\Phi(|v|) = |v|^\gamma$ for some $\gamma \in \mathbb{R}$, and the angular kernel has a singularity of the type

$$\sin^{d-2}\theta\, b(\cos\theta) \sim \frac{\kappa}{\theta^{1+2\nu}}$$

for grazing collisions $\theta \to 0$, with $\kappa > 0$ and $0 < \nu < 1$. It is customary to distinguish the cases

(i) $\gamma < -2$: very soft potentials,

(ii) $-2 < \gamma < 0$: moderately soft potentials,

(iii) $\gamma = 0$: Maxwellian molecules,

(iv) $\gamma > 0$: hard potentials,

where the Boltzmann equation behaves rather differently.[1]

Indeed, already the existence of weak solutions is a challenging problem, and for very soft potentials (due to the vanishing of $\varphi' + \varphi'_* - \varphi - \varphi_*$ only of order 2 for test functions $\varphi \in W^{2,\infty}$) one has to use an altogether different definition of weak solutions: the notion of H-solutions based on the finiteness of the entropy production functional, which gives a useful a priori estimate and allows for a definition of $\langle Q(f, f), \varphi\rangle$ even in the case of very soft potentials, at least for $\gamma > -4$, see [Vil98]. We will not discuss this case any further here, but state the relevant existence results for moderately soft and hard potentials.

Theorem 5.1 (Theorem 1 in [Vil98])**.** *Let $f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$. If $\gamma > 0$, assume in addition that $f_0 \in L^1_{2+\delta}(\mathbb{R}^d)$ for some $\delta > 0$. If $-2 \le \gamma < 2$ and $\int \sin^d \theta\, b(\cos\theta)\, \mathrm{d}\theta < \infty$, there exists a weak solution f of the Cauchy problem for the Boltzmann equation with initial datum f_0. Moreover, for all $\varphi \in W^{2,\infty}(\mathbb{R}^d)$, the map $t \mapsto \int f(t)\varphi\, \mathrm{d}v$ is Lipschitz continuous if $f_0 \in L^1_{\max\{2,2+\gamma\}}$ or in any case for $t \ge t_0 > 0$.*

The latter statement refers to the fact that for hard potentials $\gamma > 0$ polynomial moments of arbitrary order are immediately generated, i.e., if $f_0 \in L^1_2(\mathbb{R}^d)$, then for all $q > 2$ and $t_0 > 0$ there exists a finite constant $C > 0$ such that for all $t \ge t_0$

$$\int_{\mathbb{R}^d} f(t, v)\, \langle v\rangle^q\, \mathrm{d}v \le C,$$

see, for instance, [Elm83, Des93, Wen97, MW99] for details.

The following coercivity property is known in the non-Maxwellian case:

Theorem 5.2 (Proposition 2.1 in [AMUXY12])**.** *Let $0 \le g \in L^1_2 \cap L\log L(\mathbb{R}^d)$ with $\|g\|_{L^1} \ge D_0$ and $\|g\|_{L^1_2} + \|g\|_{L\log L} \le E_0$ for some $D_0, E_0 > 0$. If $\gamma + 2\nu > 0$, there exist positive constants c_0, C depending only on D_0 and E_0 such that*

$$-\langle Q(g, f), f\rangle \ge c_0 \left\|\langle\cdot\rangle^{\gamma/2} f\right\|^2_{H^\nu} - C\left\|\langle\cdot\rangle^{\gamma/2} f\right\|^2_{L^2}$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$.

Using an approach based on Fourier methods for non-Maxwellian collision kernels is considerably harder since Bobylev's identity has a much more complicated form. One can show that

$$\widehat{Q(g, f)}(\eta) = \int_{\mathbb{R}^d\times\mathbb{S}^{d-1}} \widehat{\Phi}(|\xi|)\, b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left[\widehat{g}(\eta^- + \xi)\widehat{f}(\eta^+ - \xi) - \widehat{g}(\xi)\widehat{f}(\eta - \xi)\right] \mathrm{d}\sigma\, \mathrm{d}\xi, \tag{5.1}$$

[1] Of course, γ should not become too negative, otherwise the singularity would be too strong. From our derivation in the introduction, $\gamma = \frac{n-(2d-1)}{n-1}$ with $n > 2$, so certainly $\gamma > 3 - 2d$.

where $\eta^{\pm} = \frac{1}{2}(\eta \pm |\eta|\sigma)$ as before. For $\Phi \equiv 1$ this reduces to the Bobylev identity for Maxwellian molecules, see (1.40).

In the most interesting case $\Phi(|v|) = |v|^{\gamma}$, say for hard potentials $0 < \gamma < 2$, the Fourier transform $\widehat{\Phi}$ only exists in a distributional sense, so one has to take care of the well-definedness of equation (5.1) under the singularity assumptions on the angular kernel b and the natural assumptions on f and g. This is where creation of moments can help, as this implies that the Fourier transform $\hat{f}(t,\cdot)$ of the weak solution is immediately a bounded $\mathscr{C}^{\infty}$ function.

The main difficulty in applying our strategy from the Maxwellian molecules case, however, is the fact that the quasi-locality property of the Boltzmann collision operator is lost, so that the inductive scheme we developed fails to work.

The following smoothing results are known for non-Maxwellian molecules (for the precise assumptions in each case we refer to the articles):

1. Using Littlewood-Paley theory, Alexandre and ElSafadi [AE09] were able to prove H^{∞} smoothing of weak solutions for a regularised kinetic factor $\Phi(v) = \langle v\rangle^{\gamma}$.

2. Alexandre, Morimoto, Ukai, Xu, Yang [AMUXY12] proved H^{∞} smoothing of weak solutions for the physically relevant kinetic factor $\Phi(v) = |v|^{\gamma}$.

3. Chen and He [CH11] showed H^{∞} smoothing, again in the physically relevant case $\Phi(v) = |v|^{\gamma}$ for the strong solutions constructed by Desvillettes and Mouhot [DM09].

5.2 Inhomogeneous Boltzmann Equation

In the spatially inhomogeneous case, the collision operator is highly degenerate, since it only acts on the velocity variable. Due to the presence of the transport term $v \cdot \nabla_x$, one expects a transfer of regularity from the velocity variable to the space variable, and therefore some hypoelliptic smoothing effect in *both* variables. This has been highlighted in terms of a generalised uncertainty principle for kinetic equations by Alexandre, Morimoto, Ukai, Xu and Yang [AMUXY08] under strong assumptions on the initial data and the solutions. Chen and He [CH12] were able to show an H^{∞} smoothing effect for the global classical solutions obtained by Gressmann and Strain [GS10, GS11], and Alexandre, Morimoto, Ukai, Xu, and Yang [AMUXY11a, AMUXY11b, AMUXY11c]. For the one-dimensional inhomogeneous Kac equation, Lerner, Morimoto, Pravda-Starov and Xu [LMPX15] obtained Gelfand-Shilov smoothing with respect to the velocity variable and Gevrey smoothing with respect to the space variable for *fluctuations around the global equilibrium*, i.e., close to the equilibrium.

As in the homogeneous case, replacing the bilinear Boltzmann operator by a fractional Laplacian (for power-law interactions), one is led to study the generalised Kolmogorov equation

$$\partial_t f + v \cdot \nabla_x f = -(-\Delta_v)^{\nu} f, \quad (x,v) \in \mathbb{R}^{2d}, \quad t > 0, \tag{5.2}$$

for instance with initial datum $f|_{t=0} = f_0 \in L^2(\mathbb{R}^{2d})$, for some $\nu > 0$. The Fourier transform $\widehat{f}(t,\xi,\eta) = \int_{\mathbb{R}^{2d}} \mathrm{e}^{-2\pi \mathrm{i}(x\cdot\xi + v\cdot\eta)} f(t,x,v)\,\mathrm{d}x\mathrm{d}v$ of f is then a solution of the inhomogeneous transport equation

$$\partial_t \widehat{f}(t,\xi,\eta) - \xi \cdot \nabla_\eta \widehat{f}(t,\xi,\eta) = -(2\pi)^{2\nu} |\eta|^{2\nu} \widehat{f}(t,\xi,\eta), \tag{5.3}$$

for $(\xi,\eta) \in \mathbb{R}^{2d}$, $t > 0$, which can be solved easily by the method of characteristics: the characteristic curve Ξ through (t,ξ,η) satisfies the ODE $\dot{\Xi}(s) = -\xi$, $\Xi(t) = \eta$, with solution $\Xi(s) = \eta + (t-s)\xi$. Setting $F_\xi(s) := \widehat{f}(s,\xi,\Xi(s))$, the function F_ξ solves the differential equation

$$\dot{F}_\xi(s) = -(2\pi)^{2\nu} |\Xi(s)|^{2\nu} F_\xi(s), \quad \text{with} \quad F_\xi(0) = \widehat{f}(0,\xi,\Xi(0)) = \widehat{f}_0(\xi, \eta + t\xi),$$

i.e., $F_\xi(s) = \mathrm{e}^{-(2\pi)^{2\nu} \int_0^s |\Xi(\tau)|^{2\nu}\,\mathrm{d}\tau} F_\xi(0)$. Hence, the solution $\widehat{f}(t,\xi,\eta) = F_\xi(t)$ of (5.3), that is, the Fourier transform of the solution of the generalised Kolmogorov equation (5.2), is given by

$$\widehat{f}(t,\xi,\eta) = \mathrm{e}^{-(2\pi)^{2\nu} \int_0^t |\eta + \tau\xi|^{2\nu}\,\mathrm{d}\tau} \widehat{f}_0(\xi, \eta + t\xi).$$

Since for any $\nu > 0$ there exists a constant $c_\nu > 0$ such that

$$\int_0^t |\eta + s\xi|^{2\nu}\,\mathrm{d}s \geq c_\nu \left(t|\eta|^{2\nu} + t^{2\nu+1}|\xi|^{2\nu} \right),$$

see, for instance, [MX09, Lemma 3.1], the solution $f(t,\cdot,\cdot)$ of (5.2) immediately becomes Gevrey regular of order $\frac{1}{2\nu}$ in *both* the x and v variables,

$$\mathrm{e}^{c_\nu \left(t(-\Delta_v)^\nu + t^{2\nu+1}(-\Delta_x)^\nu \right)} f(t,\cdot,\cdot) \in L^2(\mathbb{R}^{2d}),$$

i.e. $f(t,\cdot,\cdot) \in \mathcal{G}^{\frac{1}{2\nu}}(\mathbb{R}^{2d})$ for any $t > 0$. It is therefore natural to expect the

Conjecture. *Let f be a solution of the inhomogeneous Boltzmann equation with initial datum f_0 satisfying*

$$\int_{\mathbb{R}^{2d}} f_0(x,v) \left(1 + |v|^2 + \log(1 + f_0(x,v)) \right) \,\mathrm{d}x\,\mathrm{d}v < \infty.$$

Then $f(t,\cdot,\cdot) \in \mathcal{G}^{\frac{1}{2\nu}}(\mathbb{R}^{2d})$ for any positive time $t > 0$.

APPENDIX A

L^2-Reformulation of the homogeneous Boltzmann equation for weak solutions and coercivity

A reformulation of the weak form (1.25) of the Boltzmann and Kac equations is derived. We want to choose a suitable test function φ in terms of the weak solution f itself in the weak formulation of the Cauchy problem (1.2). We use $\varphi(t,\cdot) := G_\Lambda^2(t,D_v)f(t,\cdot)$ and since this involves a hard cut-off in Fourier space, we automatically have high regularity of $\varphi(t,v)$ in the velocity variable, the question is to have $\mathscr{C}^1$ regularity in the time variable. For this we follow the strategy by Morimoto *et al.* [MUXY09].

Proposition A.1. *Let f be a weak solution of the Cauchy problem (1.2) with initial datum $0 \le f_0 \in L^1_2 \cap L\log L(\mathbb{R}^d)$, and let $T_0 > 0$. Then for all $t \in (0,T_0]$, $\beta > 0$, $\alpha \in (0,1)$, and $\Lambda > 0$ we have $G_\Lambda f \in \mathscr{C}\left([0,T_0]; L^2(\mathbb{R}^d)\right)$ and*

$$\begin{aligned}&\frac{1}{2}\|G_\Lambda(t,D_v)f(t,\cdot)\|^2_{L^2(\mathbb{R}^d)} - \frac{1}{2}\int_0^t \left\langle f(\tau,\cdot), \left(\partial_t G_\Lambda^2(\tau,D_v)\right) f(\tau,\cdot)\right\rangle \,\mathrm{d}\tau \\ &= \frac{1}{2}\|\mathbb{1}_\Lambda(D_v)f_0\|^2_{L^2(\mathbb{R}^d)} + \int_0^t \left\langle Q(f,f)(\tau,\cdot), G_\Lambda^2(\tau,D_v)f(\tau,\cdot)\right\rangle \,\mathrm{d}\tau.\end{aligned} \tag{A.1}$$

To ensure that we can use $G_\Lambda^2 f$ as a test function in the weak formulation of the Boltzmann equation, we need the following bilinear estimate on $Q(g,f)$, which is a special case of a larger class of functional inequalities by Alexandre [Ale06, Ale09, AH08].

Lemma A.2 (Functional Estimate on Collision Operator)**.** *Assume that the angular collision cross-section b satisfies assumptions* **(B1)–(B3)** *or* **(K1)–(K3)**, *respectively. Then* *for any $k > \frac{d+4}{2}$ there exists a constant $C > 0$ such that*

$$\|Q(g,f)\|_{H^{-k}(\mathbb{R}^d)} \le C\|g\|_{L^1_2(\mathbb{R}^d)}\|f\|_{L^1_2(\mathbb{R}^d)}. \tag{A.2}$$

Proof. This is a direct consequence[1] of Theorem 7.4 in Alexandre's review [Ale09]: under the assumptions on b, for any $m \in \mathbb{R}$ there exists a constant $\widetilde{C} > 0$ such that

$$\|Q(g,f)\|_{H^{-m}(\mathbb{R}^d)} \le \widetilde{C}\|g\|_{L^1_{2\nu}(\mathbb{R}^d)}\|f\|_{H^{-m+2\nu}_{2\nu}(\mathbb{R}^d)}.$$

Since $L^1(\mathbb{R}^d) \subset H^{-s}(\mathbb{R}^d)$ for any $s > \frac{d}{2}$, we obtain for $k > \frac{d+4}{2}$ and $\nu \in (0,1)$,

$$\begin{aligned}\|f\|_{H^{-k+2\nu}_{2\nu}(\mathbb{R}^d)} &= \|\langle\cdot\rangle^{2\nu} f\|_{H^{-k+2\nu}(\mathbb{R}^d)} \le C\|\langle\cdot\rangle^{2\nu} f\|_{L^1(\mathbb{R}^d)} \\ &\le c\|\langle\cdot\rangle^{2} f\|_{L^1(\mathbb{R}^d)} = c\|f\|_{L^1_2(\mathbb{R}^d)},\end{aligned}$$

i.e., $L^1_2(\mathbb{R}^d) \subset H^{-k+2\nu}_{2\nu}(\mathbb{R}^d)$ for any $k > \frac{d+4}{2}$ and $\nu \in (0,1)$. Therefore,

$$\|Q(g,f)\|_{H^{-k}(\mathbb{R}^d)} \le \widetilde{C}\|g\|_{L^1_{2\nu}(\mathbb{R}^d)}\|f\|_{H^{-k+2\nu}_{2\nu}(\mathbb{R}^d)} \le C\|g\|_{L^1_2(\mathbb{R}^d)}\|f\|_{L^1_2(\mathbb{R}^d)}.$$

∎

Lemma A.2 implies that for $f,g \in L^1_2(\mathbb{R}^d)$, $\langle Q(g,f),h\rangle$ is well-defined for all $h \in H^k(\mathbb{R}^d)$, $k > \frac{d+4}{2}$, and one has $\langle Q(g,f),h\rangle = \langle \widehat{Q(g,f)}, \widehat{h}\rangle_{L^2}$.

Proof of Proposition A.1. Choosing a constant in time test function $\varphi(t,\cdot) = \psi \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ in the weak formulation (1.25) yields

$$\int_{\mathbb{R}^d} f(t,v)\psi(v)\,\mathrm{d}v - \int_{\mathbb{R}^d} f(s,v)\psi(v)\,\mathrm{d}v = \int_s^t \langle Q(f,f)(\tau,\cdot),\psi\rangle\,\mathrm{d}\tau,$$

for all $0 \le s \le t \le T_0$ and $\psi \in \mathscr{C}_0^\infty(\mathbb{R}^d)$ (this was already remarked by Villani [Vil98] as an equivalent formulation of (1.25)). By means of (A.2) this equality can be extended to test functions $\psi \in H^k$ for $k > \frac{d+4}{2}$, in particular one can choose $\psi = G^2_\Lambda f(t,\cdot)$ and $\psi = G^2_\Lambda f(s,\cdot)$ which, taking the sum of both resulting equations, yields

$$\begin{aligned}&\|G_\Lambda f(t,\cdot)\|^2_{L^2(\mathbb{R}^d)} - \|G_\Lambda f(s,\cdot)\|^2_{L^2(\mathbb{R}^d)} = \big\langle f(t,\cdot), G^2_\Lambda f(t,\cdot)\big\rangle - \big\langle f(s,\cdot), G^2_\Lambda f(s,\cdot)\big\rangle \\ &= \big\langle f(t,\cdot), \big(G^2_\Lambda(t,D_v) - G^2_\Lambda(s,D_v)\big)\, f(s,\cdot)\big\rangle \\ &\qquad + \int_s^t \big\langle Q(f,f)(\tau,\cdot), G^2_\Lambda f(t,\cdot) + G^2_\Lambda f(s,\cdot)\big\rangle\,\mathrm{d}\tau.\end{aligned} \tag{A.3}$$

[1] This result is proved in [Ale09] for $d = 3$, but the proof depends only on assumption (1.18) and general properties of Littlewood-Paley decompositions and holds in any dimension $d \ge 1$.

Using Plancherel, the first term on the right hand side of (A.3) can be estimated by

$$
\begin{aligned}
&\left|\left\langle f(t,\cdot), \left(G_\Lambda^2(t,D_v) - G_\Lambda^2(s,D_v)\right) f(s,\cdot)\right\rangle\right| = \left|\left\langle \hat f(t,\cdot), \left(G_\Lambda^2(t,\cdot) - G_\Lambda^2(s,\cdot)\right) \hat f(s,\cdot)\right\rangle\right| \\
&\quad\le \int_{\mathbb{R}^d} |\hat f(t,\eta)|\, |G_\Lambda^2(t,\eta) - G_\Lambda^2(s,\eta)|\, |\hat f(s,\eta)| \,\mathrm{d}\eta \\
&\quad\le |t-s| \int_{\mathbb{R}^d} 2\beta \langle \eta\rangle^{2\alpha} G_\Lambda^2(t,\eta)\,\mathrm{d}\eta\, \|f(t,\cdot)\|_{L^1(\mathbb{R}^d)} \|f(s,\cdot)\|_{L^1(\mathbb{R}^d)} \\
&\quad\le C_{\Lambda,T_0} |t-s|\, \|f_0\|^2_{L^1(\mathbb{R}^d)},
\end{aligned}
$$

and, using that the terms involving the collision operator can, for any $k > \frac{d+4}{2}$ (compare (A.2)), be bounded by

$$
\begin{aligned}
|\langle Q(f,f)(\tau,\cdot), G_\Lambda^2 f(t,\cdot)\rangle| &\le \|Q(f,f)(\tau,\cdot)\|_{H^{-k}(\mathbb{R}^d)} \|G_\Lambda^2 f(t,\cdot)\|_{H^k(\mathbb{R}^d)} \\
&\le C \|f\|^2_{L^1_2(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} \langle \eta\rangle^{2k} G_\Lambda^4(t,\eta) |\hat f(t,\eta)|^2 \,\mathrm{d}\eta \right)^{1/2} \\
&\le C \|f\|^2_{L^1_2(\mathbb{R}^d)} \|f(t,\cdot)\|_{L^1(\mathbb{R}^d)} \left(\int_{\mathbb{R}^d} \langle \eta\rangle^{2k} G_\Lambda^4(T_0,\eta) \,\mathrm{d}\eta \right)^{1/2} \\
&\le C'_{\Lambda,T_0} \|f_0\|^2_{L^1_2(\mathbb{R}^d)} \|f_0\|_{L^1(\mathbb{R}^d)}
\end{aligned}
$$

for any $t \in [0, T_0]$, yields

$$
\left| \int_s^t \langle Q(f,f)(\tau,\cdot), G_\Lambda^2 f(t,\cdot) + G_\Lambda^2 f(s,\cdot)\rangle \,\mathrm{d}\tau \right| \le 2C'_{\Lambda,T_0} |t-s|\, \|f_0\|^2_{L^1_2(\mathbb{R}^d)} \|f_0\|_{L^1(\mathbb{R}^d)}.
$$

Plugging the latter two bounds into (A.3) shows that $G_\Lambda f \in \mathscr{C}([0,T_0]; L^2(\mathbb{R}^d))$, in fact, the map $[0,T_0] \ni t \mapsto \|G_\Lambda f(t,\cdot)\|_{L^2(\mathbb{R}^d)}$ is even Lipschitz continuous.

For any test function $\varphi \in \mathscr{C}^1(\mathbb{R}^+; \mathscr{C}_0^\infty(\mathbb{R}^d))$ the term involving the partial derivative $\partial_t \varphi$ in the weak formulation (1.25) can be rewritten as

$$
\int_0^t \left\langle f(\tau,\cdot), \partial_\tau \varphi(\tau,\cdot)\right\rangle \,\mathrm{d}\tau = \lim_{h\to 0} \int_0^t \left\langle f(\tau,\cdot) + f(\tau+h,\cdot), \frac{\varphi(\tau+h,\cdot) - \varphi(\tau,\cdot)}{2h} \right\rangle \,\mathrm{d}\tau,
$$

since $f \in \mathscr{C}(\mathbb{R}^+; \mathscr{D}'(\mathbb{R}^d))$. The integral on the right hand side is well-defined even for $\varphi \in L^\infty([0,T_0]; W^{2,\infty}(\mathbb{R}^d))$, in particular for $\varphi = G_\Lambda^2 f$, yielding

$$
\begin{aligned}
&\int_0^t \left\langle f(\tau,\cdot) + f(\tau+h,\cdot), \frac{\varphi(\tau+h,\cdot) - \varphi(\tau,\cdot)}{2h} \right\rangle \,\mathrm{d}\tau \\
&= \int_0^t \left\langle f(\tau,\cdot) + f(\tau+h,\cdot), \frac{G_\Lambda^2 f(\tau+h,\cdot) - G_\Lambda^2 f(\tau,\cdot)}{2h} \right\rangle \,\mathrm{d}\tau \\
&= \frac{1}{2h} \int_0^t \left(\|G_\Lambda f(\tau+h,\cdot)\|^2_{L^2} - \|G_\Lambda f(\tau,\cdot)\|^2_{L^2} \right) \,\mathrm{d}\tau \\
&\quad + \int_0^t \left\langle f(\tau,\cdot), \frac{G_\Lambda^2(\tau+h,D_v) - G_\Lambda^2(\tau,D_v)}{2h} f(\tau+h,\cdot) \right\rangle \,\mathrm{d}\tau.
\end{aligned}
$$

Using $G_\Lambda f \in \mathscr{C}([0,T_0];L^2(\mathbb{R}^d))$ it follows that

$$
\begin{aligned}
&\frac{1}{2h}\int_0^t \left(\|G_\Lambda f(\tau+h,\cdot)\|^2_{L^2(\mathbb{R}^d)} - \|G_\Lambda f(\tau,\cdot)\|^2_{L^2(\mathbb{R}^d)} \right) \mathrm{d}\tau \\
&= \frac{1}{2h}\int_t^{t+h} \|G_\Lambda f(\tau,\cdot)\|^2_{L^2(\mathbb{R}^d)} \,\mathrm{d}\tau - \frac{1}{2h}\int_0^h \|G_\Lambda f(\tau,\cdot)\|^2_{L^2(\mathbb{R}^d)} \,\mathrm{d}\tau \\
&\xrightarrow{h\to 0} \frac{1}{2}\|G_\Lambda f(t,\cdot)\|^2_{L^2(\mathbb{R}^d)} - \frac{1}{2}\|G_\Lambda f(0,\cdot)\|^2_{L^2(\mathbb{R}^d)}.
\end{aligned}
$$

where $\|G_\Lambda f(0,\cdot)\|_{L^2(\mathbb{R}^d)} = \|\mathbb{1}_\Lambda(D_v) f_0\|_{L^2(\mathbb{R}^d)}$. For the second integral, an application of dominated convergence gives

$$
\begin{aligned}
&\lim_{h\to 0}\int_0^t \left\langle f(\tau,\cdot), \frac{G^2_\Lambda(\tau+h,D_v) - G^2_\Lambda(\tau,D_v)}{2h} f(\tau+h,\cdot) \right\rangle \mathrm{d}\tau \\
&\quad = \frac{1}{2}\int_0^t \left\langle f(\tau,\cdot), \left(\partial_\tau G^2_\Lambda\right)(\tau,D_v) f(\tau,\cdot) \right\rangle \mathrm{d}\tau.
\end{aligned}
$$

Putting everything together, we thus have proved equation (A.1), i.e.

$$
\begin{aligned}
\frac{1}{2}\|G_\Lambda f\|^2_{L^2(\mathbb{R}^d)} = \frac{1}{2}\|\mathbb{1}_\Lambda(D_v) f_0\|^2_{L^2(\mathbb{R}^d)} &+ \frac{1}{2}\int_0^t \left\langle f(\tau,\cdot), \left(\partial_\tau G^2_\Lambda\right)(\tau,D_v) f(\tau,\cdot) \right\rangle \mathrm{d}\tau \\
&+ \int_0^t \left\langle Q(f,f), G^2_\Lambda f \right\rangle \mathrm{d}\tau.
\end{aligned}
$$

■

APPENDIX B

H^∞ smoothing of the Boltzmann and Kac equations

We follow the strategy as in our proof of Gevrey regularity, with several simplifications. Of course, we *do not* assume that f_0 is square integrable! We have

Theorem B.1 (H^∞ smoothing for the homogeneous Boltzmann and Kac equation)**.** *Assume that the cross-section b satisfies* **(B1)–(B3)** *for $d \geq 2$, respectively* **(K1)–(K3)** *for $d = 1$. Let f be a weak solution of the Cauchy problem* (1.2) *with initial datum satisfying conditions* (1.23)*. Then*

$$f(t, \cdot) \in H^\infty(\mathbb{R}^d) \tag{B.1}$$

for all $t > 0$.

The proof is known, at least for the three dimensional Boltzmann equation see [MUXY09], we give a proof for the convenience of the reader.

Remark. Theorem B.1 applies both to the power-law type (1.18) and Debye-Yukawa type (1.19) singularities. For simplicity, we only do the proof for power-law singularities here. The proof for the latter type of singularities is completely analogous. ♦

Again, one has to use suitable time-dependent Fourier multipliers. Note that for $f_0 \in L^1(\mathbb{R}^d)$ one has

$$\|f_0\|_{H^{-\gamma}(\mathbb{R}^d)} \leq C_{d,\gamma} \|f_0\|_{L^1(\mathbb{R}^d)}$$

with $C_{d,\gamma} = \left(\int_{\mathbb{R}^d} \langle\eta\rangle^{-\gamma}\, \mathrm{d}\eta\right)^{1/2}$ which is finite for all $\gamma > d/2$. We choose $\gamma = d$, for convenience, and

$$M_\Lambda(t, \eta) := \langle\eta\rangle^{-d} e^{\beta t \log\langle\eta\rangle} \mathbb{1}_\Lambda(|\eta|)$$

as a multiplier. Then

$$\sup_{\Lambda>0} \|M_\Lambda(0,D_v)f_0\|_{L^2(\mathbb{R}^d)} = \|M_\infty(0,\cdot)\hat{f}_0\|_{L^2(\mathbb{R}^d)} = \|f_0\|_{H^{-d}(\mathbb{R}^d)} \le C_{d,d}\|f_0\|_{L^1(\mathbb{R}^d)}$$

The proof of Proposition A.1 carries over and we have

$$\begin{aligned}\frac{1}{2}\|M_\Lambda(t,D_v)f(t,\cdot)\|^2_{L^2(\mathbb{R}^d)} - \frac{1}{2}\int_0^t \left\langle f(\tau,\cdot), \left(\partial_\tau M^2_\Lambda(\tau,D_v)\right) f(\tau,\cdot)\right\rangle \, d\tau \\ = \frac{1}{2}\|M_\Lambda(0,D_v)f_0\|^2_{L^2(\mathbb{R}^d)} + \int_0^t \left\langle Q(f,f)(\tau,\cdot), M^2_\Lambda(\tau,D_v)f(\tau,\cdot)\right\rangle \, d\tau.\end{aligned} \tag{B.2}$$

and as in the proof of Corollary 3.2, we have

$$\begin{aligned}\langle Q(f,f), M^2_\Lambda f\rangle &= \langle Q(f,M_\Lambda f), M_\Lambda f\rangle + \langle M_\Lambda Q(f,f) - Q(f,M_\Lambda f), M_\Lambda f\rangle \\ &\le -\widetilde{C}_{f_0}\|M_\Lambda f\|^2_{H^\nu} + C_{f_0}\|M_\Lambda f\|^2_{L^2} + \langle M_\Lambda Q(f,f) - Q(f,M_\Lambda f), M_\Lambda f\rangle\end{aligned} \tag{B.3}$$

The replacement of Proposition 3.5 is

Proposition B.2. *The commutation error is bounded by*

$$\left|\langle M_\Lambda Q(f,f) - Q(f,M_\Lambda f), M_\Lambda f\rangle\right| \le (1+2^{d-1})c_{b,d}\|f\|_{L^1}\left(\frac{d}{2} + \frac{\beta t}{2}2^{\beta t/2}\right)\|M_\Lambda f\|^2_{L^2} \tag{B.4}$$

with the constant $c_{b,d}$ from Lemma 3.14.

Remark. Of course, for any weak solution f of the Boltzmann and Kac equations,

$$\|f\|_{L^1} = \|f(t,\cdot)\|_{L^1} = \|f_0\|_{L^1}.$$

The fact that the commutator is bounded in terms of the L^2 norm of $M_\Lambda f$ makes the proof of H^∞ smoothing for the Boltzmann and Kac equations much simpler than the proof of Gevrey regularity. ♦

Proof. As in the proof of Proposition 3.5, Bobylev's formula shows

$$\begin{aligned}&|\langle M_\Lambda Q(f,f) - Q(f,M_\Lambda f), M_\Lambda f\rangle| \\ &\quad\le \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right) M_\Lambda(\eta)|\hat{f}(\eta)||\hat{f}(\eta^-)||\hat{f}(\eta^+)| \\ &\qquad\qquad \times |M_\Lambda(t,\eta) - M_\Lambda(t,\eta^+)| \, d\sigma \, d\eta \\ &\quad\le \|\hat{f}\|_{L^\infty}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right) M_\Lambda(\eta)|\hat{f}(\eta)||\hat{f}(\eta^+)| \\ &\qquad\qquad \times |M_\Lambda(t,\eta) - M_\Lambda(t,\eta^+)| \, d\sigma \, d\eta\end{aligned} \tag{B.5}$$

where, as before, $\eta^{\pm} = \frac{1}{2}(\eta \pm |\eta|\sigma)$. To bound $|M_\Lambda(\eta) - M_\Lambda(\eta^+)|$, we let $s := |\eta|^2$ and $s^+ = |\eta^+|^2$. Recall that $|\eta^+|^2 = \frac{|\eta|^2}{2}(1 + \frac{\eta}{|\eta|} \cdot \sigma)$ and

$$1 - \frac{s^+}{s} = 1 - \frac{|\eta^+|^2}{|\eta|^2} = \frac{1}{2}\left(1 - \frac{\eta}{|\eta|} \cdot \sigma\right)$$

Again, because of the support condition on the collision kernel $b(\cos\theta)$, we have $\frac{s}{2} \le s^+ \le s$. Set $\widetilde{M}(s) := (1+s)^{-d/2} e^{\frac{\beta t}{2}\log(1+s)}$. Then, for $|\eta| \le \Lambda$,

$$\begin{aligned} M_\Lambda(\eta) - M_\Lambda(\eta^+) &= \widetilde{M}(s) - \widetilde{M}(s^+) \\ &= (1+s)^{-d/2} e^{\frac{\beta t}{2}\log(1+s)} - (1+s^+)^{-d/2} e^{\frac{\beta t}{2}\log(1+s^+)} \\ &= (1+s)^{-d/2}\left(e^{\frac{\beta t}{2}\log(1+s)} - e^{\frac{\beta t}{2}\log(1+s^+)}\right) \\ &\qquad + \left((1+s)^{-d/2} - (1+s^+)^{-d/2}\right) e^{\frac{\beta t}{2}\log(1+s^+)}. \end{aligned} \tag{B.6}$$

Since $s \le 2s^+$, we have $(1+s^+)^{-1} \le 2(1+s)^{-1}$. Hence

$$\begin{aligned} \left|(1+s)^{-d/2} - (1+s^+)^{-d/2}\right| &= \frac{d}{2}\int_{s^+}^{s} (1+r)^{-d/2-1}\,\mathrm{d}r \le \frac{d}{2}(1+s^+)^{-d/2-1}(s-s^+) \\ &\le d(1+s^+)^{-d/2}\left(1 - \frac{s^+}{s}\right) \end{aligned}$$

In addition, $\log(1+s) \le \log(2(1+s^+)) = \log 2 + \log(1+s^+)$. So

$$\begin{aligned} \left|e^{\frac{\beta t}{2}\log(1+s)} - e^{\frac{\beta t}{2}\log(1+s^+)}\right| &\le \frac{\beta t}{2}\int_{s^+}^{s} \frac{1}{1+r} e^{\frac{\beta t}{2}\log(1+r)}\,\mathrm{d}r \\ \le \frac{\beta t}{2}\frac{s}{1+s^+} e^{\frac{\beta t}{2}\log(1+s)}\left(1 - \frac{s^+}{s}\right) &\le \beta t 2^{\frac{\beta t}{2}} e^{\frac{\beta t}{2}\log(1+s^+)}\left(1 - \frac{s^+}{s}\right). \end{aligned}$$

Also $\log(1+s) \le \log(2(1+s^+)) = \log 2 + \log(1+s^+)$. These bounds together with (B.6) show

$$\left|M_\Lambda(\eta) - M_\Lambda(\eta^+)\right| \le \left(d + \beta t\, 2^{\frac{\beta t}{2}}\right)\left(1 - \frac{|\eta^+|^2}{|\eta|^2}\right) M_\Lambda(\eta^+)$$

for all $|\eta| \le \Lambda$. Since the integration in (B.5) is only over $|\eta| \le \Lambda$, plugging this together with $\|\hat{f}\|_{L^\infty} \le \|f\|_{L^1}$ into (B.5) yields

$$\begin{aligned} &|\langle M_\Lambda Q(f,f) - Q(f, M_\Lambda f), M_\Lambda f\rangle| \\ &\quad \le \|f\|_{L^1}\left(d + \beta t\, 2^{\frac{\beta t}{2}}\right)\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} b\left(\frac{\eta}{|\eta|}\cdot\sigma\right)\left(1 - \frac{|\eta^+|^2}{|\eta|^2}\right) \\ &\qquad\qquad \times M_\Lambda(\eta)|\hat{f}(\eta)|\, M_\Lambda(\eta^+)|\hat{f}(\eta^+)|\,\mathrm{d}\sigma\,\mathrm{d}\eta. \end{aligned}$$

Noting again

$$M_\Lambda(\eta)|\hat{f}(\eta)|\, M_\Lambda(\eta^+)|\hat{f}(\eta^+)| \le \frac{1}{2}\left((M_\Lambda(\eta)|\hat{f}(\eta)|)^2 + (M_\Lambda(\eta^+)|\hat{f}(\eta^+)|)^2\right)$$

and performing the same change of variables for the integral containing η^+ as in the proof of Lemma 3.6 finishes the proof of equation (B.4). ∎

Now we can finish the

Proof of Theorem B.1. Using (B.2), (B.3), Proposition B.2, and

$$\partial_\tau M_\Lambda(\tau,\eta)^2 = 2\beta \log\langle\eta\rangle\, M_\Lambda(\tau,\eta)^2$$

one sees

$$\begin{aligned}\|M_\Lambda(t,D_v)f(t,\cdot)\|_{L^2}^2 &\le \|f_0\|_{H^{-d}}^2 + 2C_{f_0}\int_0^t \|M_\Lambda(\tau,D_v)f(\tau,\cdot)\|_{L^2}^2\,\mathrm{d}\tau \\ &+ \int_0^t \Big\langle M_\Lambda(\tau,D_v)f(\tau,\cdot), \Big(\beta\log\langle D_v\rangle - 2\widetilde{C}_{f_0}\langle D_v\rangle^{2\nu}\Big)M_\Lambda(\tau,D_v)f(\tau,\cdot)\Big\rangle\,\mathrm{d}\tau \\ &+ (1+2^{d-1})c_{b,d}\|f_0\|_{L^1}\int_0^t \left(\frac{d}{2}+\frac{\beta\tau}{2}2^{\frac{\beta\tau}{2}}\right)\|M_\Lambda(\tau,D_v)f(\tau,\cdot)\|_{L^2}^2.\end{aligned}$$

Setting

$$\begin{aligned}A(\beta,\tau) &:= \sup_{\eta\in\mathbb{R}^d}\Big(\beta\log\langle\eta\rangle - 2\widetilde{C}_{f_0}\langle\eta\rangle^{2\nu}\Big) + 2C_{f_0} \\ &\qquad + (1+2^{d-1})c_{b,d}\|f_0\|_{L^1}\left(\frac{d}{2}+\frac{\beta\tau}{2}2^{\frac{\beta\tau}{2}}\right) \\ &= \frac{\beta}{2\nu}\left[\log\left(\frac{\beta}{4\nu\widetilde{C}_{f_0}}\right)-1\right] + 2C_{f_0} + (1+2^{d-1})c_{b,d}\|f_0\|_{L^1}\left(\frac{d}{2}+\frac{\beta\tau}{2}2^{\frac{\beta\tau}{2}}\right)\end{aligned}$$

the above can be bounded by

$$\|M_\Lambda(t,D_v)f(t,\cdot)\|_{L^2}^2 \le \|f_0\|_{H^{-d}}^2 + \int_0^t A(\beta,\tau)\|M_\Lambda(\tau,D_v)f(\tau,\cdot)\|_{L^2}^2\,\mathrm{d}\tau$$

and from Gronwall's lemma we get

$$\|M_\Lambda(t,D_v)f(t,\cdot)\|_{L^2}^2 \le \|f_0\|_{H^{-d}}^2 \exp\left(\int_0^t A(\beta,\tau)\,\mathrm{d}\tau\right).$$

Letting $\Lambda\to\infty$ one sees

$$\|f(t,\cdot)\|_{H^{\beta t-d}}^2 = \|M_\infty(t,D_v)f(t,\cdot)\|_{L^2}^2 \le \|f_0\|_{H^{-d}}^2 \exp\left(\int_0^t A(\beta,\tau)\,\mathrm{d}\tau\right).$$

that is, $f(t,\cdot)\in H^{\beta t-d}(\mathbb{R}^d)$. Now let $\beta\to\infty$ to see that $f(t,\cdot)\in H^\infty(\mathbb{R}^d)$ for any $t>0$. ■

Remark. Setting $\beta = \frac{\gamma+d}{t}$, one sees that $\|f(t,\cdot)\|_{H^\gamma(\mathbb{R}^d)} \lesssim t^{-\frac{\gamma+d}{4\nu}}$, so the H^γ norms, in particular the L^2 norm, of $f(t,\cdot)$ blow up at most polynomially as $t\to 0$. ♦

APPENDIX C

The Kolmogorov-Landau inequality

In this section we give a short proof of

Lemma C.1 (Kolmogorov-Landau inequality on the unit interval)**.** *Let $m \geq 2$ be an integer. There exists a constant $C_m > 0$ such that for all $w \in W^{m,\infty}([0,1])$,*

$$\|w^{(k)}\|_{L^\infty([0,1])} \leq C_m \left(\frac{\|w\|_{L^\infty([0,1])}}{u^k} + u^{m-k} \|w^{(m)}\|_{L^\infty([0,1])} \right), \quad k = 1, \ldots, m-1,$$

for all $0 < u \leq 1$.

The following argument is in part borrowed from R. A. DeVore and G. G. Lorentz's book [DL93, pp.37–39].

Proof. Since $w \in W^{m,\infty}([0,1])$, it has absolutely continuous derivatives of order up to $m-1$ and essentially bounded m^{th} derivative.

Let $x \in [0, \frac{1}{2}]$ and $h \in [0, \frac{1}{2}]$. Then, by Taylor's theorem,

$$w(x+h) = w(x) + \sum_{j=1}^{m-1} \frac{h^j}{j!} w^{(j)}(x) + R_m(x,h)$$

with the remainder $R_m(x,h) = \int_0^h \frac{(h-t)^{m-1}}{(m-1)!} w^{(m)}(x+t)\,\mathrm{d}t$, which can be bounded by

$$|R_m(x,h)| \leq \|w^{(m)}\|_{L^\infty([0,1])} \int_0^h \frac{(h-t)^{m-1}}{(m-1)!}\,\mathrm{d}t = \frac{h^m}{m!} \|w^{(m)}\|_{L^\infty([0,1])}.$$

Choosing $m-1$ real numbers $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_{m-1} \leq 1$ we obtain for $h \in [0, \frac{1}{2}]$ the system of equations

$$\sum_{j=1}^{m-1} \lambda_s^j \frac{h^j}{j!} w^{(j)}(x) = w(x + \lambda_s h) - w(x) - R_m(x, \lambda_s h) \quad \text{for } s = 1, \cdots, m-1. \tag{C.1}$$

Setting

$$V = \begin{pmatrix} \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{m-1} \\ \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{m-1} \\ \vdots & & \ddots & \vdots \\ \lambda_{m-1} & \lambda_{m-1}^2 & \cdots & \lambda_{m-1}^{m-1} \end{pmatrix}, \quad \mathbf{w}(x) = \begin{pmatrix} hw'(x) \\ \frac{h^2}{2} w''(x) \\ \vdots \\ \frac{h^{m-1}}{(m-1)!} w^{(m-1)}(x) \end{pmatrix},$$

$$\mathbf{b}(x) = \begin{pmatrix} w(x + \lambda_1 h) - w(x) - R_m(x, \lambda_1 h) \\ w(x + \lambda_2 h) - w(x) - R_m(x, \lambda_2 h) \\ \vdots \\ w(x + \lambda_{m-1} h) - w(x) - R_m(x, \lambda_{m-1} h) \end{pmatrix},$$

we have $V\mathbf{w}(x) = \mathbf{b}(x)$. Since the Vandermonde determinant

$$\det V = \prod_{i=1}^{m-1} \lambda_i \prod_{1 \leq j < l \leq m-1} (\lambda_l - \lambda_j) \neq 0,$$

V is invertible and we obtain $\mathbf{w}(x) = V^{-1}\mathbf{b}(x)$ and therefore

$$\left| \frac{h^k}{k!} w^{(k)}(x) \right| \leq \|\mathbf{w}(x)\| \leq \|V^{-1}\| \, \|\mathbf{b}(x)\|, \tag{C.2}$$

where $\|\cdot\|$ is any norm on $\mathbb{R}^{m-1}$, respectively the induced operator norm on the space of $(m-1) \times (m-1)$ real matrices. Choosing for concreteness the ℓ^1 norm on $\mathbb{R}^{m-1}$, we have

$$\begin{aligned} \|\mathbf{b}(x)\| &= \sum_{s=1}^{m-1} |w(x + \lambda_s h) - w(x) - R_m(x, \lambda_s h)| \\ &\leq (m-1) \left(2\|w\|_{L^\infty([0,1])} + \frac{h^m}{m!} \|w^{(m)}\|_{L^\infty([0,1])} \right). \end{aligned}$$

While for our application the size of $\|V^{-1}\|$ is of no importance, one can even explicitly calculate it: The inverse of the Vandermonde matrix V is explicitly known (see for instance [Gau62]),

$$\left(V^{-1}\right)_{\alpha\beta} = (-1)^{\alpha-1} \frac{\sigma_{m-1-\alpha}^{\beta}}{\lambda_\beta \prod_{\nu \neq \beta} (\lambda_\nu - \lambda_\beta)}, \quad \alpha, \beta = 1, \ldots, m-1,$$

where σ_i^j, $i, j = 1, \dots, m-2$ is the i^{th} elementary symmetric function in the $(m-2)$ variables $\lambda_1, \dots, \lambda_{j-1}, \lambda_{j+1}, \dots, \lambda_{m-1}$,

$$\sigma_i^j = \sum_{\substack{1 \le \nu_1 < \cdots < \nu_i \le m-1 \\ \nu_1, \dots, \nu_i \neq j}} \lambda_{\nu_1} \cdots \lambda_{\nu_i}, \qquad \sigma_0^j := 1.$$

By means of the identity (Lemma 1 in [Gau62])

$$\sum_{i=0}^{m-2} \sigma_i^j = \prod_{\substack{\nu=1 \\ \nu \neq j}}^{m-1} (1 + \lambda_\nu),$$

which holds since the λ_ν are all positive, we get

$$\begin{aligned} \|V^{-1}\| &= \max_{1 \le \beta \le m-1} \sum_{\alpha=1}^{m-1} \left| \left(V^{-1}\right)_{\alpha\beta} \right| = \max_{1 \le \beta \le m-1} \frac{1}{\lambda_\beta \prod_{\nu \neq \beta} |\lambda_\nu - \lambda_\beta|} \sum_{\alpha=1}^{m-1} \sigma_{m-1-\alpha}^\beta \\ &= \max_{1 \le \beta \le m-1} \frac{1}{\lambda_\beta} \prod_{\substack{\nu=1 \\ \nu \neq \beta}}^{m-1} \frac{1 + \lambda_\nu}{|\lambda_\nu - \lambda_\beta|}. \end{aligned}$$

Going back to inequality (C.2), we have so far proved that

$$\frac{h^k}{k!} \left| w^{(k)}(x) \right| \le (m-1) \|V^{-1}\| \left(2 \|w\|_{L^\infty([0,1])} + \frac{h^m}{m!} \|w^{(m)}\|_{L^\infty([0,1])} \right),$$

which yields

$$\begin{aligned} \left| w^{(k)}(x) \right| &\le (m-1) \|V^{-1}\| \left(\frac{2k!}{h^k} \|w\|_{L^\infty([0,1])} + h^{m-k} \frac{k!}{m!} \|w^{(m)}\|_{L^\infty([0,1])} \right) \\ &\le (m-1) \|V^{-1}\| \left(\frac{2m!}{h^k} \|w\|_{L^\infty([0,1])} + h^{m-k} \|w^{(m)}\|_{L^\infty([0,1])} \right) \end{aligned} \tag{C.3}$$

For $x \in [\frac{1}{2}, 1]$ the same calculations with h replaced by $-h$ proves inequality (C.3) in this case as well, so

$$\|w^{(k)}\|_{L^\infty([0,1])} \le (m-1) \|V^{-1}\| \left(\frac{2m!}{h^k} \|w\|_{L^\infty([0,1])} + h^{m-k} \|w^{(m)}\|_{L^\infty([0,1])} \right) \tag{C.4}$$

for all $h \in [0, \frac{1}{2}]$. Taking an arbitrary $u \in [0, 1]$, inequality (C.4) implies, with $h = \frac{u}{2} \in [0, \frac{1}{2}]$,

$$\|w^{(k)}\|_{L^\infty([0,1])} \le 2^m m! (m-1) \|V^{-1}\| \left(\frac{1}{u^k} \|w\|_{L^\infty([0,1])} + u^{m-k} \|w^{(m)}\|_{L^\infty([0,1])} \right),$$

which is the claimed inequality with

$$C_m = 2^m m! (m-1) \|V^{-1}\| = 2^m m! (m-1) \max_{1 \le \beta \le m-1} \frac{1}{\lambda_\beta} \prod_{\substack{\nu=1 \\ \nu \neq \beta}}^{m-1} \frac{1 + \lambda_\nu}{|\lambda_\nu - \lambda_\beta|}. \tag{C.5}$$

■

Remark. The constant C_m in equality (C.5) is far from optimal, but can be made small by minimising in the choice of the points $0 < \lambda_1 < \cdots < \lambda_{m-1} \leq 1$, suggesting that the optimal constant might be obtained by methods from approximation theory.

Indeed, by a more refined argument making use of numerical differentiation formulas, the minimisers of the associated multiplicative Kolmogorov-Landau inequality, i.e., extremisers of

$$M_k(\sigma) := \sup \left\{ \|w^{(k)}\|_{L^\infty([0,1])} : w \in W^{m,\infty}([0,1]), \|w\|_{L^\infty([0,1])} \leq 1, \|w^m\|_{L^\infty([0,1])} \leq \sigma \right\}$$

are explicitly known (at least for a wide range of parameters $m \in \mathbb{N}$ and $\sigma \geq 0$). The optimal Kolmogorov-Landau constants in these cases are given by the end-point values of certain Chebyshev type perfect splines. We refer to the papers by A. Pinkus [Pin78] and S. Karlin [Kar75], as well as the recent article by A. Shadrin [Sha14] and references therein. ♦

APPENDIX D

Properties of the function spaces $\mathscr{A}^\mu$

We prove a precise correspondence between the decay in Fourier space and the growth rate of derivatives of functions in $\mathscr{A}^\mu$.

Theorem D.1. *Let $\mu > 0$. Then*

$$\mathscr{A}^\mu(\mathbb{R}^d) = \bigcup_{\tau>0} \mathscr{D}\left(\mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}} : L^2(\mathbb{R}^d)\right).$$

Invoking a classic theorem by Denjoy and Carleman (see, for instance, [Coh68, KP02, Rud87]) one can show that the classes $\mathscr{A}^\mu$ for $\mu > 0$ are not quasi-analytic, that is, they contain non-vanishing $\mathscr{C}^\infty$ functions of arbitrarily small support.

Proof. Let $\mu > 0$ be fixed and assume first that $\|\mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}} f\|_{L^2} < \infty$ for some $\tau > 0$. Let $\alpha \in \mathbb{N}_0^d$ with $|\alpha| = n$ for some $n \in \mathbb{N}_0$. Then

$$\begin{aligned}\|\partial^\alpha f\|_{L^2}^2 &= \int_{\mathbb{R}^d} |(2\pi\mathrm{i}\eta)^\alpha \hat{f}(\eta)|^2 \,\mathrm{d}\eta \le (2\pi)^{2n} \int_{\mathbb{R}^d} \langle\eta\rangle^{2n} |\hat{f}|^2 \,\mathrm{d}\xi \\ &= 2n(2\pi)^{2n} \int_0^\infty t^{2n-1} \nu_f(\{\langle\eta\rangle > t\}) \,\mathrm{d}t\end{aligned}$$

where we introduced the (finite) measure $\nu_f(\mathrm{d}\eta) := |\hat{f}(\eta)|^2 \,\mathrm{d}\eta$. Since $\langle\eta\rangle \ge 1$ for all $\eta \in \mathbb{R}^d$, one has

$$\nu_f(\{\langle\eta\rangle > t\}) = \nu_f(\mathbb{R}^d) = \|f\|_{L^2(\mathbb{R}^d)}^2 \quad \text{for } t < 1.$$

For $t \ge 1$ we estimate

$$\begin{aligned}\nu_f(\{\langle\eta\rangle > t\}) &\le \mathrm{e}^{-2\tau(\log t)^{\mu+1}} \int_{\mathbb{R}^d} \mathrm{e}^{2\tau(\log\langle\eta\rangle)^{\mu+1}} \nu_f(\mathrm{d}\eta) \\ &= \mathrm{e}^{-2\tau(\log t)^{\mu+1}} \left\|\mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}} f\right\|_{L^2(\mathbb{R}^d)}^2,\end{aligned}$$

since $1 \le t \mapsto \mathrm{e}^{2\tau(\log t)^{\mu+1}}$ is increasing. It follows that

$$\begin{aligned}\|\partial^\alpha f\|_{L^2}^2 &\le (2\pi)^{2n}\|f\|_{L^2(\mathbb{R}^d)}^2 \\ &\quad + 2n(2\pi)^{2n}\left\|\mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}}f\right\|_{L^2(\mathbb{R}^d)}^2 \int_1^\infty t^{2n-1}\mathrm{e}^{-2\tau(\log t)^{\mu+1}}\,\mathrm{d}t.\end{aligned} \tag{D.1}$$

To extract the required growth in n from the latter integral, we essentially apply Laplace's method. Indeed, substituting the logarithm and rescaling suitably yields

$$\int_1^\infty t^{2n-1}\mathrm{e}^{-2\tau(\log t)^{\mu+1}}\,\mathrm{d}t = \left(\frac{n}{\tau}\right)^{1/\mu}\int_0^\infty \mathrm{e}^{2\tau^{-1/\mu}n^{1+1/\mu}(t-t^{\mu+1})}\,\mathrm{d}t. \tag{D.2}$$

The function $h : (0,\infty) \to \mathbb{R}$, $h(t) := t - t^{\mu+1}$ is strictly concave and attains its maximum at $t_* = (\mu+1)^{-1/\mu}$. If $\mu \ge 1$, h'' is negative and decreasing, so by Taylor's theorem we can bound

$$h(t) \le h(t_*) + \frac{h''(t_*)}{2}(t-t_*)^2 \mathbb{1}_{\{t>t_*\}}$$

and obtain with $h(t_*) = \mu(\mu+1)^{-(1+1/\mu)}$, $h''(t_*) = -\mu(\mu+1)^{1/\mu}$,

$$\begin{aligned}&\left(\frac{n}{\tau}\right)^{1/\mu}\int_0^\infty \mathrm{e}^{2\tau^{-1/\mu}n^{1+1/\mu}(t-t^{\mu+1})}\,\mathrm{d}t \\ &\quad\le \left(\frac{n}{\tau}\right)^{1/\mu}\left((\mu+1)^{-1/\mu} + \frac{\sqrt{\pi}}{2\sqrt{\mu}}\left(\frac{\tau}{\mu+1}\right)^{\frac{1}{2\mu}} n^{-\frac{\mu+1}{2\mu}}\right) \\ &\qquad\qquad \times \exp\left(2\tau^{-1/\mu}\mu(\mu+1)^{-(1+1/\mu)}n^{1+1/\mu}\right).\end{aligned} \tag{D.3}$$

Therefore, inserting the obtained bound into (D.1), there exist constants $C > 0$ and $b > 0$, depending on τ and μ, such that

$$\left\|\partial^\alpha f\right\|_{L^2} \le C^{|\alpha|+1}\mathrm{e}^{b|\alpha|^{1+1/\mu}} \quad \text{for all } \alpha \in \mathbb{N}_0^d. \tag{D.4}$$

For $\mu \in (0,1)$ the global bound (D.2) does not hold, but, as in the proof of Laplace's method for the asymptotics of integrals, one can find a suitable $\delta > 0$ such that the bound (D.2) holds on $[t_* - \delta, t_* + \delta]$ and the contribution to the integral outside of this interval is of much smaller order. So the right hand side of (D.3) still provides an upper bound modulo lower order terms and we conclude (D.4) also in this case.

For the converse assume that (D.4) holds. We want to show that there exists a $\tau > 0$ such that $\mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}}f \in L^2(\mathbb{R}^d)$. Using that

$$\mathrm{e}^{2\tau(\log\langle\eta\rangle)^{\mu+1}} = 1 + \int_1^{\langle\eta\rangle} 2\tau(\mu+1)t^{-1}(\log t)^\mu \mathrm{e}^{2\tau(\log t)^{\mu+1}}\,\mathrm{d}t$$

one obtains

$$\begin{aligned}&\left\|\mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}}f\right\|_{L^2(\mathbb{R}^d)}^2 \\ &\quad = \|f\|_{L^2(\mathbb{R}^d)}^2 + \int_1^\infty 2\tau(\mu+1)t^{-1}(\log t)^\mu \mathrm{e}^{2\tau(\log t)^{\mu+1}}\,\nu_f(\{\langle\eta\rangle > t\})\,\mathrm{d}t.\end{aligned} \tag{D.5}$$

Next we estimate for $t > 1$ and any $n \in \mathbb{N}_0$, since $|\eta|^2 \geq t^2 - 1$ on $\{\langle\eta\rangle > t\}$,

$$\nu_f(\{\langle\eta\rangle > t\} \leq \frac{1}{(2\pi)^{2n}(t^2-1)^n} \int_{\mathbb{R}^d} (2\pi)^{2n} |\eta|^{2n} |\hat{f}|^2 \,\mathrm{d}\eta.$$

By the multinomial theorem, we have (in the standard multi-index notation)

$$|\eta|^{2n} = \left(\sum_{i=1}^d \eta_i^2\right)^n = \sum_{\alpha\in\mathbb{N}_0^d : |\alpha|=n} \binom{n}{\alpha} \eta^{2\alpha},$$

so

$$\begin{aligned}
\nu_f(\{\langle\eta\rangle > t\} &\leq \frac{1}{(2\pi)^{2n}(t^2-1)^n} \sum_{|\alpha|=n} \binom{n}{\alpha} \int_{\mathbb{R}^d} \left|(2\pi\mathrm{i}\eta)^\alpha \hat{f}(\eta)\right|^2 \,\mathrm{d}\eta \\
&= \frac{1}{(2\pi)^{2n}(t^2-1)^n} \sum_{|\alpha|=n} \binom{n}{\alpha} \|\partial^\alpha f\|^2_{L^2(\mathbb{R}^d)} \\
&\leq \frac{d^n C^{2n+2}}{(2\pi)^{2n}} \frac{1}{(t^2-1)^n} \mathrm{e}^{2bn^{1+1/\mu}}
\end{aligned}$$

by assumption. Since this holds for any $n \in \mathbb{N}_0$, we even have

$$\begin{aligned}
\nu_f(\{\langle\eta\rangle > t\} &\leq \exp\left[\inf_{n\in\mathbb{N}_0} \left(2n\log A - n\log(t^2-1) + 2bn^{1+1/\mu}\right)\right] \\
&= \exp\left[2\inf_{n\in\mathbb{N}_0} \left(bn^{1+1/\mu} - n\log\frac{\sqrt{t^2-1}}{A}\right)\right]
\end{aligned}$$

where for notational convenience we set $A = \frac{C^{n+1}\sqrt{d}}{2\pi}$. If $\sqrt{t^2-1} < A$, then the infimum in the above exponent is just zero, so $\nu(\{\langle\eta\rangle > t\}) \leq 1$ in this case. If, however, $\sqrt{t^2-1} \geq A$, we get

$$\inf_{n\in\mathbb{N}_0} \left(bn^{1+1/\mu} - n\log\frac{\sqrt{t^2-1}}{A}\right) \leq bn_*^{1+1/\mu} - n_*\log\frac{\sqrt{t^2-1}}{A}$$

where $n_* = \left\lfloor \left(\frac{1}{b}\frac{\mu}{\mu+1}\log\frac{\sqrt{t^2-1}}{A}\right)^\mu \right\rfloor$, and $\lfloor a \rfloor$ denotes the greatest integer smaller or equal to $a \in \mathbb{R}$. Obviously,

$$n_* \leq \left(\frac{1}{b}\frac{\mu}{\mu+1}\log\frac{\sqrt{t^2-1}}{A}\right)^\mu < n_* + 1,$$

so we get the bound

$$\begin{aligned}
\inf_{n\in\mathbb{N}_0} \left(bn^{1+1/\mu} - n\log\frac{\sqrt{t^2-1}}{A}\right) \leq -\left(\frac{\mu}{\beta}\right)^\mu \left(\frac{1}{\mu+1}\log\frac{\sqrt{t^2-1}}{A}\right)^{\mu+1} \\
+ \log\frac{\sqrt{t^2-1}}{A}.
\end{aligned}$$

In particular, there exists $T_* > 1$ and $\beta > 0$ such that for $t > T_*$, one has

$$\inf_{n\in\mathbb{N}_0}\left(bn^{1+1/\mu} - n\log\frac{\sqrt{t^2-1}}{A}\right) \leq -\beta(\log t)^{\mu+1}.$$

This shows,

$$\nu_f(\{\langle\eta\rangle > t\}) \leq \mathrm{e}^{-\beta(\log t)^{\mu+1}}$$

for large enough t, and choosing $\tau < \beta/2$ in (D.5) we get the finiteness of $\left\|\mathrm{e}^{\tau(\log\langle D\rangle)^{\mu+1}}f\right\|_{L^2(\mathbb{R}^d)}$. ■

Bibliography

[Ale06] R. Alexandre, Integral estimates for a linear singular operator linked with the Boltzmann operator. I. Small singularities $0 < v < 1$. *Indiana University Mathematics Journal* **55** (2006), 1975–2022. MR 2284553. Zbl 1110.35062. 101

[Ale09] R. Alexandre, A review of Boltzmann equation with singular kernels. *Kinetic and Related Models* **2** (2009), 551–646. MR 2556715. Zbl 1193.35121. 18, 29, 101, 102

[ADVW00] R. Alexandre, L. Desvillettes, C. Villani, and B. Wennberg, Entropy dissipation and long-range interactions. *Archive for Rational Mechanics and Analysis* **152** (2000), 327–355. MR 1765272. Zbl 0968.76076. 26, 29, 31, 40, 43, 50, 89

[AE05] R. Alexandre, and M. ElSafadi, Littlewood-Paley theory and regularity issues in Boltzmann homogeneous equations. I. Non-cutoff case and Maxwellian molecules. *Mathematical Models and Methods in Applied Sciences* **15** (2009), 1–11. MR 2476677. Zbl 1168.35326. 32, 78

[AE09] R. Alexandre, and M. ElSafadi, Littlewood-Paley theory and regularity issues in Boltzmann homogeneous equations. II. Non cutoff case and non Maxwellian molecules. *Discrete and Continuous Dynamical Systems. Series A* **24** (2005), 907–920. MR 2149928. Zbl 1161.35331. 99

[AH08] R. Alexandre, and L. He, Integral estimates for a linear singular operator linked with Boltzmann operators. II. High singularities $1 \leq \nu < 2$. *Kinetic and Related Models* **1** (2008), 491–513. MR 2448604. Zbl 1155.35074. 101

[AMUXY08] R. Alexandre, Y. Morimoto, S. Ukai, C.-J. Xu, and T. Yang, Uncertainty principle and kinetic equations. *Journal of Functional Analysis* **255** (2008), 2013–2066. MR 2462585. Zbl 1166.35038. 99

[AMUXY10] R. Alexandre, Y. Morimoto, S. Ukai, C.-J. Xu, and T. Yang, Regularizing effect and local existence for the non-cutoff Boltzmann equation. *Archive for Rational Mechanics and Analysis* **198** (2010), 39–123. MR 2679369. Zbl 1257.76099. 40

[AMUXY11a] R. Alexandre, Y. Morimoto, S. Ukai, C.-J. Xu, and T. Yang, Global existence and full regularity of the Boltzmann equation without angular cutoff. *Communications in Mathematical Physics* **304** (2011), 513–581. MR 2795331. Zbl 1230.35082. 99

[AMUXY11b] R. Alexandre, Y. Morimoto, S. Ukai, C.-J. Xu, and T. Yang, The Boltzmann equation without angular cutoff in the whole space: I, Global existence for soft potential. *Journal of Functional Analysis* **262** (2012), 915–1010. MR 2863853. Zbl 1232.35110. 99

[AMUXY11c] R. Alexandre, Y. Morimoto, S. Ukai, C.-J. Xu, and T. Yang, The Boltzmann equation without angular cutoff in the whole space: II, Global existence for hard potential. *Analysis and Applications (Singapore)* **09** (2011), 113–134. MR 2793203. Zbl 1220.35110. 99

[AMUXY12] R. Alexandre, Y. Morimoto, S. Ukai, C.-J. Xu, and T. Yang, Smoothing effect of weak solutions for the spatially homogeneous Boltzmann equation without angular cutoff. *Kyoto Journal of Mathematics* **52** (2012), 433–463. MR 2959943. Zbl 1247.35085. 98, 99

[AV02] R. Alexandre, and C. Villani, On the Boltzmann equation for long-range interactions. *Communications in Pure and Applied Mathematics* **55** (2002), 30–70. MR 1857879. Zbl 1029.82036. 20

[Ark72] L. Arkeryd, On the Boltzmann equation. I: Existence. *Archive for Rational Mechanics and Analysis* **45** (1972), 1–16. MR 0339665. Zbl 0245.76059. 23

[Ark81] L. Arkeryd, Intermolecular forces of infinite range and the Boltzmann equation. *Archive for Rational Mechanics and Analysis* **77** (1981), 11–21. MR 0630119. Zbl 0547.76085. 17, 23

[AC90] L. Arkeryd, and C. Cercignani, On a functional equation arising in the kinetic theory of gases. *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti Lincei. Matematica e Applicazioni* Serie 9 **1** (1990), 139–149. MR 1081397. Zbl 0703.76062. 21

[Bob84] A. V. Bobylev, Exact solutions of the nonlinear Boltzmann equation and the theory of relaxation of a Maxwellian gas, translated from Teoret. Mat. Fiz. 60 No. 2 (1984), 280—310. *Theoretical and Mathematical Physics* **60** (1984), 820–841. MR 762269. Zbl 0565.76074. 26

[Bol72] L. Boltzmann, Weitere Studien über das Wärmegleichgewicht unter Gasmolecülen. *Akademie der Wissenschaften in Wien, Mathematisch-Naturwissenschaftliche Klasse, Sitzungsberichte, Abteilung IIa* **66** (1872), 275–370. Zbl 04.0566.01. 5

[Cer88] C. Cercignani, *The Boltzmann Equation and Its Applications*. *Applied Mathematical Sciences* **67**, Springer-Verlag, New York, 1988. MR 1313028. Zbl 0646.76001. 5, 7, 9, 10, 13

[CH11] Y. Chen, and L. He, Smoothing estimates for Boltzmann equation with full-range interactions: spatially homogeneous case. *Archive for Rational Mechanics and Analysis* **201** (2011), 501–548. MR 2820356. Zbl 1318.76018. 99

[CH12] Y. Chen, and L. He, Smoothing estimates for Boltzmann equation with full-range interactions: spatially inhomogeneous case. *Archive for Rational Mechanics and Analysis* **203** (2012), 343–377. MR 2885564. Zbl 06101988. 99

[Coh68] P. J. Cohen, A simple proof of the Denjoy-Carleman Theorem. *The American Mathematical Monthly* **75** (1968), 26–31. MR 0225957. Zbl 0153.08804. 113

[Des93] L. Desvillettes, Some applications of the method of moments for the homogeneous Boltzmann and Kac equations. *Archive for Rational Mechanics and Analysis* **123** (1993), 387–404. MR 1233644. Zbl 0784.76081. 98

[Des95] L. Desvillettes, About the regularizing properties of the non-cut-off Kac equation. *Communications in Mathematical Physics* **168** (1995), 417–440. MR 1324404. Zbl 0827.76081. 17, 23, 26, 31, 32, 78

[Des97] L. Desvillettes, Regularization properties of the 2-dimensional non radially symmetric non cutoff spatially homogeneous Boltzmann equation for Maxwellian molecules. *Transport Theory and Statistical Physics* **26** (1997), 341–357. MR 1475459. Zbl 0901.76072. 26, 31

[Des01] L. Desvillettes, Boltzmann's kernel and the spatially homogeneous Boltzmann equation. *Rivista di Matematica della Università di Parma (6)* **4*** (2001), 1–22. MR 1881103. Zbl 1078.76059. 23

[Des03] L. Desvillettes, About the use of Fourier transform for the Boltzmann equation. *Rivista di Matematica della Università di Parma (7)* **2*** (2003), 1–99. MR 2052786. Zbl 1072.35513. 15, 23, 32

[DFT09] L. Desvillettes, G. Furioli, and E. Terraneo, Propagation of Gevrey regularity for solutions of the Boltzmann equation for Maxwellian molecules. *Transactions of the American Mathematical Society* **361** (2009), 1731–1747. MR 2465814. Zbl 1159.76044. 39, 79, 95

[DL93] R. A. DeVore, and G. G. Lorentz, *Constructive Approximation*. Grundlehren der mathematischen Wissenschaften **303**, Springer-Verlag Berlin Heidelberg, 1993. MR 1261635. Zbl 0797.41016. 56, 109

[DM09] L. Desvillettes, and C. Mouhot, Stability and uniqueness for the spatially homogeneous Boltzmann equation with long-range interactions. *Archive for Rational Mechanics and Analysis* **193** (2009), 227–253. MR 2525118. Zbl 1169.76054. 99

[DW04] L. Desvillettes, and B. Wennberg, Smoothness of the solution of the spatially homogeneous Boltzmann equation without cutoff. *Communications in Partial Differential Equations* **29** (2004), 133–155. MR 2038147. Zbl 1103.82020. 30, 31, 78

[Elm83] T. Elmroth, Global boundedness of moments of solutions of the Boltzmann equation for forces of infinite range. *Archive for Rational Mechanics and Analysis* **82** (1983), 1–12. MR 0684411. Zbl 0503.76091. 98

[FL07] I. Fonseca, and G. Leoni, *Modern Methods in the Calculus of Variations: L^p Spaces*. Springer Monographs in Mathematics, Springer, New York, 2007. MR 2341508. Zbl 1153.49001. 24

[Gau62] W. Gautschi, On inverses of Vandermonde and confluent Vandermonde matrices. *Numerische Mathematik* **4** (1962), 117–123. MR 0139627. Zbl 0108.12501. 110, 111

[GN13] L. Glangetas, and M. Najeme, Analytical regularizing effect for the radial and spatially homogeneous Boltzmann equation. *Kinetic and Related Models* **6** (2013), 407–427. MR 3030717. Zbl 1262.35068. 32, 35

[Gou97] T. Goudon, On Boltzmann equations and Fokker–Planck asymptotics: Influence of grazing collisions. *Journal of Statistical Physics* **89** (1997), 751–776. MR 1484062. Zbl 0918.35136. 23

[GS10] P. T. Gressman, and R. M. Strain, Global classical solutions of the Boltzmann equation with long-range interactions. *Proceedings of the National Academy of Sciences of the United States of America* **107** (2010), 5744–5749. MR 2629879. Zbl 1205.82120. 99

[GS11] P. T. Gressman, and R. M. Strain, Global classical solutions of the Boltzmann equation without angular cut-off. *Journal of the American Mathematical Society* **24** (2011), 771–847. MR 2784329. Zbl 1248.35140. 99

[Kac59] M. Kac, *Probability and Related Topics in Physical Sciences*, with special lectures by G. E. Uhlenbeck, A. R. Hibbs, and B. van der Pol. Lectures in Applied Mathematics. Proceedings of the Summer Seminar, Boulder, CO, 1957. Interscience Publishers London-New York, 1959. MR 0102849. Zbl 0087.33003. 5, 21

[Kar75] S. Karlin, Interpolation properties of generalized perfect splines and the solutions of certain extremal problems. I. *Transactions of the American Mathematical Society* **206** (1975), 25–66. MR 0367512. Zbl 0303.41011. 112

[KP02] S. G. Krantz, and H. R. Parks, *A Primer of Real Analytic Functions*, 2^{nd} edition. *Birkhäuser Advanced Texts: Basler Lehrbücher*, Birkhäuser Boston Inc., Boston, MA, 2002. MR 1916029. Zbl 1015.26030. 113

[LX09] N. Lekrine, and C.-J. Xu, Gevrey regularizing effect of the Cauchy problem for non-cutoff homogeneous Kac's equation. *Kinetic and Related Models* **2** (2009), 647–666. MR 2556716. Zbl 1194.35089. 32, 78

[LL97] L. D. Landau, and E. M. Lifschitz, *Mechanik*, 14., korrigierte Auflage. Lehrbuch der theoretischen Physik, I. Harri Deutsch, Thun, Frankfurt am Main, 1997. 7

[LMPX14] N. Lerner, Y. Morimoto, K. Pravda-Starov, and C.-J. Xu, Gelfand-Shilov smoothing properties of the radially symmetric spatially homogeneous Boltzmann equation without angular cutoff. *Journal of Differential Equations* **256** (2014), 797–831. MR 3121714. Zbl 1315.35141. 32

[LMPX15] N. Lerner, Y. Morimoto, K. Pravda-Starov, and C.-J. Xu, Gelfand-Shilov and Gevrey smoothing effect for the spatially inhomogeneous non-cutoff Kac equation. *Journal of Functional Analysis* **269** (2015), 459–535. MR 3348825. Zbl 06439990. 99

[LO97] C. D. Levermore, and M. Oliver, Analyticity of solutions for a generalized Euler equation. *Journal of Differential Equations* **133** (1997), 321–339. MR 1427856. Zbl 0876.35090. 30

[Lin14] S.-Y. Lin, Gevrey regularity for the noncutoff nonlinear homogeneous Boltzmann equation with strong singularity. *Abstract and Applied Analysis* (2014), 584169 [9 pages]. MR 3166631. 32

[Lio94a] P.-L. Lions, Compactness in Boltzmann's equation via Fourier integral operators and applications. I, II. *Journal of Mathematics of Kyoto University* **34** (1994), 391–427, 429–461. MR 1284432. Zbl 0831.35139. 29

[Lio94b] P.-L. Lions, Compactness in Boltzmann's equation via Fourier integral operators and applications. III. *Journal of Mathematics of Kyoto University* **34** (1994), 539–584. MR 1295942. Zbl 0884.35124. 29

[MW99] S. Mischler, and B. Wennberg, On the spatially homogeneous Boltzmann equation. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis* **16** (1999), 467–501. MR 1697562. Zbl 0946.35075. 23, 98

[Mor87] Y. Morimoto, Hypoellipticity for infinitely degenerate elliptic operators. *Osaka Journal of Mathematics* **24** (1987), 13–35. MR 881744. Zbl 0658.35039. 38

[MU10] Y. Morimoto, and S. Ukai, Gevrey smoothing effect of solutions for spatially homogeneous nonlinear Boltzmann equation without angular cutoff. *Journal of Pseudo-Differential Operators and Applications* **1** (2010), 139–159. MR 2679746. Zbl 1207.35015. 33

[MUXY09] Y. Morimoto, S. Ukai, C.-J. Xu, and T. Yang, Regularity of solutions to the spatially homogeneous Boltzmann equation without angular cutoff. *Discrete and Continuous Dynamical Systems* **24** (2009), 187–212. MR 2476686. Zbl 1169.35315. 9, 11, 31, 32, 37, 38, 39, 40, 45, 50, 78, 81, 89, 101, 105

[MX09] Y. Morimoto, and C.-J. Xu, Ultra-analytic effect of Cauchy problem for a class of kinetic equations. *Journal of Differential Equations* **247** (2009), 596–617. MR 2523694. Zbl 1175.35024. 33, 100

[MV04] C. Mouhot, and C. Villani, Regularity theory for the spatially homogeneous Boltzmann equation with cut-off. *Archive for Rational Mechanics and Analysis* **173** (2004), 169–212. MR 2081030. Zbl 1063.76086. 30, 33

[New82] R. G. Newton, Scattering Theory of Waves and Particles, 2nd edition. Texts and Monographs in Physics, Springer, Berlin, Heidelberg, 1982. MR 0666397. Zbl 0496.47011. 7

[Pao74] Y.-P. Pao, Boltzmann collision operator with inverse-power intermolecular potentials. I, II. *Communications on Pure and Applied Mathematics* **27** (1974), 407–428, 559–581. MR 0636407. Zbl 0304.45010. 29

[Pin78] A. Pinkus, Some extremal properties of perfect splines and the pointwise Landau problem on the finite interval. *Journal of Approximation Theory* **23** (1978), 37–64. MR 0499027. Zbl 0395.41006. 112

[Rud87] W. Rudin, *Real and Complex Analysis*, 3rd edition. McGraw-Hill Book Co., New York, 1987. MR 924157. Zbl 0925.00005. 113

[Sha14] A. Shadrin, The Landau-Kolmogorov inequality revisited. *Discrete and Continuous Dynamical Systems* **34** (2014), 1183–1210. MR 3094569. Zbl 1280.41012. 112

[TV99] G. Toscani, and C. Villani, Probability metrics and uniqueness of the solution to the Boltzmann equation for a Maxwell gas. *Journal of Statistical Physics* **94** (1999), 619–637. MR 1675367. Zbl 0958.82044. 23

[Tru56] C. Truesdell, On the pressures and the flux of energy in a gas according to Maxwell's kinetic theory. II. *Journal of Rational Mechanics and Analysis* **5** (1956), 55–128. MR 0075726. Zbl 0070.23505. 32

[Uka84] S. Ukai, Local solutions in Gevrey classes to the nonlinear Boltzmann equation without cutoff. *Japan Journal of Applied Mathematics* **1** (1984), 141–156. MR 0839310. Zbl 0597.76072. 32

[Vil98] C. Villani, On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Archive for Rational Mechanics and Analysis* **143** (1998), 273–307. MR 1650006. Zbl 0912.45011. 17, 23, 98, 102

[Vil02] C. Villani, *A review of mathematical topics in collisional kinetic theory*. Handbook of Mathematical Fluid Dynamics (eds. S. Friedlander, and D. Serre) Vol. **1**, pp. 71–305. North-Holland, Amsterdam, 2002. MR 1942465. Zbl 1170.82369. 5, 6, 14, 15, 18, 21, 29, 36, 61

[Wen97] B. Wennberg, Entropy dissipation and moment production for the Boltzmann equation. *Journal of Statistical Physics* **86** (1997), 1053–1066. MR 1450762. Zbl 0935.82035. 98

[Wil51] E. Wild, On Boltzmann's equation in the kinetic theory of gases. *Mathematical Proceedings of the Cambridge Philosophical Society* **47** (1951), 602–609. MR 0042999. Zbl 0043.43703. 33

[YZ12] Z. Yin, and T.-F. Zhang, Gevrey regularity of spatially homogeneous Boltzmann equation without cutoff. *Journal of Differential Equations* **253** (2012), 1172–1190. MR 2925909. Zbl 1250.35053. 33

[YZ14] Z. Yin, and T.-F. Zhang, Gevrey regularity for solutions of the non-cutoff Boltzmann equation: The spatially inhomogeneous case. *Nonlinear Analysis: Real World Applications* **15** (2014), 246–261. MR 3110569. Zbl 1296.35113. 33

Part II

The Kac Model

CHAPTER 6

Introduction

Among the models describing a gas of interacting particles, the Kac master equation [Kac56], due to its simplicity, occupies a special place. It is useful in illuminating various issues in kinetic theory, e.g., providing a reasonably satisfactory derivation of the spatially homogeneous Boltzmann equation and giving a mathematical framework for investigating the approach to equilibrium. These issues were, in fact, the motivation for KAC's original work [Kac56]. Although it does not have a foundation in Hamiltonian mechanics, the Kac master equation is based on simple probabilistic principles and yields a linear evolution equation for the velocity distribution for N particles undergoing collisions.

The Kac master equation

Assume for simplicity that the gas is homogeneous and consists of N identical particles[1] with one-dimensional velocities. We will denote the velocity vector of the ensemble of particles by $\boldsymbol{v} = (v_1, \ldots, v_N) \in \mathbb{R}^N$.

The "collisions" in Kac's model are described by the following rule:

(1) Pick a pair of distinct indices (i, j) in $\{1, \ldots, N\}$ uniformly at random. The particles with labels i and j are declared as the particles that will collide.

(2) Randomly pick an angle $\theta \in [-\pi, \pi)$ (the "scattering angle") according to a probability distribution $\mathrm{d}\rho(\theta)$ on $[-\pi, \pi)$.

(3) Update the velocities of the two colliding particles by the rotation

$$\begin{pmatrix} v_i \\ v_j \end{pmatrix} \mapsto \begin{pmatrix} v_i^*(\theta) \\ v_j^*(\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} v_i \\ v_j \end{pmatrix}.$$

[1] whose mass m we set to $m = 2$ for convenience.

Since the collisions are just rotations in a two-dimensional plane, the total kinetic energy

$$E = \sum_{i=1}^{N} v_i^2 = |v|^2$$

is conserved.[2] Thus, the above collision rule generates a random walk on the energy sphere $\mathbb{S}^{N-1}(\sqrt{E})$ (Kac sphere). Assuming that each particle carries a unit energy, we can set $E = N$.

If the interaction times, i.e., the points in time where a collision occurs, are chosen according to a Poisson process with intensity $\lambda > 0$, we obtain a continuous-time Markov process with master equation (Kac master equation)

$$\frac{\mathrm{d}}{\mathrm{d}t} F_N(t; v) = -\mathscr{L}_N F_N(t; v) := \lambda N \left(Q_N - I\right) F_N(t; v) \tag{6.1}$$

for the probability distribution F_N of the velocities, with initial probability distribution $F_N(0; \cdot) = F_{N,0}$, where

$$(Q_N F_N)(t; v) = \frac{1}{\binom{N}{2}} \sum_{1 \le i < j \le N} \int_{-\pi}^{\pi} F_N(t; v_1, \ldots, v_i^*(\theta), \ldots, v_j^*(\theta), \ldots, v_N)\, \mathrm{d}\rho(\theta).$$

It is easy to see that if $\mathrm{d}\rho(\theta)$ is a symmetric probability measure, $\mathrm{d}\rho(\theta) = \mathrm{d}\rho(-\theta)$, then Q_N is a self-adjoint operator on $L^2(\mathbb{S}^{N-1}(\sqrt{N}), \mathrm{d}\sigma^{(N)})$. Here, $\mathrm{d}\sigma^{(N)}$ denotes the uniform probability measure on $\mathbb{S}^{N-1}(\sqrt{N})$. Moreover, one can show that $0 \le Q_N \le 1$ and

$$Q_N F = F \quad \text{if and only if} \quad F = 1,$$

that is, Q_N and $\mathrm{e}^{-\mathscr{L}_N t}$ are ergodic. As a direct consequence of the spectral theorem, we have convergence to equilibrium in $L^2(\mathbb{S}^{N-1}(\sqrt{N}))$.

Lemma 6.1 (Approach to equilibrium in L^2)**.** *Assume that F_0 is a probability distribution on $\mathbb{S}^{N-1}(\sqrt{N})$ with $F_0 \in L^2(\mathbb{S}^{N-1}(\sqrt{N}), \mathrm{d}\sigma^{(N)})$. Then*

$$\|\mathrm{e}^{-\mathscr{L}_N t} F_0 - 1\|_2 \longrightarrow 0$$

as $t \longrightarrow \infty$.

For a proof of this lemma, we refer to [CCL11], from which most of the material for this short introduction is taken.

It is in this context that Kac invented the notion of *propagation of chaos* and he used it to derive the spatially homogeneous, nonlinear Kac-Boltzmann equation.

[2] Momentum cannot be conserved in this simple model, otherwise the only admissible "collisions" would be particles going through each other or exchanging their velocities.

Recall that for a Borel probability measure ν on $\mathbb{S}^{N-1}(\sqrt{N})$, the k^{th} marginal measure $\pi_k\nu$, $1 \le k \le N$, is the unique measure such that

$$(\pi_k\nu)(A) = \nu\left(\{(v_1,\ldots,v_k) \in A\}\right)$$

for all $A \in \mathscr{B}(\mathbb{R}^k)$, the Borel σ-algebra on $\mathbb{R}^k$.

Definition 6.2 (Chaos (Boltzmann property))**.** Let μ be a Borel probability measure on $\mathbb{R}$. A sequence of probability measures $\{\mu_N\}_{N\in\mathbb{N}}$ on $\mathbb{S}^{N-1}(\sqrt{N})$ is called *μ-chaotic*, if

(i) μ_N is symmetric under permutations for any $N \in \mathbb{N}$, and

(ii) for every $k \in \mathbb{N}$, the k^{th} marginal measure $\pi_k\mu_N$ converges to $\mu^{\otimes k}$ weakly in the sense of measures as $N \to \infty$, that is,

$$\int \varphi(v_1,\ldots,v_k)\,\mathrm{d}(\pi_k\mu_N)(v) \to \int \varphi(v_1,\ldots,v_k)\,\mathrm{d}\mu(v_1)\cdots\mathrm{d}\mu(v_k)$$

for any $\varphi \in \mathscr{C}_b(\mathbb{R}^k)$ as $N \to \infty$.

In a way, chaoticity measures how independent the velocities (viewed as random variables on $\mathbb{S}^{N-1}(\sqrt{N})$) of a fixed number of particles become in the limit $N \to \infty$. The following theorem gives the connection between the Kac master equation and the Kac-Boltzmann equation.

Theorem 6.3 (Propagation of Chaos, Kac [Kac56])**.** *Let $\{F_{N,0}\,\mathrm{d}\sigma^{(N)}\}_{N\in\mathbb{N}}$ be an $f_0(v)\,\mathrm{d}v$-chaotic sequence, and let $F_N(t)$ be the solution of the Kac master equation* (6.1),

$$F_N(t) = \mathrm{e}^{N(Q_N-I)t}F_{N,0}$$

for some $t > 0$.[3] *Then $\{F_N(t)\,\mathrm{d}\sigma^{(N)}\}_{N\in\mathbb{N}}$ is an $f(t,v)\,\mathrm{d}v$-chaotic sequence, where f is a solution of the Kac-Boltzmann equation*

$$\partial_t f = 2\int_{-\pi}^{\pi}\mathrm{d}\rho(\theta)\int_{\mathbb{R}}\left[f(t,v^*(\theta))f(t,w^*(\theta)) - f(t,v)f(t,w)\right]\,\mathrm{d}w$$

with initial datum $f(0,\cdot) = f_0$.

The approach through master equations also led Kac to formulate the notion of approach to equilibrium and suggested various avenues to investigate this problem as the number of particles, N, becomes large. He emphasised that this could be done in a quantitative way if one could show, e.g., that the gap of the generator,

$$\Delta_N = \inf\{\langle F, N(I-Q_N)F\rangle : \|F\|_2 = 1, F \perp 1\}$$

is bounded below uniformly in N.[4]

This, known as Kac's conjecture [Kac56], was proved by Élise Janvresse in [Jan01] and, as a further sign of the simplicity of the model, the gap was computed explicitly in [CCL00, CCL03], see also [Mas03].

[3] We set $\lambda = 1$ for simplicity.

[4] Where $\langle F, G\rangle$ denotes the inner product on $L^2(\mathbb{S}^{N-1}(\sqrt{N}), \mathrm{d}\sigma^{(N)})$.

Theorem 6.4 (Carlen-Carvalho-Loss [CCL00, CCL03]). *Let* $d\rho(\theta) = \frac{1}{2\pi}\, d\theta$. *Then*

$$\Delta_N = \frac{1}{2}\frac{N+2}{N-1},$$

and the gap eigenfunction (unique up to a multiplicative constant) is given by

$$F_{\Delta_N} = \sum_{j=1}^{N}\left(v_j^4 - \frac{3N}{N+2}\right).$$

One of the problems in using the gap is that the approach to equilibrium is measured in terms of an L^2 distance. While this does seem to be a natural way to look at this problem, the size of the L^2 norm of approximately independent probability distributions increases exponentially with the size of the system. Thus, the half life of the L^2 norm is of order N.

A natural measure is, of course, given by the entropy, which is extensive, i.e, proportional to N. Let F be a probability distribution on $\mathbb{S}^{N-1}(\sqrt{N})$. Then the relative entropy of F with respect to the uniform probability measure $d\sigma^{(N)}$ on the Kac sphere is defined by[5]

$$H(F\,|\, d\sigma^{(N)}) := \int_{\mathbb{S}^{N-1}(\sqrt{N})} F(\boldsymbol{v}) \log F(\boldsymbol{v})\, d\sigma^{(N)}.$$

There has not been much success in proving exponential decay of the entropy with good rates. In [Vil03] Cédric Villani showed that the entropy decays exponentially, albeit with a rate that is bounded below by a quantity that is inversely proportional to N.

Theorem 6.5 (Entropy decay, Villani [Vil03]). *Let* F_0 *be a probability density on* $\mathbb{S}^{N-1}(\sqrt{N})$ *with finite relative entropy* $H(F_0|\, d\sigma^{(N)})$. *Then the solution* $F_N(t,\cdot)$ *of the Kac master equation* (6.1) *with initial datum* F_0 *satisfies the entropy inequality*

$$H(F_N(t,\cdot)|\, d\sigma^{(N)}) \le e^{-\frac{2}{N-1}t} H(F_0|\, d\sigma^{(N)}).$$

This estimate was complemented by Amit Einav [Ein11], who gave an example of a state that has entropy production essentially of order $1/N$. His example is the initial state in which most of the energy is concentrated in a few particles while most of the others have very little energy. One might surmise, based on physical intuition, that this state is physically very improbable and still has low entropy production because most of the particles are in some sort of equilibrium. This intuition can be made rigorous, see [Ein11], although by a quite difficult computation. One should add that low entropy production does not preclude exponential decay in entropy, i.e., large entropy production for the initial state might not be necessary for an exponential decay rate for the entropy.

[5] recall from Part I that this is again the negative of the physical entropy with respect to $d\sigma^{(N)}$.

A breakthrough was achieved by Stéphane Mischler and Clément Mouhot in [MM11, MM13]. They undertook a general investigation of the Kac program for gases of hard spheres and true Maxwellian molecules in three dimensions. Among the results of Mischler and Mouhot is a proof that these systems relax towards equilibrium in relative entropy as well as in Wasserstein distance with a rate that is independent of the particle number. As expected, they achieve this not for any initial condition, but rather for a natural class of chaotic states. The rate of relaxation is, however, polynomial in time.

To summarise, there is so far no mathematical evidence that the entropy in the Kac model in general decays exponentially with a rate that is independent of N and physical intuition suggests that for highly "improbable" states, such as the one used by Einav, this cannot be expected. One can restrict the class of initial conditions by considering chaotic states as done by Mischler and Mouhot, which shifts the problem of finding suitable initial conditions for proving exponential decay to the level of the non-linear Boltzmann equation.

CHAPTER 7

Entropy decay for the Kac evolution

We take a different approach here, one which is based on the idea of coupling a system of particles to a reservoir. Recall from [BLV14] the master equation of M particles with velocities $v = (v_1, v_2, \dots, v_M)$ interacting with a thermostat at temperature $1/\beta$,

$$\frac{\partial f}{\partial t} = \mathcal{L}_T f \ , \ f(v, 0) = f_0(v) \ . \tag{7.1}$$

The operator $\mathcal{L}_T$ is given by

$$\mathcal{L}_T f = \mu \sum_{j=1}^{M} (B_j - I) f \ ,$$

where

$$\begin{aligned} B_j[f](v) &:= \int_{\mathbb{R}} \mathrm{d}w \, \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}\theta \sqrt{\frac{\beta}{2\pi}} \, \mathrm{e}^{-\beta w_j^*(\theta)^2/2} f(v_j(\theta, w)) \ , \\ v_j(\theta, w) &= (v_1, \dots, v_j \cos(\theta) + w \sin(\theta), \dots, v_M) \ \text{ and} \\ w_j^*(\theta) &= -v_j \sin(\theta) + w \cos(\theta) \ . \end{aligned}$$

Thus, $B_j[f]$ describes the effect of a collision between particle j in the system and a particle in the reservoir. After the collision, the particle from the thermostat is discarded, which ensures that the thermostat stays in equilibrium. The interaction times with the thermostat are given by a Poisson process whose intensity μ is chosen so that the average time between two successive interactions of a given particle with the thermostat is independent of the number of particles in the system. Then the entropy decays exponentially fast. In fact, abbreviating $\sqrt{\frac{\beta}{2\pi}}\mathrm{e}^{-\frac{\beta}{2}v^2} = \Gamma_\beta(v)$, we know from [BLV14], that

$$S(f(\cdot,t)) := \int_{\mathbb{R}^M} f(v,t)\log\left(\frac{f(v,t)}{\Gamma_\beta(v)}\right) dv \le \mathrm{e}^{-\mu t/2} S(f_0)\,.$$

Thus, one might guess that if a "small" *system* of M particles out of equilibrium interacts with a *reservoir*, that is a large system of $N \ge M$ particles in thermal equilibrium, then the entropy decays exponentially fast in time. This intuition is also supported by the results in [BLTV17]. There it was shown that if the thermostat is replaced by a large but finite reservoir *initially* in thermal equilibrium, this evolution is close to the evolution given by the thermostat. This results holds in various norms and, in particular, it is *uniform* in time. We would like to emphasise that the reservoir will not stay in thermal equilibrium as time progresses, nevertheless it will not veer far from it.

Since this is the model that we consider in this work, we will now describe it in detail. We consider probability distributions $F : \mathbb{R}^{M+N} \to \mathbb{R}_+$ and write $F(v,w)$ where $v = (v_1,\dots,v_M)$ describes the particles in the small system, whereas $w = (w_{M+1},\dots,w_{N+M})$ describes the particles in the large system. The Kac master equation is given by

$$\frac{\partial F}{\partial t} = \mathcal{L}F\ ,\ F(v,w,0) = F_0(v,w) = f_0(v)\mathrm{e}^{-\pi|w|^2}\,, \tag{7.2}$$

where

$$\begin{aligned}\mathcal{L} = \frac{\lambda_S}{M-1}\sum_{1\le i<j\le M}(R_{ij}-I) + \frac{\lambda_R}{N-1}\sum_{M<i<j\le N+M}(R_{ij}-I)\\ + \frac{\mu}{N}\sum_{i=1}^{M}\sum_{j=M+1}^{M+N}(R_{ij}-I)\ ,\end{aligned} \tag{7.3}$$

and R_{ij} is given as follows. For $1 \le i < j \le M$ we have

$$(R_{ij}F)(v,w) = \int_{-\pi}^{\pi}\rho(\theta)\,\mathrm{d}\theta\, F(r_{ij}(\theta)^{-1}(v,w))\,,$$

where

$$r_{ij}(\theta)^{-1}(v,w) = (v_1,\dots,v_i\cos\theta - v_j\sin\theta,\dots,v_i\sin\theta + v_j\cos\theta,\dots,v_M,w)\,. \tag{7.4}$$

The other R_{ij}s are defined analogously. We assume that the probability measure ρ is smooth and satisfies

$$\int_{-\pi}^{\pi}\rho(\theta)\,\mathrm{d}\theta\ \sin\theta\cos\theta = 0\,. \tag{7.5}$$

In particular, we do *not* require $\mathcal{L}$ to be self-adjoint on $L^2(\mathbb{R}^{N+M})$, a condition called *microscopic reversibility*. The initial state of the reservoir is assumed to be a thermal equilibrium state and we have chosen units in which the inverse temperature $\beta = 2\pi$. Note that λ_S is the rate at which one particle from the system will scatter with any other particle in the system and similarly for λ_R. Likewise, μ is the rate at which a single particle of the system will scatter with any particle in the reservoir. The rate at which a particular particle from the reservoir will scatter with a particle in the system is given by $\mu M/N$. Hence, when N is large compared to M this process is suppressed and one expects that the reservoir does not move far from its equilibrium. Indeed, it is shown in [BLTV17] that the solution of the master equation (7.3) stays close to the solution of a thermostated system in the Gabetta-Toscani-Wennberg metric,

$$d_{GTW}(F,G) := \sup_{k\neq 0} \frac{|\widehat{F}(k) - \widehat{G}(k)|}{|k|^2},$$

see [GTW95]. Here, $\widehat{F}$ denotes the Fourier transform of F. More precisely, with the initial conditions of (7.1) and (7.2), it was shown that

$$d_{GTW}(f(v,t)e^{-\pi|w|^2}, F(v,w,t)) \le C(f_0)\frac{M}{N},$$

where $C(f_0)$ is a constant that depends on the initial condition but is of order one. The distance varies inversely as N, the size of the reservoir and, moreover, this estimate holds *uniformly* in time. For a detailed description of the results we refer the reader to [BLTV17]. From this result and the fact that the entropy of the system interacting with a thermostat decays exponentially in time, one might surmise that the entropy of the system interacting with a finite reservoir also decays exponentially fast in time. In fact we shall show this to be true if we consider the entropy relative to the thermal state.

7.1 Results

For the solution of the master equation (7.2) we use interchangeably the notation

$$F(v,w,t) = (e^{\mathcal{L}t}F_0)(v,w). \tag{7.6}$$

This evolution preserves the energy, hence it suffices to consider it on the space $L^1(\mathbb{S}^{N+M-1}(\sqrt{N+M}))$ with the normalised surface measure. Likewise, it is easy to see that the evolution is ergodic on $\mathbb{S}^{N+M-1}(\sqrt{N+M})$ in the sense that $e^{\mathcal{L}t}F_0 \to 1$ as $t \to \infty$ and 1 is the only normalised equilibrium state.

For our purposes it is convenient to consider the evolution in $L^1(\mathbb{R}^{M+N})$ equipped with Lebesgue measure. Then $e^{\mathcal{L}t}F_0$ converges to the spherical average of F_0 taken over spheres in $\mathbb{R}^{M+N}$. In this space we choose the initial condition

$$F_0(v,w) = f_0(v)\, e^{-\pi|w|^2}, \tag{7.7}$$

with a probability distribution f_0 on $\mathbb{R}^M$. Moreover, we introduce the function f,

$$f(v,t) := \int_{\mathbb{R}^N} \left[e^{\mathcal{L}t}F_0\right](v,w)\, dw, \tag{7.8}$$

and we call

$$S(f(\cdot,t)) := \int_{\mathbb{R}^M} f(v,t)\log\left(\frac{f(v,t)}{\mathrm{e}^{-\pi|v|^2}}\right) \mathrm{d}v$$

the entropy of f *relative* to the thermal state $\mathrm{e}^{-\pi|v|^2}$. Our main result is the following theorem.

Theorem 7.1. *Let $N \geq M$ and let ρ be a probability distribution with an absolutely convergent Fourier series such that (7.5) holds. The entropy of f* relative *to the thermal state $\mathrm{e}^{-\pi|v|^2}$ then satisfies*

$$S(f(\cdot,t)) \leq \left[\frac{M}{N+M} + \frac{N}{N+M}\mathrm{e}^{-t\mu_\rho(N+M)/N}\right] S(f_0),$$

where

$$\mu_\rho = \mu \int_{-\pi}^{\pi} \rho(\theta)\,\mathrm{d}\theta\, \sin^2(\theta),$$

and f_0 is as introduced in (7.7).

Remark. (i) Note that the theorem deals with the entropy relative to the thermal state and not with respect to the equilibrium state. The entropy relative to the equilibrium state tends to zero as $t \to \infty$. We do not know how to adapt our proof to this situation nor do we have any evidence that it does indeed tend to zero at an exponential rate. If this were the case, the rate would most likely depend on the initial condition.

(ii) The decay rate is universal in the sense that it only depends on μ and the distribution ρ. The intra-particle interactions in the system and in the reservoir do not seem to matter.

(iii) The statement of the theorem becomes particularly simple as $N \to \infty$. This corresponds to the thermostat problem treated in [BLV14] with the exact same decay rate. It is known that for the thermostat the decay rate is optimal, see [TV15], and hence the decay rate here is optimal as well.

(iv) Although we assume that ρ is smooth, our result also holds for the case where ρ is a finite sum of Dirac measures. In particular Theorem 7.1 also holds if ρ is a delta measure that has its mass at the angles $\theta = \pm\pi/2$, that is, our result does not depend on ergodicity of the evolution.

♦

As a consequence of Remark (ii), one obtains a result for the standard Kac model. Recall that the generator of the standard Kac model is given by

$$\mathcal{L}_{\mathrm{cl}} = \frac{2}{N+M-1} \sum_{1\leq i<j\leq N+M} (R_{ij} - I).$$

We may arbitrarily split the variables into two groups $(v_1, \ldots, v_M)$ and $(w_{M+1}, \ldots, w_{M+N})$. Splitting the generator accordingly,

$$\mathscr{L}_{\mathrm{cl}} = \frac{2}{N+M-1} \sum_{1 \le i < j \le M} (R_{ij} - I) + \frac{2}{N+M-1} \sum_{M+1 \le i < j \le N+M} (R_{ij} - I) + \frac{2}{N+M-1} \sum_{i=1}^{M} \sum_{j=M+1}^{N+M} (R_{ij} - I)\,,$$

we see that the standard Kac model can be cast in the from (7.3) by setting

$$\lambda_S = \frac{2(M-1)}{N+M-1}\,, \quad \lambda_R = \frac{2(N-1)}{N+M-1} \quad \text{and} \quad \mu = \frac{2N}{N+M-1}\,.$$

Hence, we obtain the following Corollary:

Corollary 7.2. *Let $N \ge M$ and consider the time evolution defined by $\mathscr{L}_{\mathrm{cl}}$ with initial condition (7.7). Assume that the function f_0 in the initial condition has finite entropy. The entropy of the function*

$$f(v,t) := \int_{\mathbb{R}^N} \left[\mathrm{e}^{\mathscr{L}_{\mathrm{cl}} t} F_0\right](v, w)\, \mathrm{d}w$$

relative *to the thermal state* $\mathrm{e}^{-\pi |v|^2}$, *satisfies*

$$S(f(\cdot, t)) \le \left[\frac{M}{N+M} + \frac{N}{N+M} \mathrm{e}^{-t\mu_\rho 2(N+M)/(N+M-1)}\right] S(f_0)\,,$$

where

$$\mu_\rho = \int_{-\pi}^{\pi} \rho(\theta)\, \mathrm{d}\theta\, \sin^2(\theta)$$

and ρ is a probability distribution such that (7.5) holds.

On a mathematical level, an efficient way of proving approach to equilibrium is through a logarithmic Sobolev inequality, which presupposes that the generator of the time evolution is given by a Dirichlet form. This kind of structure is notably absent in the Kac master equation. We shall see however, that the logarithmic Sobolev inequality in the form of Nelson's hypercontractive estimate is an important tool for the proof of Theorem 7.1. We will use an iterated version of it, which expresses the result in terms of marginals of the functions involved. This, coupled with an auxiliary computation and a sharp version of the Brascamp-Lieb inequalities [BL76] (see also [Lie90]) will lead to the result.

In our opinion, the main result of this chapter is the description of a simple mechanism for obtaining exponential relaxation towards equilibrium. One can extend the results to three dimensional momentum preserving collisions, however, so far only for a caricature of Maxwellian molecules. To carry this method over to the case of hard spheres and for true Maxwellian molecules is an open problem.

In Section 7.2 we derive a representation formula for the Kac evolution $e^{\mathcal{L}t}$ which is reminiscent of the Ornstein-Uhlenbeck process. This allows us to prove an entropy inequality based upon Nelson's hypercontractive estimate in Section 7.3. In Section 7.4 we show how the sharp version of the geometric Brascamp-Lieb inequality leads to a correlation inequality for the entropy involving marginals, which in turn proves our main entropy inequality. The fact that our Brascamp-Lieb datum is geometric relies on a sum rule which will be proved in Section 7.5. A short proof of the geometric form of the Brascamp-Lieb inequalities is deferred to Appendix E. In Chapter 8 we show how our method can be applied to three-dimensional Maxwellian collisions with a very simple angular dependence.

7.2 The representation formula

The aim of this section is to rewrite (7.6), that is $e^{\mathcal{L}t}F_0$, in a way which is reminiscent of the Ornstein-Uhlenbeck process. This representation will naturally lead to the next step in the proof of Theorem 7.1, namely the entropy inequality that will be presented in Theorem 7.5.

It is convenient to write

$$\mathcal{L} = \Lambda(Q - I)\,, \text{ where } \Lambda = \lambda_S \frac{M}{2} + \lambda_R \frac{N}{2} + \mu M\,,$$

and the operator Q is a convex combination of R_{ij}s, given by

$$Q = \frac{\lambda_S}{\Lambda(M-1)} \sum_{1 \le i < j \le M} R_{ij} + \frac{\lambda_R}{\Lambda(N-1)} \sum_{M < i < j \le N+M} R_{ij} + \frac{\mu}{\Lambda N} \sum_{i=1}^{M} \sum_{j=M+1}^{M+N} R_{ij}\,,$$

i.e., Q is an average over rotation operators. The right-hand side of (7.6) can be written as

$$(e^{\mathcal{L}t}F_0)(v,w) = e^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} Q^k F_0(v,w)\,, \tag{7.9}$$

where

$$Q^k F_0(v,w) \\ = \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\,d\theta_1 \cdots \rho(\theta_k)d\theta_k\, F_0\left(\left[\prod_{l=1}^{k} r_{\alpha_l}(\theta_l)\right]^{-1}(v,w)\right). \tag{7.10}$$

Here, α labels pairs of particles, that is, $\alpha = (i, j)$, $1 \le i < j \le M + N$, $r_\alpha(\theta)$ is defined in (7.4) and λ_α is given by the rotation corresponding to the index α, that is,

$$\begin{aligned}
\lambda_{(i,j)} &= \frac{\lambda_S}{\Lambda(M-1)} && \text{if } 1 \le i < j \le M\,, \\
\lambda_{(i,j)} &= \frac{\lambda_R}{\Lambda(N-1)} && \text{if } M+1 \le i < j \le M+N\,, \\
\lambda_{(i,j)} &= \frac{\mu}{\Lambda N} && \text{if } 1 \le i \le M\,, M+1 \le j \le M+N\,.
\end{aligned}$$

Note that the sum over *all* pairs $\sum_\alpha \lambda_\alpha \;=\; 1$.

For our purpose, it is convenient to write the function f_0, introduced in (7.7), as $f_0(v) = h_0(v)\mathrm{e}^{-\pi|v|^2}$. Since the Gaussian function is invariant under rotations, (7.9) takes the form

$$(\mathrm{e}^{\mathcal{L}t} F_0)(v, w) = \mathrm{e}^{-\pi\left(|v|^2+|w|^2\right)}\mathrm{e}^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} Q^k \left(h_0 \circ P\right)(v, w)\,.$$

We introduce the projection $P : \mathbb{R}^{N+M} \longrightarrow \mathbb{R}^M$ by $P(v, w) = v$, as a reminder that the semigroup $\mathrm{e}^{\mathcal{L}t}$ acts on functions that depend on v as well as w. If we write

$$f(v, t) = \mathrm{e}^{-\pi|v|^2} h(v, t)\,,$$

then (7.8) can be written as

$$h(v, t) = \mathrm{e}^{-\Lambda t} \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!} h_k(v)\,,$$

where the functions h_k are given by

$$h_k(v) := \int_{\mathbb{R}^N} Q^k \left(h_0 \circ P\right)(v, w)\mathrm{e}^{-\pi|w|^2}\, \mathrm{d}w\,.$$

Likewise, the entropy of f is expressed as

$$S(f(\cdot, t)) = \int_{\mathbb{R}^M} h(v, t) \log h(v, t)\mathrm{e}^{-\pi|v|^2}\, \mathrm{d}v =: \mathcal{S}(h(\cdot, t))\,.$$

Expanding the function $Q^k(h_0 \circ P)(v, w)$, we find that

$$\begin{aligned}
h_k(v) = \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\, \mathrm{d}\theta_1 \cdots \rho(\theta_k)\mathrm{d}\theta_k \times \\
\times \int_{\mathbb{R}^N} \left(h_0 \circ P\right)\left(\left[\prod_{l=1}^{k} r_{\alpha_l}(\theta_l)\right]^{-1} (v, w)\right)\mathrm{e}^{-\pi|w|^2}\, \mathrm{d}w\,, \qquad (7.11)
\end{aligned}$$

where, as before, see (7.10), $r_\alpha(\theta)$ rotates the plane given by the index pair α by an angle θ while keeping the other directions fixed. Since $P(\boldsymbol{v},\boldsymbol{w}) = \boldsymbol{v}$, it is natural to write

$$\left[\prod_{j=1}^{k} r_{\alpha_j}(\theta_j)\right]^{-1} = \begin{pmatrix} A_k(\underline{\alpha},\underline{\theta}) & B_k(\underline{\alpha},\underline{\theta}) \\ C_k(\underline{\alpha},\underline{\theta}) & D_k(\underline{\alpha},\underline{\theta}) \end{pmatrix} ,$$

where $A_k \in \mathbb{R}^{M\times M}$ is an $M \times M$ matrix, $B_k \in \mathbb{R}^{M\times N}$, $C_k \in \mathbb{R}^{N\times M}$ and $D_k \in \mathbb{R}^{N\times N}$. Further, $\underline{\alpha} = (\alpha_1, \dots, \alpha_k)$ and $\underline{\theta} = (\theta_1, \dots, \theta_k)$. This notation allows us to rewrite (7.11) as

$$\begin{aligned} h_k(\boldsymbol{v}) = \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\,\mathrm{d}\theta_1 \cdots \rho(\theta_k)\,\mathrm{d}\theta_k \times \\ \times \int_{\mathbb{R}^N} h_0\left(A_k(\underline{\alpha},\underline{\theta})\boldsymbol{v} + B_k(\underline{\alpha},\underline{\theta})\boldsymbol{w}\right) \mathrm{e}^{-\pi|\boldsymbol{w}|^2}\,\mathrm{d}\boldsymbol{w} \ . \end{aligned}$$

Note that, by the definition of rotations,

$$A_k(\underline{\alpha},\underline{\theta})A_k^T(\underline{\alpha},\underline{\theta}) + B_k(\underline{\alpha},\underline{\theta})B_k^T(\underline{\alpha},\underline{\theta}) = I_M \ . \tag{7.12}$$

Lemma 7.3. *Let $A \in \mathbb{R}^{M\times M}$ and $B \in \mathbb{R}^{M\times N}$ be matrices that satisfy $AA^T + BB^T = I_M$. Then*

$$\int_{\mathbb{R}^N} h(A\boldsymbol{v} + B\boldsymbol{w})\mathrm{e}^{-\pi|\boldsymbol{w}|^2}\,\mathrm{d}\boldsymbol{w} = \int_{\mathbb{R}^M} h\left(A\boldsymbol{v} + (I_M - AA^T)^{1/2}\boldsymbol{u}\right) \mathrm{e}^{-\pi|\boldsymbol{u}|^2}\,\mathrm{d}\boldsymbol{u}$$

for any integrable function h.

Proof. Denote the range of B by $H \subset \mathbb{R}^M$ and its kernel by $K \subset \mathbb{R}^N$. We may write

$$\begin{aligned} \int_{\mathbb{R}^N} h(A\boldsymbol{v} + B\boldsymbol{w})\mathrm{e}^{-\pi|\boldsymbol{w}|^2}\,\mathrm{d}\boldsymbol{w} &= \int_K \int_{K^\perp} h(A\boldsymbol{v} + B\boldsymbol{u})\mathrm{e}^{-\pi|\boldsymbol{u}|^2}\mathrm{e}^{-\pi|\boldsymbol{u}'|^2}\,\mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{u}' \\ &= \int_{K^\perp} h(A\boldsymbol{v} + B\boldsymbol{u})\mathrm{e}^{-\pi|\boldsymbol{u}|^2}\,\mathrm{d}\boldsymbol{u} \ . \end{aligned}$$

The symmetric map $BB^T : \mathbb{R}^M \to \mathbb{R}^M$ has H as its range and $H^\perp$, that is the orthogonal complement of H in $\mathbb{R}^M$, as its kernel. Indeed, suppose that there exists $x \in \mathbb{R}^M$ with $BB^T x = 0$, then $B^T x = 0$, i.e., $x \in \mathrm{Ker}B^T$ or x is perpendicular to H. Hence, the map $BB^T : H \to H$ is invertible. Define the linear map $R : \mathbb{R}^N \to H$ by

$$R = \left(BB^T\right)^{-1/2} B$$

and note that $RR^T = I_H$ while $R^T R$ projects the space $K^\perp$ orthogonally onto H. Since $K^\perp$ and H have the same dimension, it follows that R^T restricted to H defines an isometry between H and $K^\perp$. Hence,

$$\begin{aligned}\int_{K^\perp} h\,(A\boldsymbol{v} + B\boldsymbol{u})\,\mathrm{e}^{-\pi|\boldsymbol{u}|^2}\,\mathrm{d}\boldsymbol{u} &= \int_{K^\perp} h\left(A\boldsymbol{v} + \left(BB^T\right)^{1/2} R\boldsymbol{u}\right)\mathrm{e}^{-\pi|\boldsymbol{u}|^2}\,\mathrm{d}\boldsymbol{u}\\ &= \int_H h\left(A\boldsymbol{v} + \left(BB^T\right)^{1/2} RR^T\boldsymbol{u}\right)\mathrm{e}^{-\pi|R^T\boldsymbol{u}|^2}\,\mathrm{d}\boldsymbol{u}\\ &= \int_H h\left(A\boldsymbol{v} + \left(BB^T\right)^{1/2}\boldsymbol{u}\right)\mathrm{e}^{-\pi|\boldsymbol{u}|^2}\,\mathrm{d}\boldsymbol{u}\;.\end{aligned}$$

The assumption $AA^T + BB^T = I_M$, together with the fact that

$$\begin{aligned}&\int_H h\left(A\boldsymbol{v} + \left(BB^T\right)^{1/2}\boldsymbol{u}\right)\mathrm{e}^{-\pi|\boldsymbol{u}|^2}\,\mathrm{d}\boldsymbol{u}\\ &= \int_{H^\perp}\int_H h\left(A\boldsymbol{v} + (BB^T)^{1/2}\boldsymbol{u}\right)\mathrm{e}^{-\pi|\boldsymbol{u}|^2}\,\mathrm{d}\boldsymbol{u}\,\mathrm{e}^{-\pi|\boldsymbol{u}'|^2}\,\mathrm{d}\boldsymbol{u}'\end{aligned}$$

now implies the lemma. ■

The matrix $A_k(\underline{\alpha},\underline{\theta})$ has an orthogonal singular value decomposition,

$$A_k(\underline{\alpha},\underline{\theta}) = U_k(\underline{\alpha},\underline{\theta})\Gamma_k(\underline{\alpha},\underline{\theta})V_k^T(\underline{\alpha},\underline{\theta})\,, \tag{7.13}$$

where $\Gamma_k(\underline{\alpha},\underline{\theta}) = \mathrm{diag}[\gamma_{k,1}(\underline{\alpha},\underline{\theta}),\ldots,\gamma_{k,M}(\underline{\alpha},\underline{\theta})]$ is the diagonal matrix whose entries $\gamma_{k,j}(\underline{\alpha},\underline{\theta})$, $j = 1,\ldots,M$, are the singular values of $A_k(\underline{\alpha},\underline{\theta})$, and $U_k(\underline{\alpha},\underline{\theta})$ and $V_k(\underline{\alpha},\underline{\theta})$ are rotations in $\mathbb{R}^M$. Note that (7.12) implies $\gamma_{k,j}(\underline{\alpha},\underline{\theta}) \in [0,1]$ for $j = 1,\ldots,M$. We shall use the abbreviation

$$h_0(U_k(\underline{\alpha},\underline{\theta})\boldsymbol{v}) = h_{0,U_k(\underline{\alpha},\underline{\theta})}(\boldsymbol{v})\;.$$

These considerations can be summarized by the representation formula presented in the following theorem.

Theorem 7.4 (Representation formula). *The function h_k can be written as*

$$\begin{aligned}h_k(\boldsymbol{v}) = \sum_{\alpha_1,\ldots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\rho(\theta_1)\,\mathrm{d}\theta_1\cdots\rho(\theta_k)\,\mathrm{d}\theta_k\times\\ \times\int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}\left(\Gamma_k(\underline{\alpha},\underline{\theta})V_k^T(\underline{\alpha},\underline{\theta})\boldsymbol{v} + \left(I_M - \Gamma_k^2(\underline{\alpha},\underline{\theta})\right)^{1/2}\boldsymbol{w}\right)\mathrm{e}^{-\pi|\boldsymbol{w}|^2}\,\mathrm{d}\boldsymbol{w}\,,\end{aligned} \tag{7.14}$$

where $h_{0,U_k(\underline{\alpha},\underline{\theta})}$, $\Gamma_k(\underline{\alpha},\underline{\theta})$ and V_k are as defined above.

7.3 The hypercontractive estimate

Starting from (7.14) and using convexity of the entropy and Jensen's inequality together with

$$\sum_{\alpha_1,\ldots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\rho(\theta_1)\,\mathrm{d}\theta_1\cdots\rho(\theta_k)\,\mathrm{d}\theta_k = 1\,,$$

we get

$$\mathcal{S}(h_k) \leq \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\,\mathrm{d}\theta_1 \cdots \rho(\theta_k)\,\mathrm{d}\theta_k\, \mathcal{S}(g_k(\cdot,\underline{\alpha},\underline{\theta})),$$

where we set

$$g_k(v,\underline{\alpha},\underline{\theta}) = \int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}\left(\gamma_k(\underline{\alpha},\underline{\theta})v + \left(I_M - \gamma_k^2(\underline{\alpha},\underline{\theta})\right)^{1/2} w\right) \mathrm{e}^{-\pi|w|^2}\,\mathrm{d}w\ , \quad (7.15)$$

and we removed the rotation $V_k^T(\underline{\alpha},\underline{\theta})$ by a change of variables.

To explain the main observation in this section we look at (7.15) when $M = 1$. Since $0 \leq \gamma_k(\underline{\alpha},\underline{\theta}) \leq 1$, we can write $\gamma_k(\underline{\alpha},\underline{\theta}) = \mathrm{e}^{-t}$ and we get $g_k(v,\underline{\alpha},\underline{\theta}) = N_t(h_{0,U_k(\underline{\alpha},\underline{\theta})})$ where N_t is the Ornstein-Uhlenbeck semigroup, that is

$$N_t h(x) = \int_{\mathbb{R}} h\left(\mathrm{e}^{-t}x + \sqrt{1-\mathrm{e}^{-2t}}\,y\right) \mathrm{e}^{-\pi y^2}\,\mathrm{d}y\ .$$

Thus Theorem 7.4 renders the function h_k as a convex combination of terms reminiscent of the Ornstein-Uhlenbeck process, albeit in matrix form. We make use of this observation to find a bound for $\mathcal{S}(g_k(\cdot,\underline{\alpha},\underline{\theta}))$. This bound together with a suitable correlation inequality proved in the next section will lead to a bound for $\mathcal{S}(h_k)$.

In addition to the notation developed in the previous section, we need various marginals of the function $h_{0,U_k(\underline{\alpha},\underline{\theta})}$. Quite generally, if h is a function of M variables and $\sigma \subset \{1,\ldots,M\}$, we shall denote by h^σ the marginals of h with respect to the variables $v_j, j \in \sigma$, for instance,

$$h^{\{1,2\}}(v_3,\ldots,v_M) = \int_{\mathbb{R}^2} h(v_1,v_2,v_3,\ldots,v_M)\mathrm{e}^{-\pi\left(v_1^2+v_2^2\right)}\,\mathrm{d}v_1\mathrm{d}v_2\ .$$

It will be convenient to use the matrix $P_\sigma : \mathbb{R}^M \to \mathbb{R}^{|\sigma|}$ that projects $\mathbb{R}^M$ orthogonally onto $\mathbb{R}^{|\sigma|}$ which we will identify with subspace of $\mathbb{R}^M$. To give an example, let $v = (v_1,\ldots,v_M)$. Then $P_{\{1,2\}}v = (v_1,v_2)$. The following theorem is the main result of this section.

Theorem 7.5 (Partial entropy bound). *Let $h_0 \in L^1(\mathbb{R}^M, \mathrm{e}^{-\pi|v|^2}dv)$ be nonnegative and assume that $\mathcal{S}(h_0) < \infty$. Then*

$$\mathcal{S}(g_k(\cdot,\underline{\alpha},\underline{\theta})) \leq \sum_{\sigma\subset\{1,\ldots,M\}} \prod_{i\in\sigma^c} \gamma_{k,i}^2 \prod_{j\in\sigma}\left(1-\gamma_{k,j}^2\right) \times$$
$$\times \int_{\mathbb{R}^M} h_0(v) \log h^\sigma_{0,U_k(\underline{\alpha},\underline{\theta})}\left(P_{\sigma^c}U_k(\underline{\alpha},\underline{\theta})^T v\right) \mathrm{e}^{-\pi|v|^2}\,\mathrm{d}v\ , \quad (7.16)$$

where σ^c is the complement of the set σ in $\{1,\ldots,M\}$.

A key role in the proof of Theorem 7.5 is played by Nelson's hypercontractive estimate.

Theorem 7.6 (Nelson's hypercontractive estimate)**.** *The Ornstein-Uhlenbeck semigroup,*

$$N_t h(x) = \int_{\mathbb{R}} h\left(e^{-t}x + \sqrt{1-e^{-2t}}y\right) e^{-\pi y^2}\, dy\,,$$

for $t \geq 0$, is bounded from $L^p(\mathbb{R}, e^{-\pi x^2}\, dx)$ to $L^q(\mathbb{R}, e^{-\pi x^2}\, dx)$ if and only if

$$(p-1) \geq e^{-2t}(q-1)\,.$$

For such values of p and q,

$$\|N_t h\|_q \leq \|h\|_p$$

with equality if and only if h is constant.

Proof. For a proof we refer the reader to [Nel73]. For other proofs see [Gro75, Gro93, Fed69, CL90]. ■

Nelson's hypercontractive estimate, that is Theorem 7.6, implies the following Corollary, which will be useful in the proof of Theorem 7.5.

Corollary 7.7 (Entropic version of Nelson's hypercontractive estimate)**.** *Let $h : \mathbb{R} \to \mathbb{R}_+$ be a function in $L^1(\mathbb{R}, e^{-\pi x^2}\, dx)$ with finite entropy, i.e.,*

$$\mathcal{S}(h) = \int_{\mathbb{R}} h(x) \log h(x)\, e^{-\pi x^2} dx < \infty\,.$$

Then

$$\mathcal{S}(N_t h) \leq e^{-2t}\mathcal{S}(h) + (1-e^{-2t})\|h\|_1 \log \|h\|_1$$

for all $t \geq 0$.

Proof. Let $h \in L^p(\mathbb{R}, e^{-\pi x^2}\, dx)$, for $p \geq 1$ small, be a nonnegative function. As $\|N_t h\|_1 = \|h\|_1$, we can apply Nelson's hypercontractive estimate, which implies that for p, q that satisfy $(p-1) = e^{-2t}(q-1)$,

$$\frac{\|N_t h\|_q - \|N_t h\|_1}{q-1} \leq \frac{\|h\|_p - \|h\|_1}{q-1} = e^{-2t}\frac{\|h\|_p - \|h\|_1}{p-1}\,.$$

Sending $p \to 1$ and hence $q \to 1$, we get the claimed estimate for such functions h. If h just has finite entropy one cuts off h at large values, uses the above estimate and removes the cutoff using the monotone convergence theorem. ■

We are now ready to prove Theorem 7.5.

Proof of Theorem 7.5. Remember that $0 \le \gamma_{k,j}(\underline{\alpha}, \underline{\theta}) \le 1$ for $j = 1, \dots, M$. Thus, by inductively applying Corollary 7.7 to

$$\int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}\left(\gamma_{k,1} v_1 + \sqrt{1-\gamma_{k,1}^2}\, u_1, \dots, \gamma_{k,M} v_M + \sqrt{1-\gamma_{k,M}^2}\, u_M\right) \times \\ \times e^{-\pi \sum_{j=1}^M u_j^2}\, du_1 \cdots du_M\,,$$

we obtain

$$\begin{aligned} &\mathcal{S}(g_k(\cdot, \underline{\alpha}, \underline{\theta})) \\ &\le \sum_{\sigma \subset \{1,\dots,M\}} \prod_{i \in \sigma^c} \gamma_{k,i}^2 \prod_{j\in\sigma} (1-\gamma_{k,j}^2) \int_{\mathbb{R}^{|\sigma^c|}} h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(\boldsymbol{u}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(\boldsymbol{u})\, e^{-\pi|\boldsymbol{u}|^2}\, d\boldsymbol{u}\,. \end{aligned}$$

Inserting the definition of the marginal $h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}$, we see that

$$\begin{aligned} &\int_{\mathbb{R}^{|\sigma^c|}} h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(\boldsymbol{u}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(\boldsymbol{u})\, e^{-\pi|\boldsymbol{u}|^2}\, d\boldsymbol{u} \\ &= \int_{\mathbb{R}^M} h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}\boldsymbol{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}\boldsymbol{v})\, e^{-\pi|\boldsymbol{v}|^2}\, d\boldsymbol{v} \\ &= \int_{\mathbb{R}^M} h_{0,U_k(\underline{\alpha},\underline{\theta})}(\boldsymbol{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}\boldsymbol{v})\, e^{-\pi|\boldsymbol{v}|^2}\, d\boldsymbol{v} \\ &= \int_{\mathbb{R}^M} h_0(\boldsymbol{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c} U_k(\underline{\alpha},\underline{\theta})^T \boldsymbol{v})\, e^{-\pi|\boldsymbol{v}|^2}\, d\boldsymbol{v}, \end{aligned}$$

which finishes the proof of Theorem 7.5. ∎

7.4 The key entropy bound

Collecting the results of the previous sections we get the following bound

$$\begin{aligned} \mathcal{S}(h_k) \le \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \rho(\theta_1)\, d\theta_1 \cdots \rho(\theta_k)\, d\theta_k \sum_{\sigma\subset\{1,\dots,M\}} \prod_{i\in\sigma^c} \gamma_{k,i}^2 \times \\ \times \prod_{j\in\sigma} \left(1-\gamma_{k,j}^2\right) \int_{\mathbb{R}^M} h_0(\boldsymbol{v}) \log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}\left(P_{\sigma^c} U_k(\underline{\alpha},\underline{\theta})^T \boldsymbol{v}\right)\, e^{-\pi|\boldsymbol{v}|^2}\, d\boldsymbol{v}. \end{aligned} \tag{7.17}$$

The right-hand side of (7.17) contains a large sum over the entropy of marginals of h_0. In order to bound such a sum in terms of the entropy of h_0 one may try to apply some version of the Loomis-Whitney inequality [LW49] or, more precisely, of an inequality by Han [Han78]. This is essentially correct, but will require a substantial generalization of this inequality. Let us first formulate the main theorem of this section.

Theorem 7.8 (Entropy bound). *The estimate*

$$\mathcal{S}(h_k) \le \left[\frac{M}{N+M} + \frac{N}{N+M} \left(1 - \mu_\rho \frac{N+M}{N\Lambda}\right)^k \right] \mathcal{S}(h_0) \tag{7.18}$$

holds.

As mentioned before, to prove Theorem 7.8, we need a generalized version of an inequality by Han. This generalization was proven by Carlen-Cordero-Erausquin in [CC09]. It is based on the geometric Brascamp-Lieb inequality due to Ball [Bal89], see also [Bal91], in the rank one case, and due to Barthe [Bar98] in the general case.

Theorem 7.9 (Correlation inequality). *For $i = 1, \dots K$, let $H_i \subset \mathbb{R}^M$ be subspaces of dimension d_i and $B_i : \mathbb{R}^M \to H_i$ be linear maps with the property that $B_i B_i^T = I_{H_i}$, the identity map on H_i. Assume further that there are non-negative constants c_i, $i = 1, \dots, K$ such that*

$$\sum_{i=1}^{K} c_i B_i^T B_i = I_M \; . \tag{7.19}$$

Then, for nonnegative functions $f_i : H_i \to \mathbb{R}$,

$$\int_{\mathbb{R}^M} \prod_{i=1}^{K} f_i^{c_i}(B_i v)\, \mathrm{e}^{-\pi |v|^2}\, \mathrm{d}v \le \prod_{i=1}^{K} \left(\int_{H_i} f_i(u)\, \mathrm{e}^{-\pi |u|^2}\, \mathrm{d}u \right)^{c_i} . \tag{7.20}$$

Moreover,

$$\begin{aligned} \int_{\mathbb{R}^M} h(v) \log h(v)\, \mathrm{e}^{-\pi |v|^2}\, \mathrm{d}v \ge \sum_{i=1}^{K} c_i \Bigg[& \int_{\mathbb{R}^M} h(v) \log f_i(B_i v)\, \mathrm{e}^{-\pi |v|^2}\, \mathrm{d}v \\ & - \log \int_{H_i} f_i(u)\, \mathrm{e}^{-\pi |u|^2}\, \mathrm{d}u \Bigg], \end{aligned} \tag{7.21}$$

for any nonnegative function $h \in L^1(\mathbb{R}^M, \mathrm{e}^{-\pi |v|^2} dv)$.

Since Theorem 7.9 is very useful in a number of applications, and for the readers convenience, we will give an elementary proof in Appendix E.

Remark. By taking the trace in (7.19) one sees that

$$\sum_{i=1}^{K} c_i d_i = M \; .$$

♦

We would like to apply (7.21) to the right-hand side of (7.17). An immediate problem is that (7.17) is in terms of integrals and not sums. While there are some results available for continuous indices (see, e.g., [Bar04]), they do not apply to our situation and hence we will take a more direct approach and approximate the measure $\rho(\theta)\mathrm{d}\theta$ by a discrete measure. It is important that the approximation also satisfies the constraint (7.5). The following lemma establishes such an approximation.

Lemma 7.10. *Let ρ be a probability density on $[-\pi, \pi]$ whose Fourier series converges absolutely and assume that (7.5) is satisfied. There exists a sequence of discrete probability measures ν_K, $K = 1, 2, \ldots$, such that for every continuous function f on $[-\pi, \pi]$*

$$\lim_{K\to\infty} \int_{-\pi}^{\pi} f(\theta)\, \nu_K(\mathrm{d}\theta) = \int_{-\pi}^{\pi} f(\theta)\rho(\theta)\, \mathrm{d}\theta\, .$$

Moreover,

$$\int_{-\pi}^{\pi} \cos\theta \sin\theta\, \nu_K(\mathrm{d}\theta) = 0\, ,$$

for all $K \in \mathbb{N}$. More precisely,

$$\nu_K(\mathrm{d}\theta) = \frac{2\pi}{4K+1} \sum_{\ell=-2K}^{2K} \rho_K\left(\frac{2\pi\ell}{4K+1}\right) \delta\left(\theta - \frac{2\pi\ell}{4K+1}\right) \mathrm{d}\theta\, ,$$

where

$$\rho_K(\theta) = \int_{-\pi}^{\pi} \rho(\theta-\phi)\, p_K(\theta)\, \mathrm{d}\phi \ \textit{ and } \ p_K(\theta) := \frac{1}{2K+1}\left(\sum_{k=-K}^{K} \mathrm{e}^{\mathrm{i}k\theta}\right)^2 .$$

Proof. For K any positive integer we convolve $\rho(\theta)$ with the non-negative trigonometric polynomial

$$p_K(\theta) := \frac{1}{2K+1}\left(\sum_{k=-K}^{K} \mathrm{e}^{\mathrm{i}k\theta}\right)^2 = \sum_{m=-2K}^{2K}\left(1 - \frac{|m|}{2K+1}\right)\mathrm{e}^{\mathrm{i}m\theta}\, ,$$

and obtain a probability density $\rho_K(\theta)$. The Fourier coefficients of $\rho_K(\theta)$ are given by

$$\widehat{\rho}_K(m) = \widehat{\rho}(m)\left(1 - \frac{|m|}{2K+1}\right)$$

for $|m| \le 2K$ and are zero otherwise. In particular,

$$\widehat{\rho}_K(2) - \widehat{\rho}_K(-2) = 4\mathrm{i}\int_{-\pi}^{\pi} \rho_K(\theta) \sin\theta \cos\theta\, \mathrm{d}\theta = 0\, .$$

With ρ_K we construct the measure

$$\nu_K(d\theta) = \frac{2\pi}{4K+1} \sum_{\ell=-2K}^{2K} \rho_K\left(\frac{2\pi\ell}{4K+1}\right) \delta\left(\theta - \frac{2\pi\ell}{4K+1}\right) \mathrm{d}\theta\, .$$

The measure ν_K is positive since $\rho_K((2\pi\ell)/(4K+1)) \geq 0$. Moreover, for all $m \in \mathbb{Z}$ with $|m| \leq 2K$ the Fourier coefficients $\widehat{\nu}_K(m)$ and $\widehat{\rho}_K(m)$ coincide. In particular, we have

$$\int_{-\pi}^{\pi} \nu_K(\mathrm{d}\theta) \sin\theta \cos\theta = 0 .$$

To see this, we compute

$$\widehat{\nu}_K(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \nu_K(\theta) \mathrm{e}^{-\mathrm{i}m\theta} \, \mathrm{d}\theta = \frac{1}{4K+1} \sum_{\ell=-2K}^{2K} \rho_K\left(\frac{2\pi\ell}{4K+1}\right) \mathrm{e}^{-2\pi\mathrm{i}m\ell/(4K+1)}$$

for $|m| \leq 2K$. Observe that

$$\rho_K\left(\frac{2\pi\ell}{4K+1}\right) = \sum_{k=-2K}^{2K} \widehat{\rho}_K(k) \mathrm{e}^{2\pi\mathrm{i}k\ell/(4K+1)} ,$$

and, as a consequence,

$$\widehat{\nu}_K(m) = \frac{1}{4K+1} \sum_{\ell=-2K}^{2K} \sum_{k=-2K}^{2K} \widehat{\rho}_K(k) \mathrm{e}^{2\pi\mathrm{i}\ell(k-m)/(4K+1)} .$$

But

$$\sum_{\ell=-2K}^{2K} \mathrm{e}^{2\pi\mathrm{i}\ell(k-m)/(4K+1)} = \begin{cases} 4K+1 & \text{if } k = m \\ 0 & \text{if } k \neq m, \end{cases}$$

and hence we conclude that

$$\widehat{\nu}_K(m) = \widehat{\rho}_K(m) \tag{7.22}$$

for $|m| \leq 2K$. It is easy to see that for any continuous function f on $[-\pi, \pi]$,

$$\lim_{K\to\infty} \int_{-\pi}^{\pi} f(\theta) \nu_K(\mathrm{d}\theta) = \int_{-\pi}^{\pi} f(\theta) \rho(\theta) \, \mathrm{d}\theta .$$

Indeed, Weierstrass's theorem implies that for any $\varepsilon > 0$ there exist $K \in \mathbb{N}$ and a trigonometric polynomial q_K of degree less than $2K$, such that

$$f(\theta) = q_K(\theta) + r_k(\theta)$$

where r_k satisfies the uniform bound $|r_k(\theta)| < \frac{\varepsilon}{4}$. Further, (7.22) implies that

$$\int_{-\pi}^{\pi} q_K(\theta) \nu_K(\mathrm{d}\theta) = \int_{-\pi}^{\pi} q_K(\theta) \rho_K(\theta) \, \mathrm{d}\theta ,$$

which yields

$$\left| \int_{-\pi}^{\pi} f(\theta) \nu_K(\mathrm{d}\theta) - \int_{-\pi}^{\pi} f(\theta) \rho_K(\theta) \, \mathrm{d}\theta \right| < \frac{\varepsilon}{2} .$$

It remains to show that

$$\int_{-\pi}^{\pi} f(\theta)\,(\rho(\theta) - \rho_K(\theta))\;\mathrm{d}\theta \to 0$$

as $K \to \infty$. By assumption, ρ has an absolutely convergent Fourier series $\sum_k \widehat{\rho}(k)\mathrm{e}^{ik\theta}$. Hence,

$$|\rho(\theta) - \rho_K(\theta)| \leq \sum_{|k|\geq 2K+1} |\widehat{\rho}(k)| + \frac{1}{2K+1} \sum_{|k|\leq 2K} |k|\,|\widehat{\rho}(k)|\;.$$

As $K \to \infty$, the first term on the right-hand side tends to zero since $\{\widehat{\rho}(k)\}_k \in \ell^1(\mathbb{Z})$. Notice that by the summability of $\{\widehat{\rho}(k)\}_k$, $|k\widehat{\rho}(k)| \to 0$ as $|k| \to \infty$, so the second term also goes to zero as Cesàro mean. ■

At this point we can prepare the ground for the application of Theorem 7.9 to inequality (7.17). We first replace $\rho(\theta)\mathrm{d}\theta$ in (7.17) with $\nu_K(\mathrm{d}\theta)$. Setting

$$\omega_{\ell_j} = \rho_K(\theta_j)\,,\; \theta_{\ell_j} = \frac{2\pi\,\ell_j}{4K+1}\,,\text{ and } \underline{\theta} = (\theta_{\ell_1}, \ldots, \theta_{\ell_k})\,,$$

we obtain

$$\begin{aligned}
&\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu_K(\mathrm{d}\theta_1)\cdots\nu_K(\mathrm{d}\theta_k) \sum_{\sigma\subset\{1,\ldots,M\}} \prod_{i\in\sigma^{\mathrm{c}}} \gamma_{k,i}(\underline{\alpha},\underline{\theta})^2 \times \\
&\quad\times \prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right) \int_{\mathbb{R}^M} h_0(v)\log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^T v)\,\mathrm{e}^{-\pi|v|^2}\,\mathrm{d}v \\
&= \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \sum_{-K\leq\ell_1,\ldots,\ell_k\leq K} \prod_{j=1}^{k} \omega_{\ell_j} \sum_{\sigma\subset\{1,\ldots,M\}} \prod_{i\in\sigma^{\mathrm{c}}} \gamma_{k,i}(\underline{\alpha},\underline{\theta})^2 \times \\
&\quad\times \prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right) \int_{\mathbb{R}^M} h_0(v)\log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^T v)\,\mathrm{e}^{-\pi|v|^2}\,\mathrm{d}v\;.
\end{aligned} \tag{7.23}$$

In order to apply Theorem 7.9 to (7.23) we have to replace the sum over the index i with a sum over the indices $\alpha_1, \ldots, \alpha_k, \ell_1, \ldots \ell_k$ and all subsets $\sigma \subset \{1, \ldots, M\}$.

Moreover, we substitute

the constants c_i by $\dfrac{1}{C_{k,M}}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\prod_{j=1}^{k}\omega_{\ell_j}\prod_{i\in\sigma^c}\gamma_{k,i}(\underline{\alpha},\underline{\theta})^2\prod_{j\in\sigma}(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2)$,

the functions $f_i(w)$ by $h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(w)$,

the linear maps B_i by $P_{\sigma^c}U_k(\underline{\alpha},\underline{\theta})^T$,

the functions $f_i(B_i v)$ by $h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^c}U_k(\underline{\alpha},\underline{\theta})^T v)$,

and the subspaces H_i by $\mathbb{R}^{|\sigma^c|}$.

For any given index i the condition $B_iB_i^T = I_{H_i}$ corresponds to

$$P_{\sigma^c}U_k(\underline{\alpha},\underline{\theta})^T U_k(\underline{\alpha},\underline{\theta})P_{\sigma^c} = P_{\sigma^c}$$

which is the identity on $\mathbb{R}^{|\sigma^c|}$.

The next theorem establishes the sum rule (7.19) in our setting and hence ensures the applicability of Theorem 7.9 to (7.23).

Theorem 7.11 (The sum rule)**.** *If $\nu(\mathrm{d}\theta)$ is a probability measure satisfying (7.5), then*

$$\begin{aligned}
&\sum_{\alpha_1,\dots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{[-\pi,\pi]^k}\nu(\mathrm{d}\theta_1)\cdots\nu(\mathrm{d}\theta_k)\times\\
&\qquad\times\sum_{\sigma\subset\{1,\dots,M\}}\prod_{i\in\sigma^c}\gamma_{k,i}(\underline{\alpha},\underline{\theta})^2\prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right)U_k(\underline{\alpha},\underline{\theta})P_{\sigma^c}^TP_{\sigma^c}U_k(\underline{\alpha},\underline{\theta})^T\\
&= C_{k,M}I_M\,,
\end{aligned}\tag{7.24}$$

where

$$C_{k,M} = \left[\frac{M}{N+M}+\frac{N}{N+M}\left(1-\mu_\nu\frac{N+M}{N\Lambda}\right)^k\right]$$

with

$$\mu_\nu = \mu\int\nu(\mathrm{d}\theta)\sin^2\theta\,.$$

The proof will be given in Section 7.5. We observe here that it follows from Theorem 7.10 that $\mu_\rho = \lim_{K\to\infty}\mu_{\nu_K}$.

Proof of Theorem 7.8. First we consider the case where ρ is repaced by ν_K and use Theorem 7.9 together with Theorem 7.11 and the identification rules described above.

The entropy inequality (7.21) now says that

$$\begin{aligned}
&\int_{\mathbb{R}^M} h_0(v)\log h_0(v)\mathrm{e}^{-\pi|v|^2}\,\mathrm{d}v\\
&\geq \frac{1}{C_{k,M}}\sum_{\alpha_1,\dots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\sum_{-K\leq \ell_1,\dots,\ell_k\leq K}\prod_{j=1}^{k}\omega_{\ell_j}\sum_{\sigma\subset\{1,\dots,M\}}\prod_{i\in\sigma^{\mathrm{c}}}\gamma_{k,i}(\underline{\alpha},\underline{\theta})^2\times\\
&\qquad\times\prod_{j\in\sigma}\Big(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\Big)\Big[\int_{\mathbb{R}^M} h_0(v)\log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^T v)\,\mathrm{e}^{-\pi|v|^2}\,\mathrm{d}v\\
&\qquad\qquad -\log\int_{\mathbb{R}^{|\sigma^{\mathrm{c}}|}} h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(u)\,\mathrm{e}^{-\pi|u|^2}\,\mathrm{d}u\Big]\,.
\end{aligned}$$

However, since h_0 is normalized and $U_k(\underline{\alpha},\underline{\theta})$ is orthogonal, we find that

$$\begin{aligned}
\int_{\mathbb{R}^{|\sigma^{\mathrm{c}}|}} h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(u)\,\mathrm{e}^{-\pi|u|^2}\,\mathrm{d}\,u &= \int_{\mathbb{R}^{|\sigma^{\mathrm{c}}|}}\int_{\mathbb{R}^{|\sigma|}} h_{0,U_k(\underline{\alpha},\underline{\theta})}(v,u)\,\mathrm{e}^{-\pi|v|^2}\,\mathrm{d}v\,\mathrm{e}^{-\pi|u|^2}\,\mathrm{d}u\\
&= \int_{\mathbb{R}^M} h_0(U_k(\underline{\alpha},\underline{\theta})v)\,\mathrm{e}^{-\pi|v|^2}\,\mathrm{d}v\\
&= 1\,.
\end{aligned}$$

Thus, we find

$$\begin{aligned}
&\sum_{\alpha_1,\dots,\alpha_k}\lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\sum_{-K\leq \ell_1,\dots,\ell_k\leq K}\prod_{j=1}^{k}\omega_{\ell_j}\sum_{\sigma\subset\{1,\dots,M\}}\prod_{i\in\sigma^{\mathrm{c}}}\gamma_{k,i}(\underline{\alpha},\underline{\theta})^2\times\\
&\quad\times\prod_{j\in\sigma}\Big(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\Big)\int_{\mathbb{R}^M} h_0(v)\log h^{\sigma}_{0,U_k(\underline{\alpha},\underline{\theta})}(P_{\sigma^{\mathrm{c}}}U_k(\underline{\alpha},\underline{\theta})^T v)\,\mathrm{e}^{-\pi|v|^2}\,\mathrm{d}v\\
&\leq C_{k,M}\mathcal{S}(h_0)\,.
\end{aligned}\tag{7.25}$$

As $K\to\infty$, the left-hand side of (7.25) converges to the right-hand side of (7.17). ■

We now have all ingredients to give the proof of Theorem 7.1.

Proof of Theorem 7.1. Recall from Section 7.2, that

$$f(v,t)=\mathrm{e}^{-\pi|v|^2}\mathrm{e}^{-\Lambda t}\sum_{k=0}^{\infty}\frac{t^k\Lambda^k}{k!}h_k(v)\,,$$

and that $S(f(\cdot,t))=\mathcal{S}(h(\cdot,t))$. Combining Theorem 7.5 and Theorem 7.8, we obtain

$$\mathcal{S}(h_k)\leq C_{k,M}\mathcal{S}(h_0)\,,$$

and computing

$$\mathrm{e}^{-\Lambda t}\sum_{k=0}^{\infty}\frac{\Lambda^k t^k}{k!}C_{k,M}$$

yields Theorem 7.1. ■

7.5 The sum rule. Proof of Theorem 7.11

We have to compute the matrix

$$Z := \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1)\,\cdots\,\nu(\mathrm{d}\theta_k)\times$$
$$\times \sum_{\sigma\subset\{1,\dots,M\}} \prod_{i\in\sigma^c} \gamma_{k,i}(\underline{\alpha},\underline{\theta})^2 \prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right) U_k(\underline{\alpha},\underline{\theta}) P_{\sigma^c}^T P_{\sigma^c} U_k(\underline{\alpha},\underline{\theta})^T\,.$$

Obviously $P_{\sigma^c}^T P_{\sigma^c} = P_{\sigma^c}$ and hence

$$Z = \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1)\cdots\nu(\mathrm{d}\theta_k)\times$$
$$\times U_k(\underline{\alpha},\underline{\theta})\left[\sum_{\sigma\subset\{1,\dots,M\}} \prod_{i\in\sigma^c} \gamma_{k,i}(\underline{\alpha},\underline{\theta})^2 \prod_{j\in\sigma}\left(1-\gamma_{k,j}(\underline{\alpha},\underline{\theta})^2\right) P_{\sigma^c}\right] U_k(\underline{\alpha},\underline{\theta})^T\,.$$

The sum on σ is easily evaluated and yields the matrix $\Gamma_k^2(\underline{\alpha},\underline{\theta})$. Hence, recalling the orthogonal singular value decomposition (7.13) of $A_k(\underline{\alpha},\underline{\theta})$, that is, $A_k(\underline{\alpha},\underline{\theta}) = U_k(\underline{\alpha},\underline{\theta})\Gamma_k(\underline{\alpha},\underline{\theta})V_k^T(\underline{\alpha},\underline{\theta})$, we find that

$$Z = \sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1)\cdots\nu(\mathrm{d}\theta_k)\, A_k(\underline{\alpha},\underline{\theta})A_k^T(\underline{\alpha},\underline{\theta})\,. \tag{7.26}$$

One can think about this expression in the following fashion. Recall that

$$\left[\prod_{l=1}^{k} r_{\alpha_l}(\theta_l)\right]^{-1} = \begin{pmatrix} A_k(\underline{\alpha},\underline{\theta}) & B_k(\underline{\alpha},\underline{\theta}) \\ C_k(\underline{\alpha},\underline{\theta}) & D_k(\underline{\alpha},\underline{\theta}) \end{pmatrix}\,.$$

With this notation, the matrix Z equals the top left entry of the matrix

$$\sum_{\alpha_1,\dots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1)\cdots\nu(\mathrm{d}\theta_k) \left[\prod_{l=1}^{k} r_{\alpha_l}(\theta_l)\right]^{-1} \begin{pmatrix} I_M & 0 \\ 0 & 0 \end{pmatrix} \left[\prod_{l=1}^{k} r_{\alpha_l}(\theta_l)\right]\,.$$

The computation hinges on a repeated application of the elementary identity

$$\int_{-\pi}^{\pi} \nu(\mathrm{d}\theta) \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$
$$= \begin{pmatrix} (1-\tilde{\nu})m_1 + \tilde{\nu}m_2 & 0 \\ 0 & (1-\tilde{\nu})m_2 + \tilde{\nu}m_1 \end{pmatrix}\,,$$

where $\tilde{\nu} = \int \nu(\mathrm{d}\theta)\,\sin^2(\theta)$. For this to be true we just need (7.5). We easily check that for the rotations $r_\alpha(\theta)$

$$\begin{aligned}
&\sum_\alpha \lambda_\alpha \int_{-\pi}^{\pi} \nu(\mathrm{d}\theta)\, r_\alpha(\theta)^{-1} \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} r_\alpha(\theta) \\
&= \frac{1}{\Lambda}\left(\frac{M\lambda_S}{2} + \frac{N\lambda_R}{2}\right)\begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} \\
&\quad + \frac{\mu}{\Lambda N}\begin{pmatrix} N(M-1) + N((1-\tilde{\nu})m_1 + \tilde{\nu} m_2) I_M & 0 \\ 0 & (N-1)M + M(\tilde{\nu} m_1 + (1-\tilde{\nu})m_2) I_N \end{pmatrix} \\
&= \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix} + \frac{\mu_\nu}{\Lambda N}\begin{pmatrix} N(m_2 - m_1) I_M & 0 \\ 0 & M(m_1 - m_2) I_N \end{pmatrix}
\end{aligned} \tag{7.27}$$

, where $\mu_\nu = \tilde{\nu}\mu$. Denote by $L(\nu_1, \nu_2)$ the $(N+M)\times(N+M)$ matrix

$$L(m_1, m_2) = \begin{pmatrix} m_1 I_M & 0 \\ 0 & m_2 I_N \end{pmatrix},$$

and set

$$\mathcal{P} = I_2 - \frac{\mu_\nu}{\Lambda N}\begin{pmatrix} N & -N \\ -M & M \end{pmatrix}.$$

Then (7.27) is recast as

$$\sum_\alpha \lambda_\alpha \int_{-\pi}^{\pi} \nu(\mathrm{d}\theta)\, r_\alpha(\theta)^{-1} L(m_1, m_2) r_\alpha(\theta) = L(m_1', m_2'), \tag{7.28}$$

where

$$\begin{pmatrix} m_1' \\ m_2' \end{pmatrix} = \mathcal{P}\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}.$$

By a repeated application of (7.28) we obtain

$$\begin{aligned}
&\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{[-\pi,\pi]^k} \nu(\mathrm{d}\theta_1)\cdots\nu(\mathrm{d}\theta_k) \left[\prod_{j=1}^k r_{\alpha_j}(\theta_j)\right]^T L(\underline{m}) \left[\prod_{j=1}^k r_{\alpha_j}(\theta_j)\right] \\
&\quad = L(\mathcal{P}^k \underline{m}).
\end{aligned}$$

Thus,

$$Z = \left(\mathcal{P}^k \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)_1 I_M.$$

It is easy to see that $\mathcal{P}$ has eigenvalues $\ell_1 = 1$ and $\ell_2 = 1 - \mu_\nu(M+N)/(\Lambda N)$ with eigenvectors $\underline{m}_1 = (1,1)$ and $\underline{m}_2 = (N, -M)^T/(M+N)$. Consequently,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{M}{N+M}\underline{m}_1 + \underline{m}_2,$$

which yields

$$\left(\mathscr{P}^k \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)_1 = \frac{M}{N+M} + \frac{N}{M+N}\left(1 - \mu_\nu \frac{M+N}{\Lambda N}\right)^k .$$

This proves Theorem 7.11. ∎

CHAPTER 8

Boltzmann-Kac collisions

In this chapter we show that the above results can also be extended, at least in a particular case, to three-dimensional Boltzmann-Kac collisions.

We again consider a system of M particles coupled to a reservoir consisting of N particles, but now with velocities $v_1, \dots, v_M$, $w_1, \dots, w_N \in \mathbb{R}^3$. The collisions between a pair of particles have to conserve energy and momentum,

$$
\begin{aligned}
z_i^2 + z_j^2 &= (z_i^*)^2 + (z_j^*)^2 \,, \\
z_i + z_j &= z_i^* + z_j^* \,,
\end{aligned}
$$

where z can be either the velocity of a system particle v or of a reservoir particle w. A convenient parametrization of the post-collisional velocities in terms of the velocities before the collision is given by

$$
\begin{aligned}
z_i^*(\omega) &= z_i - \omega \cdot (z_i - z_j)\,\omega \,, \\
z_j^*(\omega) &= z_j + \omega \cdot (z_i - z_j)\,\omega \,, \quad \text{where } \omega \in \mathbb{S}^2 \,.
\end{aligned}
$$

This is the so-called ω-representation. This representation is particularly useful, because the velocities are related to each other by a *linear* transformation, and the strategy used to prove the results for the one-dimensional Kac system carries over rather directly. The direction ω will be chosen according to the uniform probability distribution on the unit sphere $\mathbb{S}^2$.

Introduce the operators

$$
(R_{ij} f)(\mathbf{z}) = \int_{\mathbb{S}^2} f(r_{ij}(\omega)^{-1}\mathbf{z})\,\mathrm{d}\omega \,,
$$

where $\mathrm{d}\omega$ denotes the uniform probability measure on the sphere and the matrices $r_{ij}(\omega)$ are symmetric involutions acting as

$$
\begin{pmatrix} z_i^* \\ z_j^* \end{pmatrix} = \begin{pmatrix} I - \omega\omega^T & \omega\omega^T \\ \omega\omega^T & I - \omega\omega^T \end{pmatrix} \begin{pmatrix} z_i \\ z_j \end{pmatrix}
$$

on the velocities of the particles i and j, and as identities otherwise. They will replace the one-dimensional Kac collision operators in (7.3) in the otherwise unchanged generator of the time evolution. Notice that the matrices $r_{ij}(\omega)$ are orthogonal, so that the expansion formula (7.10) still holds with the obvious changes in the dimension of the single-particle spaces.

We prove an analogue of Theorem 7.1 for the case of three-dimensional Boltzmann-Kac collisions and pseudo-Maxwellian molecules.

Theorem 8.1. *Let $N \geq M$ and $F_0(v,w) = f_0(v)\,\mathrm{e}^{-\pi|w|^2}$ for some probability distribution f_0 on $\mathbb{R}^{3M}$. Then the entropy of the marginal*

$$f(v,t) := \int_{\mathbb{R}^{3N}} \left(\mathrm{e}^{\mathscr{L}t} F_0\right)(v,w)\,\mathrm{d}w$$

with respect to the thermal state $\mathrm{e}^{-\pi|v|^2}$ is bounded by

$$S(f(\cdot,t)) \leq \left[\frac{N}{N+M} + \frac{N}{N+M}\mathrm{e}^{-\frac{\mu}{3}\frac{N+M}{N}t}\right] S(f_0)\,.$$

Remark. The result in three dimensions is very similar to the case of one-dimensional Kac collisions, with the difference that the rate of exponential decay is $\mu/3$ instead of μ_ρ. The factor $1/3$ comes from the fact that $\int_{\mathbb{S}^2} \mathrm{d}\omega\,\omega\omega^T = I_3/3$. It would be interesting to cover the true Maxwellian molecules interaction

$$(R_{ij}f)(z) = \int_{\mathbb{S}^2} b\left(\frac{v_i - v_j}{|v_i - v_j|}\cdot\omega\right) f(r_{ij}(\omega)^{-1}z)\,\mathrm{d}\omega\,.$$

However, the dependence of the scattering rate b on the velocities does not seem to be treatable with the above methods. ♦

The proof of Theorem 8.1 essentially deviates from the one-dimensional case in only two places: the sum rule and the discrete approximation of the integrals. We begin by proving an analogue of Theorem 7.11. Most of the steps for the computation of the matrix Z in (7.26) are the same. What remains is to compute

$$Z := \sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k} \int_{\mathbb{S}^2\times\cdots\times\mathbb{S}^2} \mathrm{d}\omega_1\cdots\mathrm{d}\omega_k\, A_k(\underline{\alpha},\underline{\omega})A_k(\underline{\alpha},\underline{\omega})^T\,,$$

which is somewhat different for the case of Boltzmann-Kac collisions. Recall that $A_k(\underline{\alpha},\underline{\omega})$ is the upper left $3M\times 3M$ block of $[\prod_{j=1}^k r_{\alpha_j}(\omega_j)]^{-1}$, i.e.,

$$A_k(\underline{\alpha},\underline{\omega}) = P_{3M}[\Pi_{j=1}^k r_{\alpha_j}(\omega_j)]^{-1}P_{3M}^T$$

with the projection $P_{3M} = \begin{pmatrix} I_{3M} & 0\end{pmatrix}$ from $\mathbb{R}^{3M+3N}$ to $\mathbb{R}^{3M}$. In particular, by linearity,

$$Z = P_{3M}\left(\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1}\cdots\lambda_{\alpha_k}\int_{(\mathbb{S}^2)^k} \mathrm{d}\underline{\omega}\left[\prod_{j=1}^k r_{\alpha_j}(\omega_j)\right]^{-1}\begin{pmatrix} I_{3M} & 0\\ 0 & 0\end{pmatrix}\left[\prod_{j=1}^k r_{\alpha_j}(\omega_j)\right]\right)P_{3M}^T\,.$$

As in the proof of Theorem 7.11, we have

Lemma 8.2. *Let $\alpha, \beta \geq 0$. Then*

$$\sum_{1\leq i<j\leq M+N} \lambda_{ij} \int_{\mathbb{S}^2} \mathrm{d}\omega \, r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) = \begin{pmatrix} \alpha' I_{3M} & 0 \\ 0 & \beta' I_{3N} \end{pmatrix} ,$$

where α', β' are related to α, β by

$$\begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = \mathcal{P} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad \mathcal{P} = I_2 - \frac{\mu}{3\Lambda} \begin{pmatrix} 1 & -1 \\ -\frac{M}{N} & \frac{M}{N} \end{pmatrix} .$$

Notice that the matrix $\mathcal{P}$ of Lemma 8.2 has eigenvalues 1 and $1-\mu/(3\Lambda)\,(1 + M/N)$ with corresponding eigenvectors $(1 \quad 1)^T$ and $(-N/M \quad 1)^T$. Repeated application of Lemma 8.2 then implies, see also the argument in the one-dimensional case,

$$\sum_{\alpha_1,\ldots,\alpha_k} \lambda_{\alpha_1} \cdots \lambda_{\alpha_k} \int_{(\mathbb{S}^2)^k} \mathrm{d}\underline{\omega} \left[\prod_{j=1}^{k} r_{\alpha_j}(\omega_j) \right]^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} \left[\prod_{j=1}^{k} r_{\alpha_j}(\omega_j) \right]$$
$$= \begin{pmatrix} \alpha^{(k)} I_{3M} & 0 \\ 0 & \beta^{(k)} I_{3N} \end{pmatrix},$$

where

$$\begin{pmatrix} \alpha^{(k)} \\ \beta^{(k)} \end{pmatrix} = \mathcal{P}^k \begin{pmatrix} \alpha \\ \beta \end{pmatrix} .$$

Before we prove Lemma 8.2, let us make an easy observation.

Corollary 8.3. *In the particular case $\alpha = 1$, $\beta = 0$, we get*

$$Z = \left[\frac{M}{M+N} + \frac{N}{M+N} \left(1 - \frac{\mu}{3\Lambda} \left(1 + \frac{M}{N} \right) \right)^k \right] I_{3M} .$$

Proof of Lemma 8.2. For $1 \leq i < j \leq M$ (respectively for $M + 1 \leq i < j \leq M + N$) the operators $r_{ij}(\omega)$ only act non-trivially in the first $3M$ (last $3N$) variables. Taking into account that $r_{ij}(\omega)^{-1}\, I \, r_{ij}(\omega) = I$, we obtain

$$\frac{\lambda_S}{M-1} \sum_{1\leq i<j\leq M} \int_{\mathbb{S}^2} \mathrm{d}\omega \, r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) = \frac{M\lambda_S}{2} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix},$$

and

$$\frac{\lambda_R}{N-1} \sum_{M+1\leq i<j\leq M+N} \int_{\mathbb{S}^2} \mathrm{d}\omega \, r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) = \frac{N\lambda_R}{2} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix}.$$

It remains to look at the interaction terms $i = 1, \dots, M$ and $j = M+1, \dots, M+N$. Notice that

$$
\begin{aligned}
& r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) \\
&= \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} + \left(\begin{array}{ccc|ccc} 0 & & & & & \\ & (\beta-\alpha)\omega\omega^T & & & 0 & \\ & & 0 & & & \\ \hline & & & 0 & & \\ & 0 & & & (\beta-\alpha)\omega\omega^T & \\ & & & & & 0 \end{array}\right),
\end{aligned}
$$

where the non-zero entries in the second summand on the right-hand side correspond to the i^{th}, respectively j^{th}, 3×3 block on the diagonal. Since $\int_{\mathbb{S}^2} \mathrm{d}\omega\, \omega\omega^T = 1/3\, I_3$, we obtain

$$
\begin{aligned}
& \frac{\mu}{N} \sum_{i=1}^{M} \sum_{j=M+1}^{M+N} \int_{\mathbb{S}^2} \mathrm{d}\omega\, r_{ij}(\omega)^{-1} \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} r_{ij}(\omega) \\
&= \mu M \begin{pmatrix} \alpha I_{3M} & 0 \\ 0 & \beta I_{3N} \end{pmatrix} + \frac{\mu}{3}(\alpha - \beta) \begin{pmatrix} -I_{3M} & 0 \\ 0 & \frac{M}{N} I_{3N} \end{pmatrix} .
\end{aligned}
$$

Recall the definition of $\Lambda = M\lambda_S/2 + N\lambda_R/2 + \mu M$. Hence, summation of all the three contributions yields the statement of the Lemma. ■

As in the one-dimensional case, in order to apply the geometric Brascamp-Lieb inequality Theorem 7.9, we need to approximate the uniform probability measure $\mathrm{d}\omega$ on the sphere by a suitable sequence of discrete measures as in the one-dimensional case (see Lemma 7.10). Additionally, in each step of the discretization, the constraint $\int_{\mathbb{S}^2} \mathrm{d}\omega\, \omega\omega^T = 1/3 I$, has to hold. This is important because it guarantees that the geometric Brascamp-Lieb condition, i.e., the sum rule (7.19), holds in each step.

In order to find such an approximation, we parametrise the sphere in the usual way by spherical coordinates

$$
\omega = \omega(\theta, \varphi) = \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ \cos\theta \end{pmatrix}
$$

for $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi]$. For $K, L \in \mathbb{N}$ we introduce the measures

$$
\Phi_K := \frac{\pi}{K} \sum_{j=0}^{2K-1} \delta_{\frac{\pi}{K} j} \quad \text{on } [0, 2\pi], \quad \text{and}
$$

$$
\Theta_L := \sum_{i=1}^{L} \frac{2}{(1-u_i^2)^{3/2} (P_L'(u_i))^2} \delta_{\arccos u_i} \quad \text{on } [0, \pi],
$$

where P_L is the Legendre polynomial of order L on $[-1, 1]$, and u_i, $i = 1, \ldots, L$, are its zeros. Then, if $f \in \mathscr{C}[0, 2\pi]$ and $g \in \mathscr{C}[-1, 1]$,

$$\int_0^{2\pi} f(\varphi)\,\Phi_k(\mathrm{d}\varphi) = \frac{\pi}{K}\sum_{j=0}^{2K-1} f\left(\frac{\pi}{K}j\right) \to \int_0^{2\pi} f(\varphi)\,\mathrm{d}\varphi$$

as $K \longrightarrow \infty$ as Riemann sum. Furthermore,

$$\int_0^{\pi} g(\cos\theta)\,\sin\theta\,\Theta_L(\mathrm{d}\theta) = \sum_{i=1}^{L} \frac{2}{(1-u_i^2)(P_L'(u_i))^2} g(u_i)$$
$$\longrightarrow \int_{-1}^{1} g(u)\,\mathrm{d}u = \int_0^{\pi} g(\cos\theta)\,\sin\theta\,\mathrm{d}\theta$$

as $L \longrightarrow \infty$ by Gauss-Legendre quadrature. The latter approximation is exact for polynomials of order less or equal to $2L-1$. In particular, we have

$$\int_0^{\pi} \cos^2\theta\,\sin\theta\,\Theta_L(\mathrm{d}\theta) = \int_0^{\pi} \cos^2\theta\,\sin\theta\,\mathrm{d}\theta = \frac{2}{3}\,, \text{ and}$$
$$\int_0^{\pi} \sin^3\theta\,\Theta_L(\mathrm{d}\theta) = \int_0^{\pi} \sin^3\theta\,\mathrm{d}\theta = \frac{4}{3},$$

for all $L \geq 2$. It is easy to check that

$$\int_0^{2\pi} \sin\varphi\cos\varphi\,\Phi_k(\mathrm{d}\varphi) = 0,$$
$$\int_0^{2\pi} \sin\varphi\,\Phi_k(\mathrm{d}\varphi) = \int_0^{2\pi} \cos\varphi\Phi_k(\mathrm{d}\varphi) = 0,$$
$$\int_0^{2\pi} \sin^2\varphi\Phi_k(\mathrm{d}\varphi) = \int_0^{2\pi} \cos^2\varphi\,\Phi_k(\mathrm{d}\varphi) = \pi,$$

for all $K \geq 2$. Consequently,

$$\frac{1}{4\pi}\int_0^{2\pi} \omega(\theta,\varphi)\omega(\theta,\varphi)^T\,\Theta_L(\mathrm{d}\theta)\Phi_k(\mathrm{d}\varphi)$$
$$= \frac{1}{2K}\sum_{j=0}^{2K-1}\sum_{i=0}^{L} \frac{\omega\left(\arccos u_i, \pi j/K\right)\omega\left(\arccos u_i, \pi j/K\right)^T}{(1-u_i^2)(P_L'(u_i))^2} = \frac{1}{3}I_3$$

for all $K, L \geq 2$. It follows that Z is not changed by replacing the uniform measure on $\mathbb{S}^2$ by the above discrete approximation, in particular, Z is still proportional to the identity matrix, which guarantees the applicability of the geometric Brascamp-Lieb inequality.

This concludes the proof of Theorem 8.1. ■

APPENDIX E

The geometric Brascamp-Lieb inequality and the entropy inequality

In this section we prove Theorem 7.9. We use the same strategy as in [CLL04] and [BCCT08] which consists of transporting the functions f_i with the heat kernel in such a way that the right-hand side of (7.20) remains fixed while the left-hand side of that inequality increases. The results in [BCCT08] are quite general but for the special case in which the sum rule (7.19) holds, the proof is quite simple and this is one of the reasons why we include it here.

Proof of Theorem 7.9. The inequality (7.20) is equivalent to

$$\int_{\mathbb{R}^M} \prod_{i=1}^{K} f_i^{c_i}(B_i v)\,\mathrm{d}v \le \prod_{i=1}^{K} \left(\int_{H_i} f_i(u)\,\mathrm{d}u \right)^{c_i} . \tag{E.1}$$

This follows from the identity

$$\prod_{i=1}^{K} \left(\mathrm{e}^{-\pi |B_i v|^2} \right)^{c_i} = \mathrm{e}^{-\pi \sum_{i=1}^{K} (v\ c_i B_i^T B_i\ v)} = \mathrm{e}^{-\pi |v|^2} .$$

We transport the functions f_i by the heat flow, that is we define

$$f_i(B_i v, t) := \frac{1}{(4\pi t)^{M/2}} \int_{\mathbb{R}^M} \mathrm{e}^{-|v-w|^2/(4t)} f_i(B_i w)\mathrm{d}w . \tag{E.2}$$

For the above definition to make sense, we have to show that the right-hand side is a function of $B_i v$ alone. The condition $B_i B_i^T = I_{H_i}$ means that the matrix $P_i = B_i^T B_i$ is an orthogonal projection onto a d_i dimensional subspace of $\mathbb{R}^M$. Moreover, $B_i P_i =$

$I_{H_i} B_i = B_i$. We rewrite the integral (E.2) by splitting it in an integral over $w' \in \operatorname{Ran} P_i$ and one over integration over $w'' \in \operatorname{Ran} P_i^\perp$. Carrying out the integration over w'' we obtain

$$\begin{aligned}
f_i(B_i v, t) &= \frac{1}{(4\pi t)^{M/2}} \int_{\operatorname{Ran} P_i} \int_{\operatorname{Ran} P_i^\perp} \mathrm{e}^{-|(P_i v - P_i w')|^2/(4t)} \mathrm{e}^{-|(P_i^\perp v - w'')|^2/(4t)} f_i(B_i P_i w) \, \mathrm{d}w' \mathrm{d}w'' \\
&= \frac{1}{(4\pi t)^{d_i/2}} \int_{\operatorname{Ran} P_i} \mathrm{e}^{-|(P_i v - P_i w')|^2/(4t)} f_i(B_i P_i w') \, \mathrm{d}w' \\
&= \frac{1}{(4\pi t)^{d_i/2}} \int_{\operatorname{Ran} P_i} \mathrm{e}^{-|(B_i v - B_i w')|^2/(4t)} f_i(B_i w') \, \mathrm{d}w' \\
&= \frac{1}{(4\pi t)^{d_i/2}} \int_{H_i} \mathrm{e}^{-|(B_i v - u)|^2/(4t)} f_i(u) \, \mathrm{d}u \,,
\end{aligned}$$

where, in the last equality, we have used that B_i maps the range of P_i isometrically onto H_i. This justifies (E.2). The above computation also shows that

$$\int_{H_i} f_i(u, t) \mathrm{d}u = \int_{H_i} f_i(u) \mathrm{d}u$$

so that the right-hand side of the inequality (E.1) does not change under the heat flow.

We now show that the left-hand side of (E.1) is an increasing function of t. It is convenient to set $\phi_i(u, t) = \log f_i(u, t)$. Differentiating the function $\phi_i(B_i v, t)$ with respect to t yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi_i(B_i v, t) = \Delta_v \phi_i(B_i v, t) + |\nabla_v \phi_i(B_i v, t)|^2 \,.$$

Moreover,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^M} \prod_{i=1}^K f_i^{c_i}(B_i v, t) \, \mathrm{d}v \\
&= \sum_{m=1}^K c_m \int_{\mathbb{R}^M} [\Delta_v \phi_m(B_m v, t) + |\nabla_v \phi_m(B_m v, t)|^2] \prod_{i=1}^K f_i^{c_i}(B_i v, t) \, \mathrm{d}v \,.
\end{aligned}$$

Integrating by parts the term containing the Laplacian yields

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^M} \prod_{i=1}^K f_i^{c_i}(B_i v, t) \, \mathrm{d}v = \\
&\quad \sum_{m=1}^K c_m \int_{\mathbb{R}^M} |\nabla_v \phi_m(B_m v, t)|^2 \prod_{i=1}^K f_i^{c_i}(B_i v, t) \mathrm{d}v \\
&\qquad - \sum_{m,\ell=1}^K c_m c_\ell \int_{\mathbb{R}^M} \nabla_v \phi_m(B_m v, t) \cdot \nabla_v \phi_\ell(B_\ell v, t) \prod_{i=1}^K f_i^{c_i}(B_i v, t) \mathrm{d}v \,.
\end{aligned}$$

Finally, using that

$$\nabla_v \phi_m(B_m v, t) = B_i^T (\nabla \phi_m)(B_m v)$$

we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^M}\prod_{i=1}^K f_i^{c_i}(B_i v, t)\,\mathrm{d}v = \\ \sum_{m=1}^K c_m \int_{\mathbb{R}^M} |B_m^T(\nabla\phi_m)(B_m v, t)|^2 \prod_{i=1}^K f_i^{c_i}(B_i v, t)\mathrm{d}v \\ - \sum_{m,\ell=1}^K c_m c_\ell \int_{\mathbb{R}^M} B_m^T(\nabla\phi_m)(B_m v, t)\cdot B_\ell^T(\nabla\phi_\ell)(B_\ell v, t)\prod_{i=1}^K f_i^{c_i}(B_i v, t)\mathrm{d}v.\end{aligned}$$

We claim that this expression is non-negative. The vectors $\nabla\phi_m \in H_m$ are arbitrary and hence the problem is reduced to proving that for any set of vectors $V_m \in H_m$, $m = 1, \dots, K$, it holds

$$\sum_{m=1}^K c_m |B_m^T V_m|^2 - \sum_{m,\ell=1}^K c_m c_\ell B_m^T V_m \cdot B_\ell^T V_\ell \ \geq\ 0\,.$$

Recalling that $B_m B_m^T = I_{H_m}$ and setting $Y = \sum_\ell c_\ell B_\ell^T V_\ell$, we conclude that it is enough to show that

$$|Y|^2 \leq \sum_{m=1}^K c_m |V_m|^2\,.$$

This follows easily since, by applying Schwarz's inequality, we find that

$$|Y|^2 = \sum_{\ell=1}^K c_\ell Y\cdot B_\ell^T V_\ell = \sum_{\ell=1}^K c_\ell B_\ell Y\cdot V_\ell \leq \left(\sum_{\ell=1}^K c_\ell |B_\ell Y|^2\right)^{1/2}\left(\sum_{\ell=1}^K c_\ell |V_\ell|^2\right)^{1/2}.$$

Combining this with (7.19), we learn that

$$|Y|^2 \leq \left(Y\cdot\sum_{\ell=1}^K c_\ell B_\ell^T B_\ell Y\right)^{1/2}\left(\sum_{\ell=1}^K c_\ell |V_\ell|^2\right)^{1/2} = |Y|\left(\sum_{\ell=1}^K c_\ell |V_\ell|^2\right)^{1/2}.$$

Thus, we have that, when applying (E.1) to the functions $f_i(u, t)$, the left-hand side is an increasing function of t while the right-hand side does not depends on t. It is thus enough to show that the inequality holds for large t. Using once more the sum rule (7.19), we see that

$$
\begin{aligned}
&\int_{\mathbb{R}^M} \prod_{i=1}^{K} \frac{1}{(4\pi t)^{c_i d_i/2}} \left[\int_{H_i} e^{-\frac{|B_i v - u|^2}{4t}} f_i(u) du \right]^{c_i} dv \\
&= \frac{1}{(4\pi)^{M/2}} \int_{\mathbb{R}^M} \prod_{i=1}^{K} \left[\int_{H_i} e^{-\frac{|B_i v - t^{-1/2} u|^2}{4}} f_i(u) du \right]^{c_i} dv \\
&\xrightarrow{t\to\infty} \frac{1}{(4\pi)^{M/2}} \int_{\mathbb{R}^M} e^{-\frac{|v|^2}{4}} \prod_{i=1}^{K} \left[\int_{H_i} f_i(u) du \right]^{c_i} dv = \prod_{i=1}^{K} \left[\int_{H_i} f_i(u) du \right]^{c_i},
\end{aligned}
$$

which proves the first part of Theorem 7.9.

To prove the entropy inequality (7.21) we follow [CC09]. Let h be a non-negative function whose L^1 norm is one and whose entropy is finite. An elementary computation then shows that

$$
\begin{aligned}
&\int_{\mathbb{R}^M} h(v) \log h(v)\, e^{-\pi|v|^2} dv \\
&= \sup_{\Phi} \left\{ \int_{\mathbb{R}^M} h(v)\Phi(v)\, e^{-\pi|v|^2}\, dv - \log \int_{\mathbb{R}^M} e^{\Phi(v)} e^{-\pi|v|^2}\, dv \right\} .
\end{aligned}
$$

Now, we set

$$
\Phi(v) = \sum_{i=1}^{K} c_i \log f_i(B_i v) .
$$

This leads to the lower bound

$$
\begin{aligned}
&\int_{\mathbb{R}^M} h(v) \log h(v)\, e^{-\pi|v|^2}\, dv \\
&\geq \sum_{i=1}^{K} c_i \int_{\mathbb{R}^M} h(v) \log f_i(B_i v)\, e^{-\pi|v|^2}\, dv - \log \int_{\mathbb{R}^M} \prod_{i=1}^{K} f_i(B_i v)^{c_i}\, e^{-\pi|v|^2}\, dv \\
&\geq \sum_{i=1}^{K} c_i \int_{\mathbb{R}^M} h(v) \log f_i(B_i v)\, e^{-\pi|v|^2}\, dv - \log \left[\prod_{i=1}^{K} \left(\int_{H_i} f_i(u)\, e^{-\pi|u|^2}\, du \right)^{c_i} \right],
\end{aligned}
$$

where the second step is a consequence of the Brascamp-Lieb inequality (7.20). ■

Bibliography

[Bal89] K. Ball, Volumes of sections of cubes and related problems. In *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 1987–88*, Volume **1376** of the series *Lecture Notes in Mathematics*, pp 251–260. Springer Berlin Heidelberg, 1989. MR 1008726. Zbl 0674.46008. 143

[Bal91] K. Ball, Volume ratios and a reverse isoperimetric inequality. *Journal of the London Mathematical Society (Second Series)* **44** (1991), 351–359. MR 1136445. Zbl 0694.46010. 143

[Bar98] F. Barthe, On a reverse form of the Brascamp-Lieb inequality. *Inventiones mathematicae* **134** (1998), 335–361. MR 1650312. Zbl 0901.26010. 143

[Bar04] F. Barthe, A continuous version of the Brascamp-Lieb inequalities. In *Geometric Aspects of Functional Analysis: Israel Seminar 2002-2003*, Volume **1850** of the series *Lecture Notes in Mathematics*, pp 53–63. Springer Berlin Heidelberg, 2004. MR 2087150. Zbl 1084.26011. 144

[BCCT08] J. Bennett, A. Carbery, M. Christ, and T. Tao, The Brascamp-Lieb inequalities: finiteness, structure and extremals. *Geometric and Functional Analysis* **17** (2008), 1343–1415. MR 2377493. Zbl 1132.26006. 159

[BLTV17] F. Bonetto, M. Loss, H. Tossounian, and R. Vaidyanathan, Uniform approximation of a Maxwellian thermostat by finite reservoirs. *Communications in Mathematical Physics* **351** (2017), 311–339. MR 3613506. Zbl 06702028. 132, 133

[BLV14] F. Bonetto, M. Loss, and R. Vaidyanathan, The Kac model coupled to a thermostat. *Journal of Statistical Physics* **156** (2014), 647–667. MR 3226836. Zbl 1302.35290. 131, 132, 134

[BL76] H.J. Brascamp, and E.H. Lieb, Best constants in Young's inequality, its converse, and its generalization to more than three functions. *Advances in Mathematics* **20** (1976), 151–173. MR 0412366. Zbl 0339.26020. 135

[CCL00] E.A. Carlen, M.C. Carvalho, and M. Loss, Many-body aspects of approach to equilibrium. In *Journées "Équations aux Dérivées Partielles" (La Chapelle sur Erdre, 2000)*, Exp. No. XI, 12 pp., Université de Nantes, Nantes, 2000. MR 1775687. Zbl 1213.81230. 127, 128

[CCL03] E.A. Carlen, M.C. Carvalho, and M. Loss, Determination of the spectral gap for Kac's master equation and related stochastic evolution. *Acta Mathematica* **191** (2003), 1–54. MR 2020418. Zbl 1080.60091. 127, 128

[CCL11] E.A. Carlen, M.C. Carvalho, and M. Loss, Kinetic theory and the Kac master equation. In *Entropy and the quantum II*, Contemporary Mathematics **552** (2011), 1–20. American Mathematical Society, Providence, RI, 2011. MR 2868039. Zbl 1244.82069. 126

[CC09] E.A. Carlen, and D. Cordero-Erausquin, Subadditivity of the entropy and its relation to Brascamp-Lieb type inequalities. *Geometric and Functional Analysis* **19** (2009), 373–405. MR 2545242. Zbl 1231.26015. 143, 162

[CL90] E.A. Carlen, and M Loss, Extremals of functionals with competing symmetries. *Journal of Functional Analysis* **88** (1990), 437–456. MR 1038450. Zbl 0705.46016. 141

[CLL04] E.A. Carlen, E.H. Lieb, and M. Loss, A sharp analog of Young's inequality on S^N and related entropy inequalities. *The Journal of Geometric Analysis* **14** (2004), 487–520. MR 2077162. Zbl 1056.43002. 159

[Ein11] A. Einav, On Villani's conjecture concerning entropy production for the Kac master equation. *Kinetic and Related Models* **4** (2011), 479–497. MR 2786394. Zbl 1335.82023. 128

[Fed69] P. Federbush, Partially alternate derivation of a result of Nelson. *Journal of Mathematical Physics* **10** (1969), 50–52. Zbl 0165.58301. 141

[GTW95] E. Gabetta, G. Toscani, and B. Wennberg, Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation. *Journal of Statistical Physics* **81** (1995), 901–934. MR 1361302. Zbl 1081.82616. 133

[Gro75] L. Gross, Logarithmic Sobolev inequalities. *American Journal of Mathematics* **97** (1975), 1061–1083. MR 0420249. Zbl 0318.46049. 141

[Gro93] L. Gross, Logarithmic Sobolev inequalities and contractivity properties of semigroups. In *Dirichlet Forms (Varenna, 1992)*, Volume **1563** of the series *Lecture Notes in Mathematics*, pp 54–88. Springer Berlin Heidelberg, 1993. MR 1292277. Zbl 0812.47037. 141

[Han78] T.S. Han, Nonnegative entropy measures of multivariate symmetric correlations. *Information and Control* **36** (1978), 133–156. MR 0464499. Zbl 0367.94041. 142

[Jan01] E. Janvresse, Spectral gap for Kac's model of Boltzmann equation. *The Annals of Probability* **29** (2001), 288–304. MR 1825150. Zbl 1034.82049. 127

[Kac56] M. Kac, Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pp 171–197. University of California Press, Berkeley and Los Angeles, 1956. MR 0084985. Zbl 0072.42802. 125, 127

[Lie90] E.H. Lieb, Gaussian kernels have only Gaussian maximizers. *Inventiones mathematicae* **102** (1990), 179–208. MR 1069246. Zbl 0726.42005. 135

[LW49] L.H. Loomis, and H. Whitney, An inequality related to the isoperimetric inequality. *Bulletin of the American Mathematical Society* **55** (1949), 961–962. MR 0031538. Zbl 0035.38302. 142

[Mas03] D.K. Maslen, The eigenvalues of Kac's master equation. *Mathematische Zeitschrift* **243** (2003), 291–331. MR 1961868. Zbl 1016.82016. 127

[MM11] S. Mischler, and C. Mouhot, About Kac's program in kinetic theory. *Comptes Rendus Mathématique (Académie des Sciences Paris)* **349** (2011), 1245–1250. MR 2861993. Zbl 1233.35153. 129

[MM13] S. Mischler, and C. Mouhot, Kac's program in kinetic theory. *Inventiones mathematicae* **193** (2013), 1–147. MR 3069113. Zbl 1274.82048. 129

[Nel73] E. Nelson, The free Markoff field. *Journal of Functional Analysis* **12** (1973), 211–227. MR 0343816. Zbl 0273.60079. 141

[TV15] H. Tossounian, and R. Vaidyanathan, Partially thermostated Kac model. *Journal of Mathematical Physics* **56** (2015), 083301 [16 pages]. MR 3455336. Zbl 1330.82036. 134

[Vil03] C. Villani, Cercignani's conjecture is sometimes true and always almost true. *Communications in Mathematical Physics* **234** (2003), 455–490. MR 1964379. Zbl 1041.82018. 128

Part III

Dispersion Management Solitons

CHAPTER 9

Introduction

The dispersion managed nonlinear Schrödinger equation (DM NLS) is by now a well-established model in nonlinear science, see, for instance, Turitsyn-Brandon-Fedoruk [TBF12] for a good review of the subject. Initially, the main motivation to study this equation came from fibre optics applications, after the introduction of the dispersion compensation technique, which itself appeared due to the invention of fibres with anomalous dispersion and vastly increased the transmission speed and capacity of optical fibre communications systems.[1] Nowadays, the DM NLS has become a paradigm of a nonlinear dispersive equation with periodically varying coefficients that in some regime, e.g. strong dispersion management, leads to a dispersion averaged nonlinearity. This nonlocal equation and its solutions can easily have properties that are qualitatively different from what one is used to from the local NLS. For example, it may have ground states which have strongly oscillating tails, see [Lus04].

One should also note an interesting related development in pure mathematics where several works have appeared on best constants in space-time inequalities, such as the celebrated Strichartz inequality [Bul10, Car09, HZ06, FVV12, Kun03, Fos07], which are related to dispersion managed solitons.

9.1 Nonlinear Optics Background

The evolutionary equation for the propagation of the wave envelope of an optical pulse in a single mode fibre is given by [GT96, AB98]

$$i\partial_t u = -d(t)\,\partial_x^2 u - p(|u|)u, \tag{9.1}$$

[1] For the practical implementation of the dispersion management technique, Andrew Chraplyvy and Robert Tkach (Bell Labs) were awarded the Marconi Prize in 2009 [http://marconisociety.org/fellows/].

where the dispersion $d(t) = \epsilon^{-1} d_0(t/\epsilon) + d_{\text{av}}$ is parametrically modulated. The constant part of the group velocity dispersion d_{av}, assumed to be nonnegative $d_{\text{av}} \geq 0$, denotes the average component (residual dispersion), and d_0 its L-periodic mean zero part. The most basic example is $d_0 = \mathbb{1}_{[0,1)} - \mathbb{1}_{[1,2)}$ for $L = 2$, that is, the dispersion is constant along parts of the fibre and changes sign periodically. $P(u) = p(|u|)u$ describes the nonlinear interaction due to the polarisability of the optical fibre.

Notice that in the above equation, t denotes the distance along the fibre, and x is the retarded time. We ignored attenuation and amplification effects which can be transformed out by an appropriate change of variables.

Let $T_r = \mathrm{e}^{\mathrm{i}r\partial_x^2}$ be the free Schrödinger evolution in one space dimension and write $u = T_{D(t/\epsilon)} v$, with $D(t) := \int_0^t d_0(r)\,\mathrm{d}r$. Then (9.1) is equivalent to

$$\mathrm{i}\partial_t v = -d_{\text{av}}\,\partial_x^2 v - T_{D(t/\epsilon)}^{-1}\left[P(T_{D(t/\epsilon)} v)\right].$$

In the limit of small ϵ, and averaging over the fast dispersion action, one obtains an averaged dispersion managed nonlinear Schrödinger equation[2],

$$\begin{aligned}\mathrm{i}\partial_\tau v &= -d_{\text{av}}\,\partial_x^2 v - \frac{1}{L}\int_0^L T_{D(r)}^{-1}\left[P(T_{D(r)} v)\right]\,\mathrm{d}r \\ &= -d_{\text{av}}\,\partial_x^2 v - \int_{\mathbb{R}} T_r^{-1}\left[P(T_r v)\right]\,\mu(\mathrm{d}r),\end{aligned} \tag{9.2}$$

where μ is the image of the uniform measure on $[0, L]$ under D,

$$\mu(B) := \frac{1}{L}\int_0^L \mathbb{1}_B(D(r))\,\mathrm{d}r,$$

for any Lebesgue measurable set $B \subset \mathbb{R}$. In the model example $d_0 = \mathbb{1}_{[0,1)} - \mathbb{1}_{[1,2)}$ from above, the measure μ has the density $\psi = \mathbb{1}_{[0,1]}$ with respect to Lebesgue measure on $\mathbb{R}$. More generally, under physically reasonable assumptions, the probability measure μ has compact support and is absolutely continuous with respect to Lebesgue measure, with density ψ in suitable L^p spaces:

Lemma 9.1 (Lemma 1.4 in [HL12]). *Assume that the dispersion profile d_0 is locally integrable. Then the following holds:*

(i) The probability measure μ has compact support.

(ii) If the set $\{d_0 = 0\}$ has zero Lebesgue measure, then μ is absolutely continuous with respect to Lebesgue measure.

(iii) If furthermore d_0 changes sign finitely many times on $[0, L]$ and is bounded away from zero, then μ has a bounded density ψ.

[2] The solutions of the averaged equation and of the original one turn out to be ϵ close on a time scale of order ϵ^{-1}. This can be shown by developing an appropriate averaging theory similar to [ZGJT01], but we do not pursue that direction here.

(iv) Moreover, if d_0 changes sign finitely many times on $[0, L]$ and for some $p > 1$

$$\int_0^L |d_0(s)|^{1-p}\, ds < \infty,$$

then μ has a density $\psi \in L^p$. More precisely, we have the bound[3]

$$\|\psi\|_{L^p} \lesssim \left(\int_0^L |d_0(s)|^{1-p}\, ds \right)^{\frac{1}{p}}, \tag{9.3}$$

where the implicit constant depends only on the number of sign changes of d_0 and the period L.

For more details, we refer to the discussion in [HL12, CHL17]. From now on, we will always assume that $\mu(\mathrm{d}r) = \psi(r)\,\mathrm{d}r$ with some compactly supported probability density $\psi \in L^p(\mathbb{R})$ for some suitable $1 \le p \le \infty$.

The averaged equation (9.2) was first obtained by Gabitov and Turitsyn [GT96] and afterwards systematically derived and solved numerically by Ablowitz and Biondini [AB98].

Standing wave solutions of the averaged DM NLS of the form $v(t, x) = \mathrm{e}^{-\mathrm{i}\omega t} f(x)$ are solutions of the nonlinear and nonlocal eigenvalue equation (dispersion management equation)

$$\omega f = -d_{\mathrm{av}} f'' - \int_{\mathbb{R}} T_r^{-1} \left[P(T_r f) \right] \mu(\mathrm{d}r). \tag{9.4}$$

The DM equation (9.4) is of variational type, as it can be considered as an Euler-Lagrange equation for an appropriate variational principle and weak solutions can be found as stationary points of a suitable energy functional, see (9.5).

We remark that one can also apply the same ideas and methods from dispersion management to the case of waveguide arrays, which are modelled by a discrete nonlinear Schrödinger equation. This was proposed in [ESMBA98, ESMA00] and the effective equation, the diffraction managed discrete nonlinear Schrödinger equation (DM DNLS), governing the regime of strong diffraction management was derived in [AM01]. It is given by

$$\omega u = -d_{\mathrm{av}} \Delta_{\mathrm{disc}} u - \int_{\mathbb{R}} T_r^{-1} \left[P(|T_r u|) \right] \mu(\mathrm{d}r),$$

on the sequence space $\ell^2(\mathbb{Z})$, and with x indexing the position of the waveguide. Here, Δ_{disc} is the discrete Laplacian, acting as $\Delta_{\mathrm{disc}} u(x) = u(x+1) - 2u(x) + u(x-1)$, and the solution operator $T_r = \mathrm{e}^{\mathrm{i} r \Delta_{\mathrm{disc}}}$ now corresponds to the discrete free Schrödinger equation.

[3]Here and in the following, we use the notation $f \lesssim g$ if there exists a finite constant $C > 0$ such that $f \le C g$.

9.2 Variational formulation

In order to find solutions of the DM equation (9.4), we study the existence of minimisers for the nonlinear and nonlocal variational problems

$$E_\lambda^{d_{\mathrm{av}}} := \inf_{f\in\mathcal{S}_\lambda^{d_{\mathrm{av}}}} H(f), \tag{9.5}$$

where

$$\mathcal{S}_\lambda^{d_{\mathrm{av}}} = \left\{u \in H^1(\mathbb{R};\mathbb{C}) : \|u\|_2^2 = \lambda\right\} \quad \text{for} \quad \lambda > 0,\ d_{\mathrm{av}} > 0,$$
$$\mathcal{S}_\lambda^{0} = \left\{u \in L^2(\mathbb{R};\mathbb{C}) : \|u\|_2^2 = \lambda\right\} \quad \text{for} \quad \lambda > 0,\ d_{\mathrm{av}} = 0,$$

and where $L^2(\mathbb{R};\mathbb{C})$, respectively $H^1(\mathbb{R};\mathbb{C})$ are the standard L^2, respectively Sobolev spaces for complex-valued functions. We will denote the usual inner product on $L^2(\mathbb{R};\mathbb{C})$ by $\langle f, g\rangle = \int_{\mathbb{R}} \overline{f}\, g\,\mathrm{d}x$, but consider $L^2(\mathbb{R};\mathbb{C})$ as *real* Hilbert space with the inner product $\mathrm{Re}\langle\cdot,\cdot\rangle$, which induces the same topology on $L^2(\mathbb{R};\mathbb{C})$. $H^1(\mathbb{R};\mathbb{C})$ will be equipped with the corresponding L^2-inner product.

We will also use the standard L^p norms denoted by $\|\cdot\|_p$ in this part of the thesis. The energy functional is given by

$$H(f) := \frac{d_{\mathrm{av}}}{2}\|f'\|_2^2 - N(f) \tag{9.6}$$

for $f \in \mathcal{H}^{d_{\mathrm{av}}}$, where we set $\mathcal{H}^{d_{\mathrm{av}}} = H^1(\mathbb{R};\mathbb{C})$ if $d_{\mathrm{av}} > 0$ and $\mathcal{H}^0 = L^2(\mathbb{R};\mathbb{C})$ for convenience. The nonlocal nonlinearity is given by

$$N(f) := \iint_{\mathbb{R}^2} V(|T_r f(x)|)\,\mathrm{d}x\,\psi(r)\,\mathrm{d}r \tag{9.7}$$

for some suitable *nonlinearity potential* $V : [0,\infty) \to \mathbb{R}$. Recall that ψ is the density of some compactly supported probability measure, and will be assumed to lie in appropriate L^p spaces.

Under rather general assumptions on the nonlinearity potential V it was recently shown in [CHL17] that there is a threshold for the existence of dispersion managed solitons. More precisely, assume that V satisfies

(CHL1) V is continuously differentiable on $(0,\infty)$ and continuous on $[0,\infty)$ with $V(0) = 0$. There exist $2 \le \gamma_1 \le \gamma_2 < \infty$ such that

$$|V'(a)| \lesssim a^{\gamma_1 - 1} + a^{\gamma_2 - 1} \quad \text{for all } a > 0.$$

(CHL2) There exists $\gamma_0 > 2$ such that

$$V'(a)a \ge \gamma_0 V(a) \quad \text{for all } a > 0.$$

(CHL3) There exists $a_* > 0$ such that $V(a_*) > 0$.

(CHL4) If $d_{\mathrm{av}} > 0$, we assume that there exist $\varepsilon > 0$ and $2 < \kappa_0 < 6$ such that

$$V(a) \gtrsim a^{\kappa_0} \qquad \text{for all } 0 < a \leq \varepsilon.$$

If $d_{\mathrm{av}} = 0$, we assume that there exists $\varepsilon > 0$ such that $V(a) > 0$ for all $0 < a \leq \varepsilon$.

Then the following holds:

Theorem 9.2 (Thresholds for existence for positive average dispersion, Theorem 1.2 in [CHL17]). *Assume $d_{\mathrm{av}} > 0$, V obeys Assumptions* **(CHL1)** *through* **(CHL3)** *for some $2 < \gamma_1 \leq \gamma_2 < 10$, and $\psi \in L^{\alpha_\delta}$ has compact support for some $\delta > 0$, where $\alpha_\delta := \alpha_\delta(\gamma_2) := \max\{1, \frac{4}{10-\gamma_2} + \delta\}$. Then*

(i) There exists a threshold $0 \leq \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}} < \infty$ such that $E_\lambda^{d_{\mathrm{av}}} = 0$ for $0 < \lambda < \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$ and $-\infty < E_\lambda^{d_{\mathrm{av}}} < 0$ for $\lambda > \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$.

(ii) If $0 < \lambda < \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$, then no minimiser for the constrained minimisation problem (9.5) exists. If $\gamma_1 \geq 6$, then $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}} > 0$.

(iii) If $\lambda > \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$, then any minimising sequence for (9.5) is up to translations relatively compact in $L^2(\mathbb{R})$. In particular, there exists a minimiser for (9.5). This minimiser is also a weak solution of the dispersion management equation (9.4) for some Lagrange multiplier $\omega < 2E_\lambda^{d_{\mathrm{av}}}/\lambda < 0$.

(iv) If V obeys in addition **(CHL4)**, *then $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}} = 0$.*

Theorem 9.3 (Threshold for existence for zero average dispersion, Theorem 1.4 in [CHL17]). *Assume $d_{\mathrm{av}} = 0$ and V obeys Assumptions* **(CHL1)** *through* **(CHL3)** *with $2 < \gamma_1 \leq \gamma_2 < 6$, and that the density ψ has compact support and $\psi \in L^{\frac{4}{6-\gamma_2}+\delta}$ for some $\delta > 0$. Then*

(i) There exists a threshold $0 \leq \lambda_{\mathrm{cr}}^0 < \infty$ such that $E_\lambda^0 = 0$ for $0 < \lambda < \lambda_{\mathrm{cr}}^0$ and $-\infty < E_\lambda^0 < 0$ for $\lambda > \lambda_{\mathrm{cr}}^0$.

(ii) If $\lambda > \lambda_{\mathrm{cr}}^0$, then any minimising sequence for (9.5) is up to translations and boosts, that is, translations in Fourier space, relatively compact in $L^2(\mathbb{R})$. In particular, there exists a minimiser for (9.5). This minimiser is also a weak solution of the dispersion management equation (9.4) for some Lagrange multiplier $\omega < 2E_\lambda^0/\lambda < 0$.

(iii) If V obeys in addition **(CHL4)**, *then $\lambda_{\mathrm{cr}}^0 = 0$.*

While the energy functional H is related to the standard NLS functional, the fact that the nonlinearity is averaged over the dispersion action produces several nice properties, one of which is that ground states can exist even in the absence of the gradient term ($d_{\mathrm{av}} = 0$).

However, the energy functional (9.6) has a lot of symmetries, which makes it difficult to establish the existence of minimisers. Indeed, for $d_{\mathrm{av}} > 0$, it is invariant under translations $f \mapsto f(\cdot + y)$, $y \in \mathbb{R}$. For $d_{\mathrm{av}} = 0$, due to the absence of the derivative term, it is invariant under translations and boosts $f \mapsto \mathrm{e}^{\mathrm{i}(\cdot)\xi} f$, $\xi \in \mathbb{R}$, that is, translations in Fourier space.

One of the key properties to assure the presence of the ground state, see Theorems 9.2 and 9.3, is that the nonlinearity potential V is "sufficiently" nonlinear. More precisely, one needs to verify the *strict sub-additivity* condition

$$E^{d_{\mathrm{av}}}_{\lambda_1+\lambda_2} < E^{d_{\mathrm{av}}}_{\lambda_1} + E^{d_{\mathrm{av}}}_{\lambda_2},$$

whenever the energy is *strictly negative*. Heuristically speaking, this means that if a hypothetical ground state is split into two parts, preserving the total energy, and these two parts are moved infinitely far away from each other, then the value of the Hamiltonian will increase. In other words, it excludes splitting of a minimising sequence into two parts of positive L^2 mass running away from another, thus restoring (pre-)compactness (modulo the symmetries) of the problem.

Strict sub-additivity in [CHL17] is guaranteed by the Ambrosetti-Rabinowitz type Assumption **(CHL2)**. One important case when the sub-additivity condition does not hold uniformly is the so-called saturated nonlinearity, such as

$$p(|u|) = \frac{|u|^2}{1 + \sigma |u|^2}.$$

This function approaches a constant for large values of $|u|$ and, as a result, the nonlinear term degenerates into a linear one.

9.3 Saturated Nonlinearities

In this part of the thesis we extend the above existence results, Theorems 9.2 and 9.3, to *saturating nonlinearities*, which violate the Ambrosetti-Rabinowitz condition **(CHL2)** in the sense that the variational problem becomes asymptotically quadratic.

From a physics viewpoint, saturated nonlinearities are relevant in modelling optical waves in nonlinear materials, as the nonlinear interaction due to the polarisability of the medium, which is cubic for low intensities (Kerr nonlinearity), saturates for large fields and approaches a regime with constant refraction index. One needs to assume some form of saturation of the nonlinearity $P(u) = p(|u|)u$, with the most common law being

$$p(|u|) = \frac{|u|^2}{1 + \sigma |u|^2},$$

where p corresponds to the intensity dependent refraction coefficient. In such models, the corresponding term in the energy functional is given by

$$V(a) = \frac{a^2}{2\sigma} - \frac{1}{2\sigma^2} \log(1 + \sigma a^2),$$

where $V'(a) = P(a)$. Another natural modification with similar behaviour is given by the nonlinearity potential $V(a) = a^4/(1 + \sigma a^2)$. We will actually consider a much broader class of saturated nonlinearity potentials which includes the above two as very special cases, but first we discuss the local saturated NLS.

The presence of solitary waves in the local NLS with saturated nonlinearity has been addressed in several studies, see e.g. Gatz-Herrmann [GH91], Usman-Osman-Tilley [UOT98] and references therein. Their results show that solitary wave solutions can be obtained numerically and sometimes analytically using phase space analysis, and one may also observe bistable (two-state) solitons.

Here, we address the question of existence of at least one solitary wave for the nonlocal NLS. By the variational methods we use, the obtained solitary wave is automatically a ground state solution. It is anticipated that multiple solitary waves may also exist in the nonlocal case, but one needs to use different methods to address this question.

While in the local case a saturable nonlinearity is often helpful by creating more favourable conditions for the existence of ground states, e.g., by arresting collapse in the supercritical regime, in the nonlocal case saturation presents difficulties in satisfying the sub-additvity condition.

The direct application of the approach of [CHL17] to saturating nonlinearities does not work because of the lack of strict sub-additivity of the energy in this case. To overcome this difficulty, the main idea is to construct a modified minimising sequence with a uniform L^∞ bound, which prevents the minimising sequence from reaching the saturation regime.

Saturable nonlinearities have also been considered in the context of coupled local nonlinear Schrödinger systems, where the existence of stationary solutions was proved by variational and bifurcation techniques, see [dAMP13, JT02, Man16].

While saturated nonlinearities are well-studied in the physics literature for the regular NLS [GH91], the DM NLS with saturated nonlinearities has not received much attention. This is perhaps due to the small values of optical power in fibre optics applications, which suggests that saturation effects are negligible. Nevertheless, theoretically speaking, it leaves an open question whether one can still construct ground states. The task is especially delicate in the case of zero average dispersion, as we explain below. The reader will also see that our argument points to some possible limitations when ground states may fail to exist.

Main results

For our main results, we shall make the following assumptions on the nonlinearity potential V:

(A1) V is continuously differentiable on $(0, \infty)$ and continuous on $[0, \infty)$ with $V(0) = 0$. There exist $2 \le \gamma_1 \le \gamma_2 < \infty$ such that

$$|V'(a)| \lesssim a^{\gamma_1 - 1} + a^{\gamma_2 - 1} \quad \text{for all} \quad a > 0.$$

(**A2**) There exists a continuous function $\kappa : [0,\infty) \to [2,\infty)$ with $\kappa > 2$ on *compact* intervals, such that for all $a > 0$,

$$V'(a)a \geq \kappa(a)V(a).$$

(**A3**) There exists $a_* > 0$ such that $V(a_*) > 0$.

Remark. Assumption (**A2**) allows for *saturation* of the potential V in the sense that $\frac{V'(a)a}{V(a)} \to 2$ as $a \to \infty$. This is in contrast to the typically assumed Ambrosetti-Rabinowitz condition [AR73]

$$V'(a)a \geq \kappa V(a) \quad \text{for all} \quad a > 0 \tag{9.8}$$

with $\kappa > 2$, which fails in the limit $a \to \infty$ for the saturated nonlinearities. In [CHL17], the Ambrosetti-Rabinowitz condition (9.8) was crucial in proving strict sub-additivity of the variational problem. ♦

Under Assumptions (**A1**)–(**A3**), with appropriate restrictions on γ_1, γ_2, we can show that there exists a threshold, that is, a critical optical power $\lambda_{\text{cr}}^{d_{\text{av}}}$, for the existence of minimisers. Under the additional assumptions

(**A4**) $d_{\text{av}} = 0$: There exists $\epsilon > 0$ such that $V(a) > 0$ for all $0 < a \leq \epsilon$.

$d_{\text{av}} > 0$: There exist $\epsilon > 0$ and $2 < \gamma_0 < 6$ such that $V(a) \gtrsim a^{\gamma_0}$ for all $0 < a \leq \epsilon$.

on V, minimisers are shown to exist for any $\lambda > 0$.

Theorem 9.4 (Existence of DM solitons for zero average dispersion)**.** *Let $d_{\text{av}} = 0$. Assume that V satisfies the conditions* (**A1**)–(**A3**)*, with $3 \leq \gamma_1 \leq \gamma_2 < 5$. Assume further that $\psi \geq 0$ is compactly supported and $\psi \in L^{\frac{4}{5-\gamma_2}+\delta}$ for some $\delta > 0$.*

Then there exists a threshold $\lambda_{\text{cr}}^0 \geq 0$ such that

1. *if $0 < \lambda < \lambda_{\text{cr}}^0$, then $E_\lambda^0 = 0$,*

2. *if $\lambda > \lambda_{\text{cr}}^0$, then $-\infty < E_\lambda^0 < 0$ and there exists a minimiser $u \in \mathcal{S}_\lambda^0 \cap L^\infty$ of the variational problem* (9.5)*. This minimiser is a weak solution of the dispersion management equation*

$$\omega f = -\int_{\mathbb{R}} T_r^{-1}\left[V'(|T_r f|)\frac{T_r f}{|T_r f|}\right] \psi \, \mathrm{d}r \tag{9.9}$$

for some Lagrange multiplier $\omega < \frac{2E_\lambda^0}{\lambda} < 0$.

If, in addition, assumption (**A4**) *holds, then $\lambda_{\text{cr}}^0 = 0$.*

Theorem 9.5 (Existence of DM solitons for positive average dispersion)**.** *Let $d_{\mathrm{av}} > 0$.*

Assume that V satisfies the conditions (**A1**)–(**A3**), *with $2 \le \gamma_1 \le \gamma_2 < 10$. Assume further that $\psi \ge 0$ is compactly supported, with $\psi \in L^{a_\delta}$ for some $\delta > 0$, where $a_\delta := \max\left\{1, \frac{4}{10-\gamma_2} + \delta\right\} < \infty$.*

Then there exists a threshold $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}} \ge 0$ such that

1. *if $0 < \lambda < \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$, then $E_\lambda^{d_{\mathrm{av}}} = 0$ and there exists no minimiser for the variational problem* (9.5),

2. *if $\lambda > \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$, then $-\infty < E_\lambda^{d_{\mathrm{av}}} < 0$ and there exists a minimiser $u \in \mathcal{S}_\lambda^{d_{\mathrm{av}}}$ of the variational problem* (9.5)*. This minimiser is a weak solution of the dispersion management equation*

$$\omega f = -d_{\mathrm{av}} f'' - \int_{\mathbb{R}} T_r^{-1}\left[V'(|T_r f|)\frac{T_r f}{|T_r f|}\right]\psi\,\mathrm{d}r \tag{9.10}$$

for some Lagrange multiplier $\omega < \frac{2E_\lambda^{d_{\mathrm{av}}}}{\lambda} < 0$.

If, in addition, assumption (**A4**) *holds, then $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}} = 0$.*

Remark. If we assume that there exists $\gamma \ge 2$ such that

$$|V'(|z+w|) - V'(|z|)| \lesssim (|w| + |z|)^{\gamma-2}|w| \quad \text{for all} \quad w, z \in \mathbb{C}, \tag{9.11}$$

the nonlinearity $N : \mathcal{H}^{d_{\mathrm{av}}} \to \mathbb{R}$ is actually a $\mathcal{C}^1$ functional, see Proposition 11.4. We can work with directional derivatives only though, including the construction of the modified minimising sequence, so this assumption is not needed for our main theorems. ♦

The main ingredient in the proof of existence of minimisers of (9.5) is the construction of a minimising sequence which satisfies an additional uniform L^∞ bound. This prevents the minimising sequence from reaching the asymptotic regime, where strict sub-additivity would fail.

While for positive average dispersion $d_{\mathrm{av}} > 0$, the uniform L^∞ bound is readily provided by the Sobolev embedding $H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$, some work has to be done in the setting of zero average dispersion. More precisely, we will construct a modified minimising sequence via Ekeland's variational principle, which provides an approximate solution of the DM equation, combined with dispersive estimates on the gradient of the nonlinearity.

Remark. In Theorems 9.2 and 9.3, which rely on a *strict Ambrosetti–Rabinowitz condition* and hence do not apply to saturated nonlinearities, it was shown that *any* minimising sequence is relatively compact in $L^2(\mathbb{R})$ modulo the natural symmetries if $\lambda > \lambda_{\mathrm{cr}}^0$. In Theorem 9.4 a slightly weaker result is proved: Given a minimising sequence, we can find a new minimising sequence that satisfies additional L^∞ bounds needed to guarantee strict sub-additivity, in particular, not every minimising sequence may be relatively compact in $L^2(\mathbb{R})$ modulo translations and boosts. This is irrelevant, however, for the existence of minimisers.

For positive average dispersion $d_{\text{av}} > 0$, due to the Sobolev embedding, any minimising sequence, being bounded in H^1, is also bounded in L^∞. So in this case any minimising sequence is relatively compact in L^2 modulo translations. ♦

In Chapter 11 below, we establish the existence of ground state solutions in a general class of averaged DM NLS with saturable nonlinearities. The existence of diffraction managed solitons and the necessary changes of our argument to cover the discrete case is addressed in the last section of Chapter 10.

CHAPTER 10

Existence of DM Solitons for Saturated Potentials: an easy example

Before we prove our main theorem regarding the existence of DM solitons in the presence of saturating nonlinearities, we turn to an easier example, which highlights some of the difficulties and shows how they can be overcome.

A natural strategy to establish the existence of a ground state is to show that there is a minimiser of the constrained variational principle by constructing a converging subsequence. One difficulty with the saturable nonlinearity is that the sub-additivity property does not hold in the saturation regime, that is, where the amplitude gets large. However, if the minimising sequence has bounded amplitude, which we show, then sub-additivity holds in the relevant region.

We modify our saturated nonlinearity in such a way that the approach of Choi-Hundertmark-Lee [CHL17] can be applied and establish a bound on the maximum of the ground state in the modified problem. Next we show that the ground state in the modified problem is also a ground state in the original problem.

10.1 Zero residual dispersion

We start with the case $d_{\mathrm{av}} = 0$ and consider the minimisation problem

$$E_\lambda^0 = \inf_{\|u\|_2^2=\lambda} \left\{ -\int_0^1 \int_{\mathbb{R}} F(|T_t u|)\mathrm{d}x\mathrm{d}t \right\},$$

where the nonlinearity potential F is assumed to satisfy

(E1) **(polynomial bound)** F is continuously differentiable on $(0, \infty)$ and continuous on $[0, \infty)$ with $F(0) = 0$. Further, F satisfies the inequality

$$F'(s) \le C(s^{\gamma_1 - 1} + s^{\gamma_2 - 1}),$$

where $3 \le \gamma_1 \le \gamma_2 < 5$ and $C > 0$ is a constant.

(E2) **(superquadratic growth)** For any $A > 0$, there exists γ_0 that can depend on A such that

$$W(s) := \frac{F'(s)s}{F(s)} \ge \gamma_0(A) > 2, \quad s \in (0, A).$$

(E3) **(saturation condition)** $W(s)$ is a monotonically decreasing function with limit

$$\lim_{s \to \infty} W(s) = 2.$$

(E4) **(positivity)** $F(s) > 0$ for any $s > 0$.

Remark. (i) The growth rate function $W(s)$ is an important measure of polynomial growth and $W(s) > 2$ implies locally faster than quadratic growth. The two saturated nonlinearities that we mentioned in the introduction satisfy these four conditions. One can also construct many more examples.

(ii) For positive residual dispersion (DM NLS) and arbitrary non-negative residual diffraction (DM DNLS), we can extend the range of the parameters in Assumption **(E1)** to $2 < \gamma_1 \le \gamma_2 < \infty$.

♦

Notice that we chose $\mu(\mathrm{d}r) = \mathbb{1}_{[0,1]}(r)\,\mathrm{d}r$ in this example, which corresponds to the dispersion profile $d_0 = \mathbb{1}_{[0,1)} - \mathbb{1}_{[1,2)}$. This is merely for convenience and clarity of presentation, as the arguments can easily be generalised to more general dispersion profiles.

10.2 The modified functional

Under Assumptions **(E1)** – **(E4)** we can show the existence of DM solitons by modifying the energy functional

$$N(u) = -\int_0^1 \int_{\mathbb{R}} F(|T_t u|)\,\mathrm{d}x\,\mathrm{d}t,$$

which is done in such a way that the Ambrosetti-Rabinowitz condition holds. Let $\mu > 0$ and set $\delta = W(\mu) - 2$, where W is the growth rate function introduced in condition **(E2)**. Choosing μ large enough, we can assume that $\delta \in (0, 1)$ by Assumption **(E3)**.

We now define the modified growth rate function W_m by

$$W_m(s) = \begin{cases} W(s), & s \in [0, \mu] \\ a(s-\sigma)^2 + 2 + \frac{\delta}{2}, & s \in [\mu, \sigma] \\ 2 + \frac{\delta}{2}, & s > \sigma, \end{cases}$$

where

$$a = \frac{W'(\mu)^2}{2(W(\mu)-2)}, \quad \sigma = \mu - \frac{W(\mu)-2}{W'(\mu)},$$

see Figure 10.1.

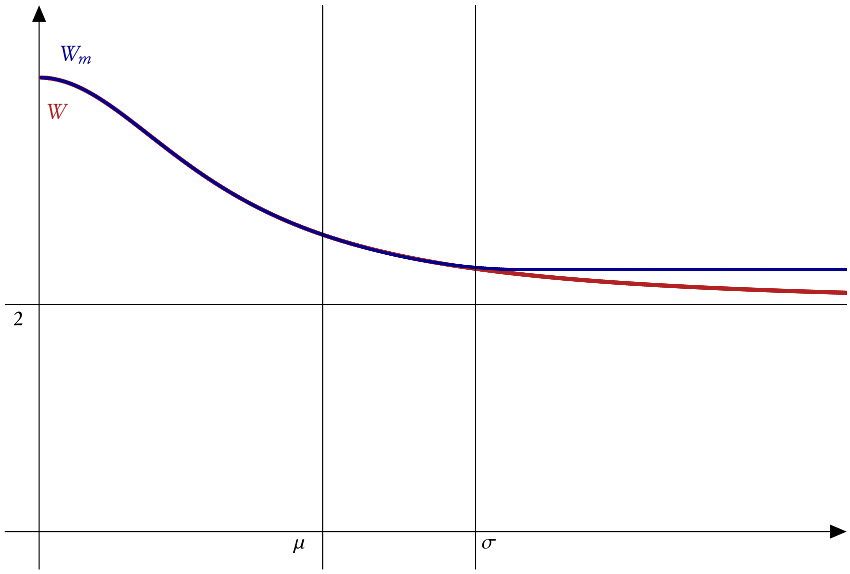

Figure 10.1: The modified growth rate function W_m.

The modified potential F_m is then given by

$$F_m(s) = F(s), \quad 0 \leq s \leq \mu,$$

and for $s > \mu$ as the solution of the ODE $W_m(s) = (\log F_m(s))' s$ with initial condition $F_m(\mu) = F(\mu)$. By construction, the modified potential $F_m(s)$ satisfies a super-quadratic growth condition

$$F_m'(s)s \geq \left(2 + \frac{\delta}{2}\right) F_m(s)$$

with $\delta = W(\mu) - 2 > 0$, but now *for all* $s > 0$. Together with the following lemma this shows that F_m satisfies the appropriate conditions **(CHL1)–(CHL4)** that guarantee the existence of ground states, see Theorem 9.3.

Lemma 10.1. *The modified potential satisfies the bounds*

$$F_m(s) \le \frac{C}{3}(s^{\gamma_1} + s^{\gamma_2})$$
$$F_m'(s) \le C(s^{\gamma_1 - 1} + s^{\gamma_2 - 1})$$

for $\mu > 0$, where C is the constant from Assumption (E1) on F.

Proof. First, by integrating the bound from Assumption (E1) from zero to $s \le \mu$, we have

$$F(s) \le \frac{C}{\gamma_1}s^{\gamma_1} + \frac{C}{\gamma_2}s^{\gamma_2} \le \frac{C}{3}\left(s^{\gamma_1} + s^{\gamma_2}\right)$$

for all $0 < s \le \mu$. Solving the ODE $(\log F_m(s))'s = W_m(s)$ for F_m, we obtain for $s \ge \mu$

$$\begin{aligned} F_m(s) &= F(\mu)\exp\left(\int_\mu^s \frac{W_m(\tau)}{\tau}\mathrm{d}\tau\right) \le F(\mu)\left(\frac{s}{\mu}\right)^{2+\delta} \le \frac{C}{3}(\mu^{\gamma_1} + \mu^{\gamma_2})\left(\frac{s}{\mu}\right)^{2+\delta} \\ &= \frac{C}{3}(\mu^{\gamma_1-\delta-2} + \mu^{\gamma_2-\delta-2})s^{2+\delta} \le \frac{C}{3}(s^{\gamma_1} + s^{\gamma_2}), \end{aligned}$$

and then using the defining differential equation for F_m, we also have

$$F_m'(s) = F_m(s)\cdot\frac{W_m(s)}{s} \le \frac{C}{3}(s^{\gamma_1} + s^{\gamma_2})\,\frac{2+\delta}{s} \le C(s^{\gamma_1-1} + s^{\gamma_2-1}),$$

which implies the result. ∎

The modified functional

$$H_m(u) = -\int_0^1\int_{\mathbb{R}} F_m(|T(t)u|)\mathrm{d}x\mathrm{d}t,$$

subject to the energy constraint, possesses a ground state u^* according to Theorem 9.3. This ground state depends on μ and λ, but we suppress this dependence in the following for simplicity of notation.

The ground state u^* must satisfy the corresponding Euler-Lagrange equation

$$\omega u^* = Q_m(u^*) := \int_0^1 T^{-1}(t)\left(F_m'(|T(t)u^*|)\frac{T(t)u^*}{|T(t)u^*|}\right)\mathrm{d}t.$$

Lemma 10.2. *The ground state u^* is uniformly bounded independently of μ, and, moreover, $|(T(r)u^*)(x)| \le K$ for all x and r for some constant $K < \infty$.*

Proof. First we show that the Lagrange multiplier $\omega = \omega(\mu) \ge c > 0$, independently of the modification. By multiplying the Euler-Lagrange equation with $\overline{u^*}$ and integrating, we have

$$\omega\int_{\mathbb{R}}|u^*|^2\mathrm{d}x = \int_{\mathbb{R}} Q_m(u^*)\overline{u^*}\,\mathrm{d}x = \int_0^1\int_{\mathbb{R}} F_m'(|T(t)u^*|)|T(t)u^*|\mathrm{d}x\mathrm{d}t.$$

To assure that the Lagrange multiplier ω is bounded away from zero, we look at the ratio

$$\omega = \frac{\int_{\mathbb{R}} Q_m(u^*)\overline{u^*}\mathrm{d}x}{\int_{\mathbb{R}} |u^*|^2\mathrm{d}x}. \tag{10.1}$$

The lower bound of the ratio depends on the energy $\int |u^*|^2\mathrm{d}x$ but it is fixed in the minimisation procedure. Considering the numerator, we see that

$$\begin{aligned}\int_{\mathbb{R}} Q_m(u^*)\overline{u^*}\mathrm{d}x &= \int_0^1 \int_{\mathbb{R}} F_m'(|T(t)u^*|)|T(t)u^*|\mathrm{d}x\mathrm{d}t \\ &\geq 2\int_0^1 \int_{\mathbb{R}} F_m(|T(t)u^*|)\mathrm{d}x\mathrm{d}t,\end{aligned}$$

where we used the superquadratic growth property of F_m, that is, $F_m'(s)s \geq (2 + \delta/2)F_m(s) \geq 2F_m(s)$. Since u^* is an energy minimiser, we have

$$\begin{aligned}\int_0^1 \int_{\mathbb{R}} F_m(|T(t)u^*|)\mathrm{d}x\mathrm{d}t &\geq \int_0^1 \int_{\mathbb{R}} F_m(|T(t)g|)\mathrm{d}x\mathrm{d}t \\ &\geq \int_0^1 \int_{\mathbb{R}} F(|T(t)g|)\mathrm{d}x\mathrm{d}t\end{aligned}$$

for any test function g, where we also used that $F_m(s) \geq F(s)$ for all $\mu > 1$ and $s \geq 0$. Choosing, for instance, g to be a Gaussian test function, we simply observe that the last integral is strictly positive (by Assumption (E4)), and independent of μ.

Next we show that $Q_m(u^*)$ is bounded by using an argument due to Kunze [Kun03]. First recall the well-known bound on the solution of the linear Schrödinger equation in one dimension

$$|T(t)u| \leq \frac{1}{|t|^{1/2}} \int_{\mathbb{R}} |u|\,\mathrm{d}x, \tag{10.2}$$

and consider

$$T(r)Q_m(u) = \int_0^1 T(r)T^{-1}(t)\left(F_m'(|T(t)u|)\frac{T(t)u}{|T(t)u|}\right)\mathrm{d}t.$$

Proposition 10.3. *Let $\nu \in [2, 4)$ and assume that $F'(s) \leq Cs^\nu$ for all $s \geq 0$ and some constant C. Then*

$$|T(r)Q_m(u)| \leq C(\lambda)$$

for all r, where $C(\lambda)$ is independent of the modification, but may depend on ν and the optical power $\lambda = \int_{\mathbb{R}} |u(x)|^2\mathrm{d}x$.

Proof. Using the dispersive estimate (10.2),

$$\begin{aligned}|T(r)Q_m(u)| &\le \int_0^1 \frac{1}{|r-t|^{1/2}} \|F_m'(|T(t)u|)\|_{L^1_x} \mathrm{d}t \le \int_0^1 \frac{C}{|r-t|^{1/2}} \||T(t)u|^\nu\|_{L^1_x} \mathrm{d}t \\ &= C\int_0^1 \frac{1}{|r-t|^{1/2}} \|T(t)u\|_{L^\nu_x}^\nu \mathrm{d}t.\end{aligned}$$

Using Hölder's inequality, we obtain

$$|T(r)Q_m(u)| \le C\left(\int_0^1 \frac{\mathrm{d}t}{|r-t|^{p/2}}\right)^{1/p} \cdot \left(\int_0^1 \|T(t)u\|_{L^\nu_x}^{\nu p'} \mathrm{d}t\right)^{1/p'},$$

where

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

To bound the first integral uniformly in r we need $p < 2$, i.e. $p' > 2$. For the convergence of the second integral, we use a general case of the Strichartz inequality [GV85] that we recall here:

$$\left(\int_{-\infty}^{+\infty} \|T(t)u\|_{L^\sigma_x}^\rho \mathrm{d}t\right)^{1/\rho} \le S_\sigma \|u\|_{L^2}$$

as long as

$$\frac{1}{\sigma} + \frac{2}{\rho} = \frac{1}{2},$$

where $2 \le \sigma \le \infty$ and $4 \le \rho \le \infty$ and S_σ is some constant. Therefore, the relation

$$\frac{1}{\nu} + \frac{2}{\nu p'} = \frac{1}{2}, \quad \text{i.e.} \quad \nu = 2 + \frac{4}{p'}, \tag{10.3}$$

has to be fulfilled. This holds for appropriate $p' > 2$ if $\nu \in [2, 4)$, which ends the proof of Proposition 10.3. ∎

Since u^* satisfies the Euler-Lagrange equation

$$u^* = \frac{1}{\omega} Q_m(u^*)$$

and F satisfies Assumption **(E1)**, we can use Proposition 10.3 to obtain

$$|T(r)u^*(x)| = \frac{1}{\omega}|T(r)Q_m(u^*)(x)| \le \frac{C_1(\lambda) + C_2(\lambda)}{c}$$

for all r and x, where we also used the uniform lower bound $\omega = \omega(\mu) \ge c > 0$, which we established in the beginning of the proof of Lemma 10.2. This also concludes the proof of Lemma 10.2. ∎

Finally, if μ is chosen sufficiently large, so that $|(T(r)u^*)(x)| < \mu$ for all x and $r \in [0, 1]$, which is possible by Lemma 10.2, then u^* is also a critical point of H. Further observing that $H_m(u) \le H(u)$ and $H_m(u^*) = H(u^*)$, u^* has to be a ground state of H.

10.3 Positive residual dispersion

We now consider the functional

$$H(u) = \frac{d_{\text{av}}}{2} \int_{-\infty}^{+\infty} |u_x|^2 \mathrm{d}x - \int_0^1 \int_{-\infty}^{+\infty} F(|T(t)u|) \mathrm{d}x \mathrm{d}t$$

for positive residual dispersion, where the nonlinearity potential F satisfies Assumptions **(E1)**–**(E4)**, but now with $2 < \gamma_1 \leq \gamma_2 < +\infty$. Modifying the functional in the same way as in the case $d_{\text{av}} = 0$, by the argument in [CHL17], there exists a minimiser $u^* \in H^1$ of the modified functional.

By the Sobolev imbedding theorem and unitarity of the free Schrödinger evolution on H^1, we have

$$\sup_{x,r} |(T(r)u)(x)| \leq \|T(r)u^*\|_{H^1} = \|u^*\|_{H^1} \leq C(\lambda).$$

So taking again the modification parameter μ large enough, we obtain that u^* is a critical point of the original functional. Note that for this argument to work, we do not need Lemma 10.2, which is why we can allow for arbitrary $2 < \gamma_1 \leq \gamma_2 < +\infty$ in Assumption **(E1)** on F if $d_{\text{av}} > 0$.

10.4 Diffraction management

In the case of diffraction management, the ground state solutions can again be found as minimisers of the averaged variational principle

$$P_\lambda = \inf \left\{ H(u) : \sum_{x\in\mathbb{Z}} |u(x)|^2 = \lambda \right\}, \tag{10.4}$$

where, with the forward difference $D_+u(x) = u(x+1) - u(x)$,

$$H(u) = \frac{d_{\text{av}}}{2} \sum_{x\in\mathbb{Z}} |D_+u(x)|^2 - \int_0^1 \sum_{x\in\mathbb{Z}} F(|T(t)u(x)|)\, \mathrm{d}t.$$

Again, for clarity of the presentation, we consider only the square wave diffraction profile corresponding to $\mu(\mathrm{d}r) = \mathbb{1}_{[0,1]}(r)\,\mathrm{d}r$. The discussion in [CHL16] shows that one can in fact include any diffraction profile one can think of.

Since $\|u\|_{\ell^2}^2 = \sum_{x\in\mathbb{Z}} |u(x)|^2$, we have the very simple estimate

$$\|u\|_{\ell^\infty} = \sup_{x\in\mathbb{Z}} |u(x)| \leq \|u\|_{l^2}\,.$$

Thus for any $u \in \ell^2(\mathbb{Z})$ with optical power $\|u\|_{\ell^2}^2 = \lambda$, we have the bound

$$\|T(t)u\|_{\ell^\infty} \leq \|T(t)u\|_{\ell^2} = \|u\|_{\ell^2} = \lambda^{1/2}$$

for all t. So taking $\mu > \lambda^{1/2}$ and modifying the functional according to the same rule as before, we get, according to [CHL16], a solution of the modified functional, which, since $\mu > \lambda^{1/2}$, is also a solution of the unmodified functional, as before. Note that this works for any residual diffraction $d_{\text{av}} \geq 0$ and any $2 < \gamma_1 \leq \gamma_2 < +\infty$ in Assumption **(E1)** on F.

CHAPTER 11

Existence of DM Solitons for Saturated Potentials II

After this short example, we now turn to the proof of Theorems 9.4 and 9.5.

11.1 Preparatory and technical remarks

In this section we review some important properties of the nonlinearity N. Most of these properties are adapted from [CHL17], which the reader may consult for a more complete background. The basic ingredient in most of these estimates is

Lemma 11.1. *Let $f \in L^2(\mathbb{R})$, $2 \leq q \leq 6$, and $\psi \in L^{\frac{4}{6-q}}$. Then*

$$\|T_r f\|_{L^q(\mathbb{R}^2, \mathrm{d}x\, \psi \mathrm{d}r)} \lesssim \|\psi\|_{\frac{4}{6-q}} \|f\|_2. \tag{11.1}$$

Proof. The inequality follows from interpolation between the unitary case $q = 2$ and the Strichartz inequality in one space dimension for $q = 6$, that is,

$$\iint_{\mathbb{R}^2} |T_r f(x)|^6 \,\mathrm{d}x\,\mathrm{d}r \leq 12^{-\frac{1}{2}} \|f\|_2^6.$$

For more details, see [CHL17, Lemma 2.1]. ∎

Without proof, we also cite

Lemma 11.2 (Lemma 4.7 in [CHL17]).

$\mathbf{d}_{\mathrm{av}} = 0$: *If $2 \leq \gamma_1 \leq \gamma_2 \leq 6$ and $\psi \in L^{\frac{4}{6-\gamma_2}}$, then the nonlinear and nonlocal functional $N : L^2(\mathbb{R}) \to \mathbb{R}$ given by*

$$L^2(\mathbb{R}) \ni f \mapsto N(f) = \iint_{\mathbb{R}^2} V(|T_r f|) \,\mathrm{d}x\, \psi \mathrm{d}r$$

is locally Lipshitz continuous on L^2 *in the sense that*

$$|N(f_1) - N(f_2)| \lesssim \left(1 + \|f_1\|_2^{\gamma_2 - 1} + \|f_2\|_2^{\gamma_2 - 1}\right) \|f_1 - f_2\|_2,$$

where the implicit constant depends only on the $L^{\frac{4}{6-\gamma_2}}$ *norm of* ψ.

$\mathbf{d}_{\mathrm{av}} > 0$: *If* $2 \le \gamma_1 \le \gamma_2 < \infty$ *and* $\psi \in L^1$, *then the nonlinear and nonlocal functional* $N : H^1(\mathbb{R}) \to \mathbb{R}$ *given by*

$$H^1(\mathbb{R}) \ni f \mapsto N(f) = \iint_{\mathbb{R}^2} V(|T_r f|)\,\mathrm{d}x\,\psi \mathrm{d}r$$

is locally Lipschitz continuous in the sense that

$$|N(f_1) - N(f_2)| \lesssim \left(1 + \|f_1\|_{H^1}^{\gamma_2 - 2} + \|f_2\|_{H^1}^{\gamma_2 - 2}\right) \left(\|f_1\|_2 + \|f_2\|_2\right) \|f_1 - f_2\|_2.$$

The directional derivatives of the nonlinearity are given by

Lemma 11.3. *If* $2 \le \gamma_1 \le \gamma_2 \le 6$ *and* $\psi \in L^1 \cap L^{\frac{4}{6-\gamma_2}}$ *(*$d_{\mathrm{av}} = 0$*), respectively if* $2 \le \gamma_1 \le \gamma_2 < \infty$ *and* $\psi \in L^1$ *(*$d_{\mathrm{av}} > 0$*), then for any* $f, h \in L^2(\mathbb{R})$, *respectively* $f, h \in H^1(\mathbb{R})$, *the functional* N *has directional derivatives given by*

$$D_h N(f) = \int_{\mathbb{R}} \mathrm{Re}\left\langle V'(|T_r f|)\tfrac{T_r f}{|T_r f|}, T_r h\right\rangle \psi \mathrm{d}r. \tag{11.2}$$

In particular, $h \mapsto D_h N(f)$ *is real linear and continuous.*

Proof. Let $f \in L^2(\mathbb{R})$ and $t \neq 0$. For any $h \in L^2(\mathbb{R})$ the difference quotient of N is

$$\begin{aligned}\frac{N(f + th) - N(f)}{t} &= \frac{1}{t}\left[\iint_{\mathbb{R}^2} V(|T_r(f + th)|) - V(|T_r f|)\,\mathrm{d}x\,\psi \mathrm{d}r\right] \\ &= \frac{1}{t}\iint_{\mathbb{R}^2}\int_0^1 \frac{\mathrm{d}}{\mathrm{d}s} V(|T_r(f + sth)|)\,\mathrm{d}s\,\mathrm{d}x\,\psi \mathrm{d}r. \end{aligned} \tag{11.3}$$

Since V is differentiable, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}s} V(|T_r(f + sth)|) = V'(|T_r(f + sth)|)\frac{t(T_r f\overline{T_r h} + T_r h\overline{T_r f} + 2st|T_r h|^2)}{2|T_r(f + sth)|}$$

and thus

$$(11.3) = \iint_{\mathbb{R}^2}\int_0^1 V'(|T_r(f + sth)|)\frac{T_r f\overline{T_r h} + T_r h\overline{T_r f} + 2st|T_r h|^2}{2|T_r(f + sth)|}\,\mathrm{d}s\,\mathrm{d}x\,\psi \mathrm{d}r.$$

Under the assumptions $2 \le \gamma_1 \le \gamma_2 \le 6$, respectively $2 \le \gamma_1 \le \gamma_2 < \infty$, on the nonlinearity, Lebesgue's dominated convergence theorem together with the continuity of V' implies that for $t \to 0$ we have

$$\begin{aligned} D_h N(f) &= \iint_{\mathbb{R}^2} \int_0^1 V'(|T_r f|) \frac{\mathrm{Re}(T_r f \overline{T_r h})}{|T_r f|} \, \mathrm{d}s \, \mathrm{d}x \, \psi \mathrm{d}r \\ &= \iint_{\mathbb{R}^2} V'(|T_r f|) \frac{\mathrm{Re}(T_r f \overline{T_r h})}{|T_r f|} \, \mathrm{d}x \, \psi \mathrm{d}r, \end{aligned}$$

which completes the proof of (11.2). Linearity of the map $h \mapsto D_h N(f)$ is immediate from (11.2), to see the continuity observe that by assumption (A1),

$$\begin{aligned} |D_h N(f)| &\le \iint_{\mathbb{R}^2} |V'(|T_r f(x)|)| \, |T_r h(x)| \, \mathrm{d}x \, \psi \, \mathrm{d}r \\ &\le \iint_{\mathbb{R}^2} \left[|T_r f(x)|^{\gamma_1 - 1} + |T_r f(x)|^{\gamma_2 - 1} \right] |T_r h(x)| \, \mathrm{d}x \, \psi \, \mathrm{d}r. \end{aligned}$$

For $2 \le \gamma \le 6$, Hölder's inequality (with exponents $\frac{\gamma}{\gamma - 1}$ and γ) implies the bound

$$\begin{aligned} \iint_{\mathbb{R}^2} |T_r f(x)|^{\gamma - 1} \, |T_r h(x)| \, \mathrm{d}x \, \psi \, \mathrm{d}r &\le \|T_r f\|^{\gamma - 1}_{L^\gamma(\mathrm{d}x \, \psi \mathrm{d}r)} \, \|T_r h\|_{L^\gamma(\mathrm{d}x \, \psi \mathrm{d}r)} \\ &\lesssim \|f\|_2^{\gamma - 1} \, \|h\|_2 \end{aligned}$$

by Lemma 11.1. By linearity, this already shows continuity of $h \mapsto D_h N(f)$ in the case $d_{\mathrm{av}} = 0$.

In the case of positive average dispersion, $d_{\mathrm{av}} > 0$, we can use $f \in H^1$ and Cauchy-Schwarz to bound

$$\begin{aligned} \iint_{\mathbb{R}^2} |T_r f(x)|^{\gamma - 1} \, |T_r h(x)| \, \mathrm{d}x \, \psi \, \mathrm{d}r &\le \sup_r \|T_r f\|_\infty^{\gamma - 2} \, \|T_r f\|_{L^2(\mathrm{d}x \, \psi \mathrm{d}r)} \, \|T_r h\|_{L^2(\mathrm{d}x \, \psi \mathrm{d}r)} \\ &\lesssim \|f\|_{H^1}^{\gamma - 2} \, \|f\|_2 \, \|h\|_2 \end{aligned}$$

for $2 \le \gamma < \infty$, since

$$\sup_{r \in \mathbb{R}} \|T_r f\|_\infty \le \sup_{r \in \mathbb{R}} \|T_r f\|_{H^1} = \|f\|_{H^1}$$

by the simple estimate $\|g\|_\infty \le \left(\|g\|_2 \|g'\|_2 \right)^{1/2} \le \|g\|_{H^1}$ and unitarity of the free Schrödinger evolution T_r on H^1. ∎

Remark. In the setting of $d_{\mathrm{av}} = 0$, that is, when working in $L^2(\mathbb{R})$, Lemma 11.3 also identifies the unique Riesz representative $\nabla N(f)$ (with respect to the real inner product $\mathrm{Re}\langle \cdot, \cdot \rangle$) of the continuous linear functional $h \mapsto D_h N(f)$ for fixed $f \in L^2(\mathbb{R})$,

$$\mathrm{Re}\langle \nabla N(f), h \rangle = D_h N(f) = \mathrm{Re} \left\langle \int_{\mathbb{R}} T_r^{-1} \left[V'(|T_r f|) \tfrac{T_r f}{|T_r f|} \right] \psi \, \mathrm{d}r, h \right\rangle,$$

so

$$\nabla N(f) = \int_{\mathbb{R}} T_r^{-1} \left[V'(|T_r f|) \tfrac{T_r f}{|T_r f|} \right] \psi \, \mathrm{d}r .$$

♦

Even though we do not need the following for our main results, we state and prove

Proposition 11.4. *Assume that (9.11) holds in addition to the assumptions of Lemma 11.3, with $2 \le \gamma \le 6$, respectively, $2 \le \gamma < \infty$. Then the functional $H : \mathcal{H}^{d_{\mathrm{av}}} \to \mathbb{R}$ is of class $\mathcal{C}^1(\mathcal{H}^{d_{\mathrm{av}}}, \mathbb{R})$.*

Proof. Since $f \mapsto \|f'\|_2^2$ is a $\mathcal{C}^1$ functional on H^1 and the directional derivatives of N are (real) linear, see Lemma 11.3, it suffices to show that $f \mapsto D_h N(f)$ is continuous for each $h \in \mathcal{H}^{d_{\mathrm{av}}}$. We start by estimating

$$\begin{aligned} &\left| D_h N(f+g) - D_h N(f) \right| \\ &\le \iint_{\mathbb{R}^2} |T_r h| \left| V'(|T_r f + T_r g|) \frac{T_r f + T_r g}{|T_r f + T_r g|} - V'(|T_r f|) \frac{T_r f}{|T_r f|} \right| \mathrm{d}x \, \psi \mathrm{d}r . \end{aligned}$$

Observe that by assumption (A1) and inequality (9.11), for any $z, w \in \mathbb{C}$,

$$\begin{aligned} &\left| V'(|z+w|) \frac{z+w}{|z+w|} - V'(|z|) \frac{z}{|z|} \right| \\ &\le |V'(|z+w|) - V'(|z|)| + \frac{|V'(|z+w|)|}{|z+w|} |w| + |V'(|z+w|)| \left| \frac{z}{|z+w|} - \frac{z}{|z|} \right| \\ &= |V'(|z+w|) - V'(|z|)| + \frac{|V'(|z+w|)|}{|z+w|} |w| + \frac{|V'(|z+w|)|}{|z+w|} \big| |z| - |z+w| \big| \\ &\lesssim |w| \left[(|z| + |w|)^{\gamma-2} + (|z| + |w|)^{\gamma_1 - 2} + (|z| + |w|)^{\gamma_2 - 2} \right] . \end{aligned}$$

It follows that $\left| D_h N(f+g) - D_h N(f) \right|$ can be bounded by a sum of terms of the form

$$\iint_{\mathbb{R}^2} |T_r h| |T_r g| \left(|T_r f| + |T_r g| \right)^{\gamma-2} \mathrm{d}x \, \psi \, \mathrm{d}r .$$

Using Hölder's inequality, with exponents $\gamma, \gamma, \frac{\gamma}{\gamma-2}$, and Lemma 11.1, we obtain the bound

$$\begin{aligned} &\left| D_h N(f+g) - D_h N(f) \right| \\ &\lesssim \|h\|_2 \|g\|_2 \left[\left(\|f\|_2 + \|g\|_2 \right)^{\gamma-2} + \left(\|f\|_2 + \|g\|_2 \right)^{\gamma_1-2} + \left(\|f\|_2 + \|g\|_2 \right)^{\gamma_2-2} \right], \end{aligned}$$

as in the proof of Lemma 11.3, which shows that all directional derivatives $D_h N$ are locally Lipshitz for each fixed $h \in \mathcal{H}^{d_{\mathrm{av}}}$. Therefore, $H \in \mathcal{C}^1(\mathcal{H}^{d_{\mathrm{av}}}; \mathbb{R})$ if $\gamma \in [2, 6]$ for $d_{\mathrm{av}} = 0$. Similarly, one proves the case $d_{\mathrm{av}} > 0$ with $\gamma \ge 2$. ■

11.2 Strict sub-additivity of the energy

The crucial ingredient in establishing existence of minimisers is restoring (pre-)compactness of minimising sequences modulo the natural symmetries of the problem. In this section we prove sub-additivity of the ground state energy with respect to $\lambda > 0$.

While strict sub-additivity was established under the Ambrosetti-Rabinowitz condition (9.8) in [CHL17], in general it fails in the saturation regime, where $\frac{V'(a)a}{V(a)} \to 2$. For any $C > 0$, we define the quantity

$$E_\lambda^{d_{\mathrm{av}}}(C) := \inf\left\{ H(f) : f \in \mathcal{S}_\lambda^{d_{\mathrm{av}}},\ \sup_{r \in \operatorname{supp}\psi} \|T_r f\|_\infty \le C \right\}. \tag{11.4}$$

The following proposition says that strict sub-additivity still holds in the case of saturated nonlinearities, at least if minimising sequences do not reach the saturation regime.

Proposition 11.5 (Strict sub-additivity). *Assume that (A1) and (A2) hold, and that for any $\lambda > 0$ there exists a constant $C > 0$ such that*

$$E_\lambda^{d_{\mathrm{av}}} = E_\lambda^{d_{\mathrm{av}}}(C). \tag{11.5}$$

Then for any $0 < \delta < \frac{\lambda}{2}$, and $\lambda_1, \lambda_2 \ge \delta$ with $\lambda_1 + \lambda_2 \le \lambda$, one has

$$E_{\lambda_1}^{d_{\mathrm{av}}} + E_{\lambda_2}^{d_{\mathrm{av}}} \ge \left[1 - \left(2^{\frac{\kappa^*(C)}{2}} - 2 \right) \left(\frac{\delta}{\lambda} \right)^{\frac{\kappa^*(C)}{2}} \right] E_\lambda^{d_{\mathrm{av}}}$$

whenever $E_\lambda^{d_{\mathrm{av}}} \le 0$, where $\kappa^*(C) := \inf_{0 < a \le C} \kappa(a) > 2$.

Remark. We will show in Propositions 11.11 and 11.14 that in fact for any $\lambda > 0$, the ground state energy is negative, $E_\lambda^{d_{\mathrm{av}}} \le 0$. Proposition 11.5 implies that

$$E_{\lambda_1}^{d_{\mathrm{av}}} + E_{\lambda_2}^{d_{\mathrm{av}}} > E_{\lambda_1+\lambda_2}^{d_{\mathrm{av}}}$$

whenever $E_{\lambda_1+\lambda_2}^{d_{\mathrm{av}}} < 0$, i.e. $E_\lambda^{d_{\mathrm{av}}}$ is *strictly sub-additive* if the ground state energy is *strictly negative.*

As shown in [CHL17], the strict sub-additivity of the ground state energy prevents minimising sequences from splitting. In particular, minimising sequences can be shown to be tight modulo the natural symmetries of the problem (shifts for $d_{\mathrm{av}} > 0$ or shifts and boosts for $d_{\mathrm{av}} = 0$). ♦

Proof of Proposition 11.5. Set

$$\chi(a) := \exp\left(-\int_{a_0}^{a} \frac{\kappa(b)}{b}\, \mathrm{d}b \right)$$

for some $0 < a_0 \le a$. Then $\chi(a_0) = 1$, $\chi'(a) = -\frac{\kappa(a)}{a}\chi(a)$, and therefore

$$\chi(a)V(a) - V(a_0) \ge 0 \tag{11.6}$$

since $(\chi V)' \ge 0$ by Assumption (A2). Setting $a_0 = sa$ for some $s \in (0, 1]$, we obtain

$$\begin{aligned} V(sa) &\le \exp\left(-\int_{sa}^{a} \frac{\kappa(b)}{b}\,\mathrm{d}b\right) V(a) = \exp\left(-\int_{s}^{1} \frac{\kappa(ab)}{b}\,\mathrm{d}b\right) V(a) \\ &\le \exp\left(-\inf_{\beta\in(0,1]} \kappa(a\beta) \int_s^1 \frac{\mathrm{d}b}{b}\right) V(a) = s^{\kappa^*(a)} V(a). \end{aligned}$$

Using that by Assumption (A2)

$$\inf_{0<b\le a} \kappa(b) \ge \inf_{0<b\le A} \kappa(b) = \kappa^*(A) > 2$$

for any finite $A \ge a > 0$, we get

$$V(sa) \le s^{\kappa^*(A)} V(a), \quad \text{for all} \quad s \in (0, 1],\ 0 < a \le A.$$

At this point the L^∞ bound comes into play, which guarantees that we always stay in a regime where saturation is not reached, that is, $\kappa^* > 2$!

Indeed, since $|T_r f(x)| \le \|T_r f\|_\infty \le C$ for all $r \in \operatorname{supp}\psi$, we get for $0 < \mu \le 1$ that

$$N(\mu^{1/2} f) = \iint_{\mathbb{R}^2} V(\mu^{1/2}|T_r f(x)|)\,\mathrm{d}x\,\psi(r)\,\mathrm{d}r \le \mu^{\kappa^*(C)/2} N(f),$$

and thus

$$\begin{aligned} E^{d_{\mathrm{av}}}_{\mu\lambda} &= \inf_{\|f\|_2^2=\mu\lambda} \left(\frac{d_{\mathrm{av}}}{2}\|f'\|_2^2 - N(f)\right) \ge \inf_{\|g\|_2^2=\lambda} \left(\mu\frac{d_{\mathrm{av}}}{2}\|g'\|_2^2 - \mu^{\kappa^*(C)/2} N(g)\right) \\ &\ge \mu^{\kappa^*(C)/2} E^{d_{\mathrm{av}}}_{\lambda}. \end{aligned}$$

As in [CHL17, Proposition 3.3], we can now take $\lambda_j = \mu_j\lambda$, $j = 1, 2$, with $\mu_1 + \mu_2 \le 1$, $\mu_1, \mu_2 \ge \frac{\delta}{\lambda}$. It then follows that

$$E^{d_{\mathrm{av}}}_{\lambda_1} + E^{d_{\mathrm{av}}}_{\lambda_2} = E^{d_{\mathrm{av}}}_{\mu_1\lambda} + E^{d_{\mathrm{av}}}_{\mu_2\lambda} \ge \left(\mu_1^{\kappa^*(C)/2} + \mu_2^{\kappa^*(C)/2}\right) E^{d_{\mathrm{av}}}_{\lambda},$$

and, since the function $t \mapsto (1+t)^{\kappa^*(C)/2} - 1 - t^{\kappa^*(C)/2}$ is increasing on $[1, \infty)$, we have

$$\mu_1^{\kappa^*(C)/2} + \mu_2^{\kappa^*(C)/2} \le 1 - \left(2^{\frac{\kappa^*(C)}{2}} - 2\right)\left(\frac{\delta}{\lambda}\right)^{\frac{\kappa^*(C)}{2}} < 1$$

for $\delta > 0$ and $\kappa^*(C) > 2$.

Now, if $E^{d_{\mathrm{av}}}_{\lambda} \le 0$, the sub-additivity

$$E^{d_{\mathrm{av}}}_{\lambda_1} + E^{d_{\mathrm{av}}}_{\lambda_2} \ge \left[1 - \left(2^{\frac{\kappa^*(C)}{2}} - 2\right)\left(\frac{\delta}{\lambda}\right)^{\frac{\kappa^*(C)}{2}}\right] E^{d_{\mathrm{av}}}_{\lambda},$$

follows. ∎

11.3 Thresholds

It turns out that under the Assumptions (A1)–(A3) on the nonlinear potential V, minimisers for $E_\lambda^{d_{\rm av}}$ may only exist for large enough λ. This is due to the fact that minimising sequences can be shown to be pre-compact modulo translations, respectively, translations and modulations, *if* the energy is *strictly negative*. The reason for this is sub-additivity: the ground state energy is *strictly* sub-additive only if $E_\lambda^{d_{\rm av}} < 0$!

This motivates

Definition 11.6 (Threshold)**.**

$$\lambda_{\rm cr}^{d_{\rm av}} := \inf\{\lambda > 0 : E_\lambda^{d_{\rm av}} < 0\}.$$

Assume that $E_\lambda^{d_{\rm av}} \le 0$ for all $\lambda > 0$ and $d_{\rm av} \ge 0$ (see Propositions 11.11 and 11.14 about the validity of this assumption). By the sub-additivity of the ground state energy, it immediately follows that

$$E_{\lambda_1}^{d_{\rm av}} \ge E_{\lambda_1}^{d_{\rm av}} + E_{\lambda_2}^{d_{\rm av}} \ge E_{\lambda_1+\lambda_2}^{d_{\rm av}},$$

where, by Proposition 11.5, the latter inequality is strict whenever $E_{\lambda_1+\lambda_2}^{d_{\rm av}} < 0$. In particular, the map $0 < \lambda \mapsto E_\lambda^{d_{\rm av}}$ is decreasing and it is strictly decreasing where $E_\lambda^{d_{\rm av}} < 0$.

Thus, $E_\lambda^{d_{\rm av}} = 0$ if $0 < \lambda < \lambda_{\rm cr}^{d_{\rm av}}$ and $E_\lambda^{d_{\rm av}} < 0$ if $\lambda > \lambda_{\rm cr}^{d_{\rm av}}$.

Lemma 11.7. *If V satisfies Assumptions (A2) and (A3), then $\lambda_{\rm cr}^{d_{\rm av}} < \infty$.*

Proof. By definition, $\lambda_{\rm cr}^{d_{\rm av}} < \infty$ if and only if $E_\lambda^{d_{\rm av}} < 0$ for some $\lambda > 0$. The claim therefore follows if we can find a suitable trial function with negative energy H, at least for large enough $\lambda > 0$.

Observe that by (A2) we again have the bound (11.6) on V. Let $a_* > 0$ be such that $V(a_*) > 0$, which exists by (A3). Then

$$V(a) \ge \exp\left(\int_{a_*}^{a} \frac{\kappa(b)}{b}\,\mathrm{d}b\right) V(a_*)\,\mathbb{1}_{[a_*,\infty)}(a),$$

where for $0 < a < a_*$ we just used the fact that $V(a) \ge 0$. Since by (A2) $\inf_{b>0}\kappa(b) \ge 2$, we get the lower bound

$$V(a) \ge \left(\frac{a}{a_*}\right)^2 V(a_*)\,\mathbb{1}_{[a_*,\infty)}(a). \tag{11.7}$$

Consider now centered Gaussian test functions

$$g_{\sigma_0}(x) = A_0\,\mathrm{e}^{-\frac{x^2}{\sigma_0}}, \quad \sigma_0 > 0, \tag{11.8}$$

where $A_0 = \left(\frac{2\lambda^2}{\pi\sigma_0}\right)^{1/4}$ is chosen in such a way that $\|g_{\sigma_0}\|_2^2 = \lambda$. Then $\|g'_{\sigma_0}\|_2^2 = \frac{\lambda}{\sigma_0}$ and the time evolution is given by

$$T_r g_{\sigma_0}(x) = A_0 \left(\frac{\sigma_0}{\sigma(r)}\right)^{1/2} \mathrm{e}^{-\frac{x^2}{\sigma(r)}}, \quad \sigma(r) = \sigma_0 + 4ir, \tag{11.9}$$

see [CHL16, Lemma B.3]. Thus,

$$|T_r g_{\sigma_0}(x)| = A_0 \left(\frac{\sigma_0^2}{\sigma_0^2 + (4r)^2}\right)^{1/4} \mathrm{e}^{-\frac{\sigma_0 x^2}{\sigma_0^2+(4r)^2}}.$$

We therefore have $|T_r g_{\sigma_0}(x)| \leq A_0$ for all $x \in \mathbb{R}$ and $r \in \mathbb{R}$. If $|x| \leq \sqrt{\sigma_0}$, we also have the lower bound

$$|T_r g_{\sigma_0}(x)| \geq A_0 \left(\frac{\sigma_0^2}{\sigma_0^2 + (4r)^2}\right)^{1/4} \mathrm{e}^{-\frac{\sigma_0^2}{\sigma_0^2+(4r)^2}}.$$

Hence choosing $R > 0$ such that $\operatorname{supp}\psi \subset [-R, R]$, we have

$$\frac{A_0}{2} \leq |T_r g_{\sigma_0}(x)| \leq A_0$$

for all $|x| \leq \sqrt{\sigma_0}$ and all $|r| \leq R$, assuming $\sigma_0 > 4R$.

Now set $\sigma_0 = \lambda$ for λ large enough. Then $\|g'_\lambda\|_2 = 1$ and $A_0 = \left(\frac{2\lambda}{\pi}\right)^{1/4}$. It follows with (11.7) that

$$\begin{aligned}
\int_{\mathbb{R}} V(|T_r g_\lambda(x)|)\,\mathrm{d}x &= \int_{|x|\leq\sqrt{\lambda}} V(|T_r g_\lambda(x)|)\,\mathrm{d}x + \int_{|x|>\sqrt{\lambda}} V(|T_r g_\lambda(x)|)\,\mathrm{d}x \\
&\geq \int_{|x|\leq\sqrt{\lambda}} \left(\frac{|T_r g_\lambda(x)|}{a_*}\right)^2 V(a_*)\,\mathbb{1}_{[a_*,\infty)}(|T_r g_\lambda(x)|)\,\mathrm{d}x \\
&\geq 2\sqrt{\lambda}\left(\frac{A_0}{2a_*}\right)^2 V(a_*)\,\mathbb{1}_{[a_*,\infty)}\left(\tfrac{A_0}{2}\right).
\end{aligned}$$

So for λ large enough, since $A_0 \sim \lambda^{1/4}$,

$$N(g_\lambda) = \iint_{\mathbb{R}^2} V(|T_r g_\lambda(x)|)\,\mathrm{d}x\,\psi\mathrm{d}r \gtrsim \lambda$$

and the energy is bounded by

$$H(g_\lambda) = \frac{d_{\mathrm{av}}}{2}\|g'_\lambda\|_2^2 - N(g_\lambda) \leq \frac{d_{\mathrm{av}}}{2} - C\lambda,$$

for some constant $C > 0$. Thus, choosing $\lambda > 0$ large enough, we can always achieve $H(g_\lambda) < 0$, so $E_\lambda^{d_{\mathrm{av}}} = \inf_{\|f\|_2^2=\lambda} H(f) \leq H(g_\lambda) < 0$. ■

Lemma 11.8. *If V satisfies Assumptions (A1), (A2), and (A4), then $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}} = 0$ for all $d_{\mathrm{av}} \geq 0$.*

Proof. Let $\lambda > 0$. We begin with $d_{\mathrm{av}} = 0$, that is, assume that there exists $\epsilon > 0$ such that $V(a) > 0$ for all $0 < a \leq \epsilon$. Let g_{σ_0} be the centered Gaussian (11.8) with $\|g_{\sigma_0}\|_2^2 = \lambda$. Then, by (11.9),

$$|T_r g_{\sigma_0}(x)| \leq A_0 = \left(\frac{2\lambda^2}{\pi\sigma_0}\right)^{1/4} \tag{11.10}$$

for all $x \in \mathbb{R}$ and $r \in \mathbb{R}$. Choosing σ_0 large enough, we can make $|T_r g_{\sigma_0}(x)| \leq \epsilon$, which implies $H(g_{\sigma_0}) = -N(g_{\sigma_0}) < 0$ by (A4), so $E_\lambda^{d_{\mathrm{av}}} < 0$. Since $\lambda > 0$ was arbitrary, it follows that $\lambda_{\mathrm{cr}}^0 = 0$.

For $d_{\mathrm{av}} > 0$ assume that there exist $\epsilon > 0$ and $2 < \gamma_0 < 6$ such that $V(a) \gtrsim a^{\gamma_0}$ for all $0 < a \leq \epsilon$. We consider the same centered Gaussian g_{σ_0} as above, with σ_0 so large that $|T_r g_{\sigma_0}(x)| \leq \epsilon$. It follows that

$$\begin{aligned} N(g_{\sigma_0}) &= \iint_{\mathbb{R}^2} V(|T_r g_{\sigma_0}(x)|)\,\mathrm{d}x\,\psi\,\mathrm{d}r \gtrsim \iint_{\mathbb{R}^2} |T_r g_{\sigma_0}(x)|^{\gamma_0}\,\mathrm{d}x\,\psi\mathrm{d}r \\ &= \left(\frac{\pi}{\gamma_0}\right)^{1/2}\left(\frac{2\lambda^2}{\pi}\right)^{\gamma_0/4} \sigma_0^{\frac{2-\gamma_0}{4}} \int_{\mathbb{R}} \frac{\psi(r)}{[1+(4r/\sigma_0)^2]^{\frac{\gamma_0-2}{4}}}\,\mathrm{d}r. \end{aligned}$$

Since $\|g_{\sigma_0}'\|_2^2 = \frac{\lambda}{\sigma_0}$, the energy of the Gaussian g_{σ_0} is bounded by

$$H(g_{\sigma_0}) \leq \frac{d_{\mathrm{av}}\lambda}{2\sigma_0}\left[1 - \frac{C}{d_{\mathrm{av}}\lambda}\left(\frac{\pi}{\gamma_0}\right)^{1/2}\left(\frac{2\lambda^2}{\pi}\right)^{\gamma_0/4}\sigma_0^{\frac{6-\gamma_0}{4}}\int_{\mathbb{R}}\frac{\psi(r)}{[1+(4r/\sigma_0)^2]^{\frac{\gamma_0-2}{4}}}\,\mathrm{d}r\right]$$

for some constant $C > 0$. In particular, since $2 < \gamma_0 < 6$ and

$$\int_{\mathbb{R}} \frac{\psi(r)}{[1+(4r/\sigma_0)^2]^{\frac{\gamma_0-2}{4}}}\,\mathrm{d}r \to \|\psi\|_1 > 0$$

as $\sigma_0 \to \infty$ by Lebesgue's dominated convergence theorem, we can make σ_0 sufficiently large such that $H(g_{\sigma_0}) < 0$. As $\lambda > 0$ was arbitrary, this yields $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}} = 0$. ∎

The following quantity will be useful in proving the non-existence of minimisers in the positive average dispersion case for sub-critical $0 < \lambda < \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$. Fix $C > 0$ and define

$$R_C(\lambda) := \sup\left\{\frac{N(\sqrt{\lambda}h)}{\lambda\|h'\|_2^2} : h \in H^1(\mathbb{R})\setminus\{0\}, \|h\|_2 = 1, \|h'\|_2 \leq C\right\}.$$

Lemma 11.9. *Let $C > 0$. If V satisfies Assumption (A2), then*

$$R_C(\lambda) \geq \left(\frac{\lambda}{\lambda_0}\right)^{\frac{1}{2}\kappa^*\left(\sqrt{\lambda C}\right)-1} R_C(\lambda_0)$$

for all $\lambda \geq \lambda_0 > 0$, with $\kappa^*\left(\sqrt{\lambda C}\right) = \inf_{a \leq \sqrt{\lambda C}} \kappa(a) > 2$.

This scaling property immediately implies

Corollary 11.10. *Let $C > 0$ and assume that V obeys Assumption (A2). If $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}} > 0$, then*

$$R_C(\lambda) < \frac{d_{\mathrm{av}}}{2} \quad \textit{for all} \quad 0 < \lambda < \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}.$$

Proof. Let $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}} > 0$, and assume that there exists $0 < \lambda_1 < \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$ such that $R_C(\lambda_1) \geq \frac{d_{\mathrm{av}}}{2}$. Pick $\lambda_2 \in (\lambda_1, \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}})$. Then Lemma 11.9 implies

$$R_C(\lambda_2) \geq \left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{2}\kappa^*\left(\sqrt{\lambda_2 C}\right)-1} R_C(\lambda_1) > R_C(\lambda_1) \geq \frac{d_{\mathrm{av}}}{2}$$

since $\kappa^*\left(\sqrt{\lambda C}\right) > 2$. In particular,

$$\begin{aligned}
E_{\lambda_2}^{d_{\mathrm{av}}} &= \inf_{\|g\|_2=1} \left(\frac{d_{\mathrm{av}}}{2}\lambda_2\|g'\|_2^2 - N(\sqrt{\lambda_2}g)\right) \leq \inf_{\substack{\|g\|_2=1\\ \|g'\|_2\leq C}} \left(\frac{d_{\mathrm{av}}}{2}\lambda_2\|g'\|_2^2 - N(\sqrt{\lambda_2}g)\right)\\
&= \inf_{\substack{\|g\|_2=1\\ \|g'\|_2\leq C}} \lambda_2\|g'\|_2^2\left(\frac{d_{\mathrm{av}}}{2} - \frac{N(\sqrt{\lambda_2}g)}{\lambda_2\|g'\|_2^2}\right) \leq \lambda_2 C^2 \inf_{\substack{\|g\|_2=1\\ \|g'\|_2\leq C}} \left(\frac{d_{\mathrm{av}}}{2} - \frac{N(\sqrt{\lambda_2}g)}{\lambda_2\|g'\|_2^2}\right)\\
&= \lambda_2 C^2\left(\frac{d_{\mathrm{av}}}{2} - \sup_{\substack{\|g\|_2=1\\ \|g'\|_2\leq C}} \frac{N(\sqrt{\lambda_2}g)}{\lambda_2\|g'\|_2^2}\right) = \lambda_2 C^2\left(\frac{d_{\mathrm{av}}}{2} - R_C(\lambda_2)\right) < 0,
\end{aligned}$$

in contradiction to $\lambda_2 < \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$ and the definition of $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$. ∎

Proof of Lemma 11.9. Let $h \in H^1 \setminus \{0\}$ with $\|h\|_2 = 1$ and $\|h'\|_2 \leq C$, and define the function

$$A(s) := s^{-2}N(sh)$$

for $s > 0$. Then

$$A'(s) = s^{-3}\left(sD_hN(sh) - 2N(sh)\right),$$

and by Assumption (A2) we have

$$\begin{aligned}
sD_hN(sh) - 2N(sh) &= \iint_{\mathbb{R}^2}\Big[V'(|T_r(sh)|)|T_r(sh)| - 2V(|T_r(sh)|)\Big]\,\mathrm{d}x\,\psi\,\mathrm{d}r\\
&\geq \iint_{\mathbb{R}^2}\Big[\kappa(|T_r(sh)|) - 2\Big]V(|T_r(sh)|)\,\mathrm{d}x\,\psi\,\mathrm{d}r.
\end{aligned}$$

Since for any $f \in H^1(\mathbb{R})$ the simple inequality $\|f\|_\infty^2 \leq \|f\|_2\|f'\|_2$ holds, we get

$$\|T_r(sh)\|_\infty = s\|T_rh\|_\infty \leq s\|T_rh\|_2^{1/2}\|T_rh'\|_2^{1/2} = s\|h\|_2^{1/2}\|h'\|_2^{1/2} \leq s\sqrt{C},$$

where we made use of the fact that T_r commutes with differentiation and is unitary on L^2, as well as the properties of h. It follows that

$$\begin{aligned} sD_h N(sh) - 2N(sh) &\geq \left(\inf_{a \leq s\sqrt{C}} \kappa(a) - 2 \right) N(sh) \\ &\geq \left(\inf_{a \leq t\sqrt{C}} \kappa(a) - 2 \right) N(sh) = \left(\kappa^*(t\sqrt{C}) - 2 \right) N(sh) \end{aligned}$$

for all $0 < s \leq t$, $t > 0$, and thus the function A satisfies the differential inequality

$$A'(s) \geq \left(\kappa^*(t\sqrt{C}) - 2 \right) s^{-1} A(s),$$

which yields

$$A(t) \geq \left(\frac{t}{t_0} \right)^{\kappa^*\left(t\sqrt{C}\right)-2} A(t_0)$$

for any $t \geq t_0 > 0$. In particular, we have

$$\begin{aligned} R_C(\lambda) = \sup_{\substack{\|h\|_2=1 \\ \|h'\|_2 \leq C}} \frac{N(\sqrt{\lambda}h)}{\lambda\|h'\|_2^2} &\geq \left(\frac{\lambda}{\lambda_0} \right)^{\frac{1}{2}\kappa^*\left(\sqrt{\lambda C}\right)-1} \sup_{\substack{\|h\|_2=1 \\ \|h'\|_2 \leq C}} \frac{N(\sqrt{\lambda_0}h)}{\lambda_0\|h'\|_2^2} \\ &= \left(\frac{\lambda}{\lambda_0} \right)^{\frac{1}{2}\kappa^*\left(\sqrt{\lambda C}\right)-1} R_C(\lambda_0) \end{aligned}$$

for all $\lambda \geq \lambda_0 > 0$. ■

11.4 Existence of minimisers for zero average dispersion

We start by establishing the existence of minimisers in the singular case $d_{\mathrm{av}} = 0$. Throughout this section, we assume that (A1), (A2), and (A3) hold with $3 \leq \gamma_1 \leq \gamma_2 < 5$, and that ψ is compactly supported with $\psi \in L^{\frac{4}{5-\gamma_2}+\delta}$ for some $\delta > 0$. This L^p condition on ψ ensures that the L^p condition in [CHL17] holds, in particular, all their multilinear estimates and splitting estimates continue to hold in our setting.

Proposition 11.11. *For any $\lambda > 0$, the energy functional $H = -N$ is bounded below on $\mathcal{S}_\lambda^0$ and*

$$-\infty < E_\lambda^0 \leq 0.$$

Proof. Let $\lambda > 0$. Integrating the bound on V' in (A1) yields

$$|V(a)| \lesssim a^{\gamma_1} + a^{\gamma_2} \tag{11.11}$$

and therefore

$$N(f) \lesssim \int_{\mathbb{R}} \|T_r f\|_{\gamma_1}^{\gamma_1} \psi \, \mathrm{d}r + \int_{\mathbb{R}} \|T_r f\|_{\gamma_2}^{\gamma_2} \psi \, \mathrm{d}r \lesssim \|f\|_2^{\gamma_1} + \|f\|_2^{\gamma_2}$$

by Lemma 11.1. It follows that

$$E_\lambda^0 = \inf_{\|f\|_2^2=\lambda} H(f) = - \sup_{\|f\|_2^2=\lambda} N(f) \gtrsim - \left(\lambda^{\frac{\gamma_1}{2}} + \lambda^{\frac{\gamma_2}{2}}\right) > -\infty.$$

Since $V(a) \geq 0$ for any $a > 0$, clearly $H(f) = -N(f) \leq 0$ for any $f \in \mathcal{S}_\lambda^0$ and therefore $E_\lambda^0 \leq 0$. ■

The following lemma is a generalization of a result by Kunze [Kun04, Lemma 2.12], and establishes L^∞ bounds on the time evolved gradient of H.

Lemma 11.12. *Let $f \in L^2(\mathbb{R})$, $3 \leq \gamma_1 \leq \gamma_2 < 5$, $\psi \in L^{\frac{4}{5-\gamma_2}+\delta}$ for some $\delta > 0$, and ψ compactly supported. Then $T_s \nabla H(f) \in L^\infty(\mathbb{R})$ and*

$$\sup_{s \in \mathbb{R}} \|T_s \nabla H(f)\|_\infty \lesssim \|f\|_2^{\gamma_1 - 1} + \|f\|_2^{\gamma_2 - 1}, \tag{11.12}$$

where the implicit constant depends on $\|\psi\|_{\frac{4}{5-\gamma_2}+\delta}$.

Proof. We have

$$\begin{aligned} \|T_s \nabla H(f)\|_\infty &= \sup_{\|g\|_1=1} \left|\mathrm{Re}\langle T_s \nabla H(f), g\rangle\right| = \sup_{\|g\|_1=1} \left|\mathrm{Re}\langle \nabla H(f), T_{-s} g\rangle\right| \\ &= \sup_{\|g\|_1=1} \left| \mathrm{Re} \int_{\mathbb{R}} \left\langle V'(|T_r f|) \frac{T_r f}{|T_r f|}, T_{r-s} g \right\rangle \psi(r) \, \mathrm{d}r \right|. \end{aligned}$$

Using the basic dispersive estimate for the free Schrödinger evolution, $\|T_s g\|_\infty \lesssim |s|^{-1/2} \|g\|_1$ for all $s \neq 0$, we obtain, together with assumption (A1),

$$\begin{aligned} \|T_s \nabla H(f)\|_\infty &\lesssim \int_{\mathbb{R}} \frac{\psi(r)}{|r-s|^{1/2}} \int_{\mathbb{R}} |V'(|T_r f|)| \, \mathrm{d}x \, \mathrm{d}r \\ &\lesssim \int_{\mathbb{R}} \frac{\psi(r)}{|r-s|^{1/2}} \left(\|T_r f\|_{\gamma_1 - 1}^{\gamma_1 - 1} + \|T_r f\|_{\gamma_2 - 1}^{\gamma_2 - 1} \right) \mathrm{d}r. \end{aligned} \tag{11.13}$$

An application of Hölder's inequality then yields

$$\int_{\mathbb{R}} \frac{\psi(r)}{|r-s|^{1/2}} \|T_r f\|_{\gamma-1}^{\gamma-1} \, \mathrm{d}r \leq \||\cdot - s|^{-1/2} \psi\|_{\frac{p}{p-1}} \left(\int_{\mathbb{R}} \|T_r f\|_{\gamma-1}^{p(\gamma-1)} \, \mathrm{d}r \right)^{1/p}.$$

The pair $(\gamma - 1, p(\gamma - 1))$ is Strichartz admissible if $\gamma - 1 \geq 2$ and $\frac{2}{p(\gamma-1)} = \frac{1}{2} - \frac{1}{\gamma-1}$, that is, $p = \frac{4}{\gamma-3}$. Note that $p \geq 1$ if $\gamma \leq 7$. In this case,

$$\int_{\mathbb{R}} \|T_r f\|_{\gamma-1}^{p(\gamma-1)} \, \mathrm{d}r \lesssim \|f\|_2^{p(\gamma-1)}$$

by Strichartz' inequality, and thus

$$\int_{\mathbb{R}} \frac{\psi(r)}{|r-s|^{1/2}} \|T_r f\|_{\gamma-1}^{\gamma-1} \,\mathrm{d}r \lesssim \| | \cdot - s|^{-1/2} \psi \|_{\frac{4}{7-\gamma}} \|f\|_2^{\gamma-1}.$$

Setting $\alpha = \frac{2}{7-\gamma}$, we see that we have to bound $\int |r-s|^{-\alpha} \psi^{2\alpha}(r)\,\mathrm{d}r$ uniformly in s. Let $\theta > 1$ and apply Hölder's inequality once more to see

$$\int |r-s|^{-\alpha}\psi^{2\alpha}(r)\,\mathrm{d}r \le \left(\int_{\operatorname{supp}\psi} |r-s|^{-\alpha\theta}\,\mathrm{d}r \right)^{\frac{1}{\theta}} \left(\int \psi(r)^{\frac{2\alpha\theta}{\theta-1}} \right)^{\frac{\theta-1}{\theta}}.$$

As long as $\alpha\theta < 1$, we have

$$\sup_{s\in\mathbb{R}} \int_{\operatorname{supp}\psi} |r-s|^{-\alpha\theta}\,\mathrm{d}r < \infty$$

since $\operatorname{supp}\psi$ is compact. So we need $\alpha < 1/\theta$, which is equivalent to

$$\frac{2\alpha\theta}{\theta-1} > \frac{2\alpha}{1-\alpha} = \frac{4}{5-\gamma}.$$

Since ψ is compactly supported and $\psi \in L^{\frac{4}{5-\gamma_2}+\delta}$ for some $\delta > 0$, we see that, setting $\alpha_j = \frac{4}{7-\gamma_j}$, there exist $\theta_j > 1$ with $\alpha_j\theta_j < 1$ and $\frac{2\alpha_j\theta_j}{\theta_j - 1} = \frac{4}{5-\gamma_j} + \delta$. This shows that both terms on the right hand side of (11.13) can be bounded uniformly in $s \in \mathbb{R}$. ∎

Lemma 11.13. *Assume that $3 \le \gamma_1 \le \gamma_2 < 5$ and that $\psi \in L^{\frac{4}{5-\gamma_2}+\delta}$ for some $\delta > 0$. Let $(u_n)_{n\in\mathbb{N}} \subset L^2(\mathbb{R})$, $\|u_n\|_2^2 = \lambda$ for all $n \in \mathbb{N}$, be a minimising sequence for E_λ^0. If $E_\lambda^0 < 0$, then there exists another minimising sequence $(v_n)_{n\in\mathbb{N}} \subset L^2 \cap L^\infty(\mathbb{R})$ with*

$$\sup_{r\in\mathbb{R}} \|T_r v_n\|_\infty \le C_\lambda.$$

Proof. **Step 1 (Construction of a modified minimising sequence).** Since H satisfies all the requirements of Ekeland's variational principle (see Appendix F), there exists another minimising sequence $(w_n)_{n\in\mathbb{N}} \subset \mathcal{S}_\lambda^0$ such that $H(w_n) \le H(u_n)$ for all $n \in \mathbb{N}$, $\|w_n - u_n\|_2 \to 0$ as $n \to \infty$, and

$$\nabla H(w_n) - \left\langle \nabla H(w_n), \frac{w_n}{\|w_n\|_2} \right\rangle \frac{w_n}{\|w_n\|_2} \to 0 \quad \text{as} \quad n \to \infty \tag{11.14}$$

strongly in L^2, where $\nabla H(f) = -\int_{\mathbb{R}} T_r^{-1} \left[V'(|T_r f|) \frac{T_r f}{|T_r f|} \right] \psi \mathrm{d}r$, see Remark 11.1. Write

$$g_n := \nabla H(w_n) + \sigma_n \frac{w_n}{\|w_n\|_2} = \nabla H(w_n) + \sigma_n \frac{w_n}{\sqrt{\lambda}}, \quad n \in \mathbb{N}, \tag{11.15}$$

with $\sigma_n := -\left\langle \nabla H(w_n), \frac{w_n}{\|w_n\|_2} \right\rangle$. Then $g_n \to 0$ strongly in L^2 for $n \to \infty$ by (11.14).

By assumption (A2),

$$-\langle \nabla H(w_n), w_n\rangle = D_{w_n} N(w_n) = \iint_{\mathbb{R}^2} V'(|T_r w_n|)\,|T_r w_n|\,\mathrm{d}x\,\psi(r)\,\mathrm{d}r$$
$$\geq 2N(w_n) = -2H(w_n) \xrightarrow{n\to\infty} -2E_\lambda^0 > 0.$$

So, picking a subsequence if necessary, we can assume that $\sigma_n \geq -\frac{E_\lambda^0}{\sqrt{\lambda}} > 0$ for all $n \in \mathbb{N}$. Therefore, σ_n^{-1} is uniformly bounded and $\sigma_n^{-1} g_n \to 0$ as $n \to \infty$.

Now define the sequence

$$v_n := -\sqrt{\lambda}\,\frac{\nabla H(w_n)}{\|\nabla H(w_n)\|_2}, \quad \|v_n\|_2^2 = \lambda, \quad n \in \mathbb{N}.$$

We will show that $(v_n)_{n\in\mathbb{N}} \subset \mathcal{S}_\lambda^0$ is again a minimising sequence for H. Indeed,

$$\|v_n - w_n\|_2 = \frac{\sqrt{\lambda}}{\sigma_n}\left\|\left(1 - \frac{\sigma_n}{\|\nabla H(w_n)\|_2}\right)\nabla H(w_n) - g_n\right\|_2$$
$$\leq \sqrt{\lambda}\left|1 - \frac{\sigma_n}{\|\nabla H(w_n)\|_2}\right|\frac{\|\nabla H(w_n)\|_2}{\sigma_n} + \sqrt{\lambda}\,\frac{\|g_n\|_2}{\sigma_n}.$$

Since $\sigma_n^{-1} g_n \to 0$ in L^2, it remains to show that

$$\frac{\sigma_n}{\|\nabla H(w_n)\|_2} \to 1$$

as $n \to \infty$. But from (11.15) we have

$$\|\nabla H(w_n)\|_2^2 = \|g_n\|_2^2 + \sigma_n^2 - 2\sigma_n \mathrm{Re}\left\langle \frac{w_n}{\|w_n\|_2}, g_n\right\rangle,$$

so that

$$\frac{\|\nabla H(w_n)\|_2^2}{\sigma_n^2} = 1 + \frac{\|g_n\|_2^2}{\sigma_n^2} - 2\,\mathrm{Re}\left\langle \frac{w_n}{\|w_n\|_2}, \frac{g_n}{\sigma_n}\right\rangle \to 1$$

as $n \to \infty$ since $\sigma_n^{-1} g_n \to 0$ in L^2.

Step 2 (L^∞ boundedness of the modified minimising sequence). By Lemma 11.12 and the bound

$$\|w_n\|_2 \|\nabla H(w_n)\|_2 \geq |\mathrm{Re}\langle \nabla H(w_n), w_n\rangle| \geq -2E_\lambda^0 > 0,$$

we obtain

$$\|T_s v_n\|_\infty = \frac{\sqrt{\lambda}}{\|\nabla H(w_n)\|_2}\|T_s \nabla H(w_n)\|_\infty$$
$$\lesssim \frac{\lambda}{|E_\lambda^0|}\left(\|w_n\|_2^{\gamma_1 - 1} + \|w_n\|_2^{\gamma_2 - 1}\right) = \left(\lambda^{\frac{\gamma_1+1}{2}} + \lambda^{\frac{\gamma_2+1}{2}}\right)/|E_\lambda^0|.$$

∎

We can now turn to the proof of existence of dispersion managed solitons in the case of zero average dispersion.

Proof of Theorem 9.4. We start with $0 < \lambda < \lambda^0_{\mathrm{cr}}$. Since by Proposition 11.11 $E^0_\lambda \leq 0$, the definition of the threshold (Definition 11.6) implies that $E^0_\lambda = 0$, proving part (i) of the theorem.

Assume now that $\lambda > \lambda^0_{\mathrm{cr}}$. Then, by definition, $E^0_\lambda < 0$.

Let $(v_n)_{n\in\mathbb{N}} \subset \mathcal{S}^0_\lambda \cap L^\infty(\mathbb{R})$ be the minimising sequence constructed in Lemma 11.13, such that $\|T_r v_n\|_\infty \leq C_\lambda$ for some uniform constant C_λ.

By Proposition 11.5, the ground state energy E^0_λ is strictly sub-additive along $(v_n)_{n\in\mathbb{N}}$. Once we have strict sub-additivity, the bound (4.7) from [CHL17, Proposition 4.4] again holds. Then one can use this, similarly to the proof of [CHL17, Proposition 4.6], to show that the sequence $(v_n)_{n\in\mathbb{N}}$ is tight (that is, $|v_n(x)|^2\,\mathrm{d}x$ and $|\widehat{v}_n(\eta)|^2\,\mathrm{d}\eta$ are tight in the sense of measures) modulo shifts and boosts, i.e. there exist shifts y_n and boosts ξ_n such that

$$\lim_{R\to\infty} \sup_{n\in\mathbb{N}} \int_{|x-y_n|>R} |v_n(x)|^2\,\mathrm{d}x = 0,$$
$$\lim_{L\to\infty} \sup_{n\in\mathbb{N}} \int_{|\eta-\xi_n|>L} |\widehat{v}_n(\eta)|^2\,\mathrm{d}\eta = 0.$$

Let $f_n(x) := \mathrm{e}^{\mathrm{i}\xi_n x} v_n(x-y_n)$, $n\in\mathbb{N}$, be the shifted and boosted minimising sequence. Then by the invariance of H under shifts and boosts, $(f_n)_{n\in\mathbb{N}}$ is again a minimising sequence with $\|f_n\|_2^2 = \|v_n\|_2^2 = \lambda$. Since $|f_n(x)| = |v_n(x-y_n)|$ and $|\widehat{f_n}(\eta)| = |\widehat{v}_n(\eta-\xi_n)|$, the sequence $(f_n)_{n\in\mathbb{N}}$ is also tight.

Since the sequence $(f_n)_{n\in\mathbb{N}}$ is bounded in $L^2(\mathbb{R})$, there exists a weakly convergent subsequence (again denoted by $(f_n)_{n\in\mathbb{N}}$) by the weak compactness of the unit ball. Since this subsequence is also tight, it converges even strongly in $L^2(\mathbb{R})$ to some $f \in L^2$. By continuity of the L^2 norm and the nonlinearity N under strong L^2-convergence, we have

$$E^0_\lambda \leq H(f) = -N(f) = \lim_{n\to\infty} -N(f_n) = E^0_\lambda$$

since $(f_n)_{n\in\mathbb{N}}$ is minimising. Thus f is a minimiser of the variational problem (9.5) for $d_{\mathrm{av}} = 0$.

The Euler-Lagrange equation of the constrained minimisation problem is the dispersion management equation (9.9) and it is a standard exercise to show that the minimiser h found above is a weak solution of the Euler-Lagrange equation,

$$\omega\,\langle f, g\rangle = -D_g N(f) = -\int_{\mathbb{R}} \left\langle V'(|T_r f|)\frac{T_r f}{|T_r f|}, T_r g\right\rangle \psi\,\mathrm{d}r \tag{11.16}$$

for all $g \in L^2(\mathbb{R})$, see also [CHL17] for more details. In particular, Lemma 11.12 implies that $T_s f \in L^\infty(\mathbb{R})$ for all $s \in \operatorname{supp}\psi$. Inserting $g = f$ as test function in (11.16) yields

$$\begin{aligned}\omega\|f\|_2^2 = \omega\lambda &= -\iint_{\mathbb{R}^2} V'(|T_r f(x)|)\,|T_r f(x)|\,\mathrm{d}x\,\psi\,\mathrm{d}r \\ &\le -\iint_{\mathbb{R}^2} \kappa(|T_r f(x)|)\,V(|T_r f(x)|)\,\mathrm{d}x\,\psi\,\mathrm{d}r \\ &\le -\kappa^*(C_\lambda)N(f) < -2N(f) = 2E_\lambda^0\end{aligned}$$

by assumption (A2) and the uniform bound on the minimiser. So $\omega < \frac{2E_\lambda^0}{\lambda}$. ■

11.5 Existence of minimisers for positive average dispersion

The situation is much easier in the positive average dispersion case since the uniform L^∞ bound is directly provided by the simple bound

$$\|h\|_\infty^2 \le \|h\|_2\|h'\|_2 \le \|h\|_{H^1}^2 \tag{11.17}$$

for any $h \in H^1(\mathbb{R})$, i.e., the Sobolev embedding $H^1(\mathbb{R}) \subset L^\infty(\mathbb{R})$. We will assume throughout this section that Assumptions (A1), (A2), and (A3) hold with $2 < \gamma_1 \le \gamma_2 < 10$. We further assume that ψ is compactly supported and $\psi \in L^{a_\delta}$ for some $\delta > 0$, where $a_\delta := \max\{1, \frac{4}{10-\gamma_2} + \delta\}$.

Proposition 11.14. *The energy functional H is bounded below on $\mathcal{S}_\lambda^{d_{\mathrm{av}}}$ for any $\lambda > 0$ and coercive in $\|f'\|$, that is,*

$$\lim_{\substack{\|f'\|\to\infty \\ \|f\|^2=\lambda}} H(f) = +\infty.$$

Moreover, $-\infty < E_\lambda^{d_{\mathrm{av}}} \le 0$.

Proof. For $2 < \gamma_1 \le \gamma_2 \le 6$ we can, as in Proposition 11.11, estimate the nonlinearity by

$$N(f) \lesssim \|f\|_2^{\gamma_1} + \|f\|_2^{\gamma_2}.$$

In case $\gamma_j > 6$ for some $j = 1, 2$, we can extract the excess part in the L^∞ norm, estimating

$$\int_{\mathbb{R}} \|T_r f\|_\gamma^\gamma\,\psi\,\mathrm{d}r \le \sup_{r\in\mathbb{R}} \|T_r f\|_\infty^\kappa \int_{\mathbb{R}} \|T_r f\|_{\gamma-\kappa}^{\gamma-\kappa}\,\psi\,\mathrm{d}r$$

for some $2 \le \gamma - \kappa \le 6$. Using (11.17),

$$\sup_{r\in\mathbb{R}} \|T_r f\|_\infty \le \left(\|f\|_2\|f'\|_2\right)^{1/2},$$

where we used the unitarity of T_r on L^2 and the fact that T_r commutes with ∂_x. Together with Lemma 11.1 this yields

$$N(f) \lesssim \|f'\|_2^{\frac{\kappa_1}{2}} \|f\|_2^{\gamma_1 - \frac{\kappa_1}{2}} + \|f'\|_2^{\frac{\kappa_2}{2}} \|f\|_2^{\gamma_2 - \frac{\kappa_2}{2}}$$

for suitable $(\gamma_j - 6)_+ \le \kappa_j \le \gamma_j - 2$, $j = 1, 2$, and an implicit constant that can be chosen in such a way that it only depends on the L^{a_δ} norm of ψ. It is easy to see that for given $a_\delta \ge 1$ one can always choose $\kappa_j < 4$. Therefore,

$$H(f) \ge \frac{d_{\text{av}}}{2} \|f'\|_2^2 - C \left(\|f'\|_2^{\frac{\kappa_1}{2}} \|f\|_2^{\gamma_1 - \frac{\kappa_1}{2}} + \|f'\|_2^{\frac{\kappa_2}{2}} \|f\|_2^{\gamma_2 - \frac{\kappa_2}{2}} \right) \tag{11.18}$$

for some constant $C = C(\|\psi\|_{a_\delta})$. In particular, if $\|f\|_2^2 = \lambda$, then $H(f) \to \infty$ as $\|f'\|_2 \to \infty$. Moreover,

$$E_\lambda^{d_{\text{av}}} \ge \inf_{t>0} \left(\frac{d_{\text{av}}}{2} t^2 - C \left(t^{\frac{\kappa_1}{2}} \lambda^{\frac{1}{2}(\gamma_1 - \frac{\kappa_1}{2})} + t^{\frac{\kappa_2}{2}} \lambda^{\frac{1}{2}(\gamma_2 - \frac{\kappa_2}{2})} \right) \right) > -\infty.$$

To prove that $E_\lambda^{d_{\text{av}}} \le 0$ we again calculate the energy of suitable centered Gaussians (11.8). Since by (11.11)

$$N(g_{\sigma_0}) \lesssim \|\psi\|_1 \sup_{r \in \operatorname{supp} \psi} \left(\|T_r g_{\sigma_0}\|_{\gamma_1}^{\gamma_1} + \|T_r g_{\sigma_0}\|_{\gamma_2}^{\gamma_2} \right),$$

where $2 < \gamma_1 \le \gamma_2$, it is not hard to see that

$$\lim_{\sigma_0 \to \infty} H(g_{\sigma_0}) = 0,$$

which implies $E_\lambda^0 \le 0$. ■

Proof of Theorem 9.5. Fix $0 < \lambda < \lambda_{\text{cr}}^{d_{\text{av}}}$. By definition of the threshold and $E_\lambda^{d_{\text{av}}} \le 0$, we must then have $E_\lambda^{d_{\text{av}}} = 0$. Assume now that there exists a minimiser $f \in \mathcal{S}_\lambda^{d_{\text{av}}}$ with $H(f) = E_\lambda^{d_{\text{av}}} = 0$. Then

$$\begin{aligned} 0 = H(f) &= \frac{d_{\text{av}}}{2} \|f'\|_2^2 - N(f) = \|f'\|_2^2 \left(\frac{d_{\text{av}}}{2} - \frac{N(f)}{\|f'\|_2^2} \right) \\ &\ge \|f'\|_2^2 \left(\frac{d_{\text{av}}}{2} - \sup_{\substack{\|g\|_2 = 1 \\ \|g'\|_2 \le \lambda^{-1/2} \|f'\|_2}} \frac{N(\sqrt{\lambda} g)}{\lambda \|g'\|_2^2} \right) \\ &= \|f'\|_2^2 \left(\frac{d_{\text{av}}}{2} - R_{\lambda^{-1/2}\|f'\|_2}(\lambda) \right). \end{aligned} \tag{11.19}$$

Since $\lambda < \lambda_{\text{cr}}^{d_{\text{av}}}$, Corollary 11.10 implies that $R_C(\lambda) < \frac{d_{\text{av}}}{2}$ for any $C > 0$, in particular for $C = \lambda^{-1/2} \|f'\|_2$. So

$$\frac{d_{\text{av}}}{2} - R_{\lambda^{-1/2}\|f'\|_2}(\lambda) > 0,$$

which by (11.19) implies that $\|f'\|_2 = 0$. But as the kernel of ∂_x is trivial on $H^1(\mathbb{R})$, we must have $f \equiv 0$, in contradiction to $\|f\|_2^2 = \lambda$, which shows that there cannot exist a minimiser if we are below the threshold $\lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$.

Assume now $\lambda > \lambda_{\mathrm{cr}}^{d_{\mathrm{av}}}$ and let $(v_n)_{n\in\mathbb{N}} \subset \mathcal{S}_\lambda^{d_{\mathrm{av}}}$ be a minimising sequence for $E_\lambda^{d_{\mathrm{av}}}$. Since H is coercive on $\mathcal{S}_\lambda^{d_{\mathrm{av}}}$, the sequence (v_n) is bounded. Indeed, since $\|v_n\|_2^2 = \lambda$ and $H(v_n) \to E_\lambda^{d_{\mathrm{av}}} > -\infty$, the bound (11.18) implies that $\|v_n'\|_2$ stays bounded, thus $\|v_n\|_{H^1}$ is bounded *uniformly* in $n \in \mathbb{N}$.

Together with (11.17) and the unitarity of T_r on H^1, we have

$$\|T_r v_n\|_\infty \le \|T_r v_n\|_{H^1} = \|v_n\|_{H^1} \le C_\lambda$$

for any $r \in \operatorname{supp}\psi$, and some constant $C_\lambda > 0$, and Proposition 11.5 implies that the ground state energy $E_\lambda^{d_{\mathrm{av}}}$ is strictly sub-additive. Hence arguing as in the proofs of [CHL17, Propositions 4.3 and 4.5], the minimising sequence is tight modulo shifts and tight in Fourier space, that is there exist shifts $(y_n)_{n\in\mathbb{N}}$ such that for the sequence $w_n := v_n(\cdot - y_n)$, $n \in \mathbb{N}$, we have

$$\lim_{R\to\infty} \sup_{n\in\mathbb{N}} \int_{|x|>R} |w_n(x)|^2 \,\mathrm{d}x = 0,$$

and there exists a constant $K < \infty$ such that for any $L > 0$

$$\sup_{n\in\mathbb{N}} \int_{|\eta|>L} |\widehat{w_n}(\eta)|^2 \,\mathrm{d}\eta = \sup_{n\in\mathbb{N}} \int_{|\eta|>L} |\widehat{v_n}(\eta)|^2 \,\mathrm{d}\eta \le \frac{K}{L^2}.$$

Since $H(w_n) = H(v_n)$ for all $n \in \mathbb{N}$ by translation invariance, $(w_n)_{n\in\mathbb{N}}$ is also a minimising sequence with $\|w_n\|_2^2 = \|v_n\|_2^2 = \lambda$, which is bounded in H^1, $\|w_n\|_{H^1} = \|v_n\|_{H^1} \le C_\lambda$. So the weak compactness of the unit ball implies that there exists a subsequence $w_{n_k} \rightharpoonup v \in H^1$ weakly in H^1 and in L^2. By tightness, we even have strong convergence in L^2. It follows that

$$\|v\|_2^2 = \lim_{k\to\infty} \|w_{n_k}\|_2^2 = \lambda > 0$$

and together with the weak sequential lower semi-continuity of the H^1 norm this also implies

$$\|v'\|_2^2 \le \liminf_{k\to\infty} \|w_{n_k}'\|_2^2.$$

Finally, since $\{w_{n_k}\}_{k\in\mathbb{N}}$ is bounded in H^1 and converges in L^2, the continuity of the nonlinearity N with respect to strong L^2-convergence (Lemma 11.2) yields

$$\lim_{k\to\infty} N(w_{n_k}) = N(v).$$

Altogether, we thus have shown that H is weakly lower semi-continuous along $\{w_{n_k}\}$, in particular

$$E_\lambda^{d_{\mathrm{av}}} \le H(v) \le \liminf_{k\to\infty} H(w_{n_k}) = E_\lambda^{d_{\mathrm{av}}},$$

since $\{w_{n_k}\}$ is minimising. It follows that f is a minimiser of the variational problem (9.5). The rest of the proof is analogous to the zero average dispersion case $d_{\mathrm{av}} = 0$. ■

APPENDIX F

Ekeland's variational principle

In this chapter we briefly derive the following corollary of Ekeland's variational principle [Eke74, see also the Appendix in [Cos07]] needed in the construction of our modified minimising sequence. Note that we do not require the functional to be $\mathscr{C}^1$, but only that all its directional derivatives exists and depend linearly and continuously on the direction.

Proposition F.1. *Let $\mathscr{H}$ be a real Hilbert space and $\varphi : \mathscr{H} \to \mathbb{R}$ a continuous functional with the property that all directional derivatives exist and the functional $h \mapsto D_h\varphi(f)$ is linear and continuous for all $f \in \mathscr{H}$.*

Assume that φ is bounded from below on $\mathcal{S}_\lambda = \{u \in \mathscr{H} : \|u\|^2 = \lambda\}$, and let $(f_n)_{n\in\mathbb{N}} \subset \mathcal{S}_\lambda$ be a minimising sequence for $\varphi|_{\mathcal{S}_\lambda}$. Then there exists another minimising sequence $(g_n)_{n\in\mathbb{N}} \subset \mathcal{S}_\lambda$ such that

$$\varphi(g_n) \leq \varphi(f_n), \quad \|g_n - f_n\| \to 0$$

and

$$|(D_{h_n}\varphi|_{\mathcal{S}_\lambda})(g_n)| \to 0 \quad \textit{as} \quad n \to \infty$$

for any $h_n \in T_{g_n}\mathcal{S}_\lambda$ with $\sup_n \|h_n\| < \infty$.

Remark. (i) As will be clear from the proof, linearity of the map $h \mapsto D_h\varphi(f)$ is not needed, the only important property is that the one-sided derivatives from left and right coincide, respectively, that $D_{-h}\varphi(f) = -D_h\varphi(f)$ for all $f \in \mathscr{H}$. Linearity allows us to represent, by reflexivity, the directional derivative at a given point f in $\mathcal{S}_\lambda$ by a vector $\nabla\varphi(f) \in \mathscr{H}$.

(ii) Let $u \in \mathcal{S}_\lambda$. Since by assumption, the map

$$h \mapsto D_h\varphi(u)$$

is linear and continuous, there exists by the Riesz representation theorem a uniquely determined vector $\nabla\varphi(u)$ such that

$$\langle \nabla\varphi(u), h\rangle = D_h\varphi(u).$$

Since $\mathcal{S}_\lambda$ is a sphere in $\mathcal{H}$, we have $\mathcal{H} = T_u\mathcal{S}_\lambda \oplus \mathbb{R}u$ for all $u \in \mathcal{S}_\lambda$. Therefore, the projection of $\nabla\varphi(u)$ onto $T_u\mathcal{S}_\lambda$ is given by

$$\nabla\varphi(u) - \left\langle \nabla\varphi(u), \frac{u}{\|u\|}\right\rangle \frac{u}{\|u\|}.$$

By Proposition F.1, we thus have

$$\left|\left\langle \nabla\varphi(g_n) - \langle\nabla\varphi(g_n), \tfrac{g_n}{\|g_n\|}\rangle \tfrac{g_n}{\|g_n\|}, h_n\right\rangle\right| = \left|(D_{h_n}\varphi|_{\mathcal{S}_\lambda})(g_n)\right| \to 0$$

as $n \to \infty$ for all $h_n \in T_{g_n}\mathcal{S}_\lambda$ with $\|h_n\| \leq 1$ (and therefore also for all $\widetilde{h}_n \in T_{g_n}\mathcal{S}_\lambda \oplus \mathbb{R}g_n = \mathcal{H}$ with $\|\widetilde{h}_n\| \leq 1$), so that

$$\nabla\varphi(g_n) - \langle\nabla\varphi(g_n), \tfrac{g_n}{\|g_n\|}\rangle \tfrac{g_n}{\|g_n\|} \to 0, \quad n \to \infty$$

strongly in $\mathcal{H}$.

♦

Proof. Let $c = \inf_{\mathcal{S}_\lambda}\varphi$ and set $\epsilon_n = \max\left\{\frac{1}{n}, \varphi(f_n) - c\right\}$. By Ekeland's variational principle there exists a sequence $(g_n)_{n\in\mathbb{N}} \subset \mathcal{S}_\lambda$ such that $\varphi(g_n) \leq \varphi(f_n)$ for all $n \in \mathbb{N}$, $\|g_n - f_n\| \to 0$ as $n \to \infty$, and

$$\varphi(g_n) < \varphi(u) + \sqrt{\epsilon_n}\,\|g_n - u\| \quad \text{for all} \quad u \neq g_n. \tag{F.1}$$

Now let $\gamma : (-1, 1) \to \mathcal{S}_\lambda$ be a $\mathcal{C}^1$ curve with $\gamma(0) = g_n$ and $\gamma'(0) = h_n$, for some arbitrary $h_n \in T_{g_n}\mathcal{S}_\lambda$. Then, by means of the continuity of $h \mapsto D_h\varphi(f)$ for all $f \in \mathcal{H}$, we have

$$\begin{aligned}
\lim_{t\to 0}\frac{\varphi(\gamma(t)) - \varphi(\gamma(0))}{t} &= \lim_{t\to 0}\frac{\varphi(\gamma(0) + t\gamma'(0) + o(t)) - \varphi(\gamma(0))}{t}\\
&= \lim_{t\to 0}\frac{\varphi(\gamma(0)) + tD_{\gamma'(0)+t^{-1}o(t)}\varphi(\gamma(0)) + o(t) - \varphi(\gamma(0))}{t}\\
&= \lim_{t\to 0} D_{\gamma'(0)+t^{-1}o(t)}\varphi(\gamma(0)) = D_{\gamma'(0)}\varphi(\gamma(0)) = D_{h_n}\varphi(g_n).
\end{aligned}$$

As the curve γ was arbitrary, this implies

$$(D_{h_n}\varphi|_{S_\lambda})(g_n) = \lim_{t\to 0}\frac{\varphi(\gamma(t)) - \varphi(\gamma(0))}{t}.$$

By (F.1), for all $t > 0$ we have

$$\varphi(\gamma(t)) - \varphi(\gamma(0)) > -\sqrt{\epsilon_n}\|\gamma(0) - \gamma(t)\|,$$

and dividing by $t > 0$ and letting $t \to 0$ yields

$$(D_{h_n}\varphi|_{S_\lambda})(g_n) = \lim_{t\downarrow 0} \frac{\varphi(\gamma(t)) - \varphi(\gamma(0))}{t} \geq -\sqrt{\epsilon_n}\|\gamma'(0)\| = -\sqrt{\epsilon_n}\|h_n\|.$$

Similarly, exchanging t by $-t$, one obtains

$$(D_{h_n}\varphi|_{S_\lambda})(g_n) = \lim_{t\downarrow 0} \frac{\varphi(\gamma(-t)) - \varphi(\gamma(0))}{-t} \leq \sqrt{\epsilon_n}\|\gamma'(0)\| = \sqrt{\epsilon_n}\|h_n\|,$$

and therefore

$$\left|(D_{h_n}\varphi|_{S_\lambda})(g_n)\right| \leq \sqrt{\epsilon_n}\|h_n\| \to 0 \quad \text{as} \quad n \to \infty.$$

∎

Bibliography

[AB98] M.J. Ablowitz, and G. Biondini, Multiscale pulse dynamics in communication systems with strong dispersion management. *Optics Letters* **23** (1998), 1668–1670. 169, 171

[dAMP13] L. de Almeida Maia, E. Montefusco, and B. Pellacci, Weakly coupled nonlinear Schrödinger systems: the saturation effect. *Calculus of Variations and Partial Differential Equations* **46** (2013), 325—351. MR 3016511. Zbl 1257.35171. 175

[AM01] M.J. Ablowitz, and Z.H. Musslimani, Discrete diffraction managed spatial solitons. *Physical Review Letters* **87** (2001), 254102 [4 pages]. 171

[AR73] A. Ambrosetti, and P. H. Rabinowitz, Dual variational methods in critical point theory and applications. *Journal of Functional Analysis* **14** (1973), 349–381. MR 0370183. Zbl 0273.49063. 176

[Bul10] A. Bulut, Maximizers for the Strichartz inequalities for the wave equation. *Differential and Integral Equations* **23** (2010), 1035–1072. MR 2742477. Zbl 1240.35314. 169

[Car09] E. Carneiro, A sharp inequality for the Strichartz norm. *International Mathematics Research Notices* **2009** (2009), 3127–3145. MR 2533799. Zbl 1178.35090. 169

[CHL16] M.-R. Choi, D. Hundertmark, and Y.-R. Lee, Discrete diffraction managed solitons: Threshold phenomena and rapid decay for general nonlinearities. *Preprint* (2016), 49 pages. arXiv 1602.05755. 185, 194

[CHL17] M.-R. Choi, D. Hundertmark, and Y.-R. Lee, Thresholds for existence of dispersion management solitons for general nonlinearities. *SIAM Journal on Mathematical Analysis* **49** (2017), 1519–1569. MR 3639574. 171, 172, 173, 174, 175, 176, 179, 185, 187, 191, 192, 197, 201, 202, 204

[Cos07] D. G. Costa, An Invitation to Variational Methods in Differential Equations. Birkhäuser Boston, 2007. MR 2321283. Zbl 1123.35001. 205

[ESMBA98] H. Eisenberg, Y. Silverberg, R. Morandotti, A. Boyd, and J. Aitchison, Discrete spatial optical solitons in waveguide arrays. *Physical Review Letters* **81** (1998), 3383–3386. 171

[ESMA00] H. Eisenberg, Y. Silverberg, R. Morandotti, and J. Aitchison, Diffraction management. *Physical Review Letters* **85** (2000), 1863–1866. 171

[Eke74] I. Ekeland, On the variational principle. *Journal of Mathematical Analysis and Applications* **47** (1974), 324–353. MR 0346619. Zbl 0286.49015. 205

[FVV12] L. Fanelli, L. Vega, and N. Visciglia, Existence of maximizers for Sobolev Strichartz inequalities. *Advances in Mathematics* **229** (2012), 1912–1923. MR 2871161. Zbl 1235.35012. 169

[Fos07] D. Foschi, Maximizers for the Strichartz inequality. *Journal of the European Mathematical Society* **9** (2007), 739–774. MR 2341830. Zbl 1231.35028. 169

[GT96] I.R. Gabitov, and S.K. Turitsyn, Averaged pulse dynamics in a cascaded transmission system with passive dispersion compensation. *Optics Letters* **21** (1996), 327–329. 169, 171

[GH91] S. Gatz, and J. Herrmann, Soliton propagation in materials with saturable nonlinearity. *Journal of the Optical Society of America B* **8** (1991), 2296–2302. 175

[GV85] J. Ginibre, and G. Velo, The global Cauchy problem for the nonlinear Schrödinger equation revisited. *Annales de l'Institut Henri Poincaré (C) Analyse Non Linéaire* **2** (1985), 309–327. MR 0801582. Zbl 0586.35042. 184

[HL12] D. Hundertmark, and Y.-R. Lee, On non-local variational problems with lack of compactness related to non-linear optics. *Journal of Nonlinear Science* **22** (2012), 1–38. MR 2878650. Zbl 1244.49008. 170, 171

[HZ06] D. Hundertmark, and V. Zharnitsky, On sharp Strichartz inequalities in low dimensions. *International Mathematics Research Notices* **2006** (2006), 34080 [18 pages]. MR 2219206. Zbl 1131.35308. 169

[JT02] L. Jeanjean, and K. Tanaka, A positive solution for an asymptotically linear elliptic problem on $\mathbb{R}^N$ autonomous at infinity. *ESAIM: Control, Optimisation and Calculus of Variations* **7** (2002), 597—614. MR 1925042. Zbl 1225.35088. 175

[Kun03] M. Kunze, On the existence of a maximizer for the Strichartz inequality. *Communications in Mathematical Physics* **243** (2003), 137–162. MR 2020224. Zbl 1060.35133. 169, 183

[Kun04] M. Kunze, On a variational problem with lack of compactness related to the Strichartz inequality. *Calculus of Variations and Partial Differential Equations* **19** (2004), 307–336. MR 2033144. 198

[Lus04] P.M. Lushnikov, Oscillating tails of a dispersion-managed soliton. *Journal of the Optical Society of America B* **21** (2004), 1913–1918. 169

[Man16] R. Mandel, Minimal energy solutions and infinitely many bifurcating branches for a class of saturated nonlinear Schrödinger systems. *Advanced Nonlinear Studies* **16** (2016), 95—113. MR 3456750. Zbl 1334.35319. 175

[TBF12] S.K. Turitsyn, B.G. Bale, and M.P. Fedoruk, Dispersion-managed solitons in fibre systems and lasers. *Physics Reports* **521** (2012), 135–203. 169

[UOT98] A. Usman, J. Osman, and D. R. Tilley, Analytical solitary wave solutions for nonlinear Schrödinger equations with denominating saturating nonlinearity. *Journal of Physics A: Mathematical and General* **31** (1998), 8397–8403. Zbl 0986.81025. 175

[ZGJT01] V. Zharnitsky, E. Grenier, C. K. R. T. Jones, and S. K. Turitsyn, Stabilizing effects of dispersion management. *Physica D: Nonlinear Phenomena* **152–153** (2001), 794–817. MR 1837940. Zbl 0972.78021. 170