Abstract

Title: Capturing the Large Scale Behavior of Many Particle Systems Through Coarse-Graining

Samuel Punshon-Smith
Doctor of Philosophy, 2017

Directed by: Professor Charles David Levermore
Department of Mathematics

This dissertation is concerned with two areas of investigation: the first is understanding the mathematical structures behind the emergence of macroscopic laws and the effects of small scales fluctuations, the second involves the rigorous mathematical study of such laws and related questions of well-posedness. To address these areas of investigation the dissertation involves two parts:

Part I concerns the theory of coarse-graining of many particle systems. We first investigate the mathematical structure behind the Mori-Zwanzig (projection operator) formalism by introducing two perturbative approaches to coarse-graining of systems that have an explicit scale separation. One concerns systems with little dissipation, while the other concerns systems with strong dissipation. In both settings we obtain an asymptotic series of 'corrections' to the limiting description which are small with respect to the scaling parameter, these corrections represent the effects of small scales. We determine that only certain approximations give rise to dissipative effects in the resulting evolution. Next we apply this framework to the problem of coarse-graining the locally conserved quantities of a classical Hamilto-

nian system. By lumping conserved quantities into a collection of mesoscopic cells, we obtain, through a series of approximations, a stochastic particle system that resembles a discretization of the non-linear equations of fluctuating hydrodynamics. We study this system in the case that the transport coefficients are constant and prove well-posedness of the stochastic dynamics.

Part II concerns the mathematical description of models where the underlying characteristics are stochastic. Such equations can model, for instance, the dynamics of a passive scalar in a random (turbulent) velocity field or the statistical behavior of a collection of particles subject to random environmental forces. First, we study general well-posedness properties of stochastic transport equation with rough diffusion coefficients. Our main result is strong existence and uniqueness under certain regularity conditions on the coefficients, and uses the theory of renormalized solutions of transport equations adapted to the stochastic setting. Next, in a work undertaken with collaborator Scott-Smith we study the Boltzmann equation with a stochastic forcing. The noise describing the forcing is white in time and colored in space and describes the effects of random environmental forces on a rarefied gas undergoing instantaneous, binary collisions. Under a cut-off assumption on the collision kernel and a coloring hypothesis for the noise coefficients, we prove the global existence of renormalized (DiPerna/Lions) martingale solutions to the Boltzmann equation for large initial data with finite mass, energy, and entropy. Our analysis includes a detailed study of weak martingale solutions to a class of linear stochastic kinetic equations. Tightness of the appropriate quantities is proved by an extension of the Skorohod theorem to non-metric spaces.

Capturing the Large Scale Behavior of Many Particle Systems Through Coarse-Graining

by

Samuel Punshon-Smith

Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park in partial fulfillment of the requirements for the degree of Doctor of Philosophy
2017

Advisory Committee:
Professor Charles David Levermore
Professor Jacob Bedrossian
Professor Pierre-Emmanuel Jabin
Professor Christopher Jarzynski
Professor Konstantina Trivisa

Dedication

To my father, for introducing me to the pleasures of intellectual pursuit.

Acknowledgments

First and foremost, I would like to thank my advisor, Dave Levermore, for suggesting the problem in part I of this dissertation. If not for his constant encouragement and insistence on my stepping outside my comfort zone, this work would have just a tiny fraction of what it became. Dave's willingness to allow me to make mistakes and discover them on my own have allowed me to grow enormously as a researcher, and has taught me many valuable lessons in conducting independent research. I can only hope to obtain the breadth of knowledge and level of insight that Dr. Levermore has achieved during his mathematical career.

I would also like to thank my dissertation committee members for devoting the time and energy to reading this lengthy dissertation and providing valuable feedback: Professor Jacob Bedrossian has served as a source of inspiration and confidence during our overlap at University of Maryland. He gave me valuable and detailed feedback on the initial draft of this dissertation including many of the typos that I had missed. For this, I am extremely grateful. Pierre-Emmanuel Jabin has been a tremendous support and, in many ways, has been an secondary mentor during my Ph.D. It is through his encouragement and many discussions that the work in Part II of this dissertation came to fruition. His careful reading of the manuscript for the work on Stochastic Boltzmann has proven immensely useful. Christopher Jarzynski gave detailed feedback on the initial manuscript of this dissertation finding an enormous number of typos that I daresay might have never been found by my eye. His positive feedback and genuine interest in the problems in this dissertation has been very reassuring. Konstantina Trivisa has provided me with unwavering support and enthusiasm during my Ph.D. If not for her role as director of AMSC I may have never been a part of the Mathematics department of University of Maryland. For this I am extremely grateful.

In addition to my committee, I would like to acknowledge the staff of the Mathematics and AMSC departments, particularly Alverda Mcoy and her successor Jessica Sadler for keeping me on track to graduate and putting up with my propensity to miss important deadlines.

I would also like to thank my collaborator, peer, and close friend, Scott Smith, for many long discussions on a variety of mathematical topics. I am extremely grateful for his patience, and interest in the many stupid questions I have posed to him over the year. His careful methodical nature has provided me with a model for how to conduct systematic incremental research.

Finally, and perhaps most importantly I would like to thanks my many friends and family for their unwavering support during the most difficult times of my Ph.D. Particularly my Parents David and Lesley, my Brothers Ben and Daniel, and my lovely girlfriend Rhiannon. I am not sure that I would be where I am today without them.

Table of Contents

List of Figures

Introduction to the Dissertation

The physical world is multi-scale. Natural laws tend to exhibit drastically different structures at various time and space scales. Quite remarkably, it is often possible to describe the behavior at each of these scales independently of the other scales and with significantly fewer degrees of freedom than are present at the smaller scales. Such *effective equations* can emerge in unusual ways and are often not immediately accessible from the underlying microscopic laws. The equations of fluid mechanics, like the Euler equations or the Navier-Stokes equations are examples of effective equations governing hydrodynamic fields associated to a system of many classical particles. Other examples of effective equation include equations in kinetic theory, like the Boltzmann equation or Vlasov equation, which govern the evolution of a kinetic density of particles over a one-particle phase space.

In this dissertation, we will mostly follow two main lines of inquiry. The first involves the process of representing a system with many degrees of freedom by one with fewer degrees of freedom, known as coarse-graining. Here we are interested in questions like: Can one always derive a given effective description directly from the microscopic system? How does one pass from one set of effective equations to another? Is there a general procedure for determining a set of effective equations at any scale of interest? How does one take into account the influence of smaller scale fluctuations in an effective model? The second line of inquiry involves the study of the qualitative and quantitative behavior of the equations arising from such effective descriptions. Here, several natural questions come to mind: Are the equations of a given effective description well-posed? What is the long time behavior of the solutions? How well do a set of effective equations hold outside of their given scale? How does one incorporate the effects of 'small' scales outside an effective equations prescribes scale?

The dissertation is broken up into two parts with distinct conceptual contributions, the first is largely *formal* and attempts to address questions along first line of inquiry by exploring the mathematical structure in a setting where very few rigorous results are available, the other is entirely *rigorous* and addresses questions along the second line of inquiry, studying well-posedness of certain stochastic perturbations of macroscopic equations using well-developed mathematical tools from the theory of stochastic partial differential equations.

More specifically, Part I concerns the theory of coarse-graining. In the first half, we study the problem abstractly through the Mori-Zwanzig (projection operator) framework, viewing the procedure of coarse-graining as the application of a certain projection operator $\mathcal{P}$ on the solution f of a linear evolution equation

$$\frac{\mathrm{d}}{\mathrm{d}t} f = \mathcal{L} f,$$

where $\mathcal{L}$ is a certain linear operator generating the microscopic evolution. This framework, though formal, has broad applications to a wide variety of problems in classical and quantum statistical mechanics. Our contribution is to develop two perturbative approaches for obtaining dissipative corrections to the, (leading order) Galerkin truncated system

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}f = \mathcal{P}\mathcal{L}\mathcal{P}f.$$

The first approach is useful for when the system has no dissipation and relies on a specific decomposition of the fluctuations into a fast and a slow part. The second approach is more relevant when the starting system has some dissipation and is scaled so that the dissipation dominates the evolution of the small scale fluctuations. In this setting, we obtain a sequence of approximations to the Galerkin truncated system and show that only every 4th term in the sequence leads to an approximation that is dissipative.

The second half of Part I involves the more concrete problem of coarse-graining a one dimensional classical particle system with nearest neighbor interactions and Hamiltonian

$$H = \sum_i \tfrac{1}{2}v_i^2 + V(x_i - x_{i-1}),$$

where (x_i, v_i) are the position and velocity of the $i$the particle and $V(r)$ is a singular repulsive interaction potential. The coarse-graining procedure involves dividing the particles into mesoscopic cells and averaging the inter-particle spacing $r_i = x_i - x_{i-1}$, momentum v_i, and energy $e_i = \frac{1}{2}v_i^2 + V(r_i)$ of the particles inside each cell. Using the perturbative Mori-Zwanzig approach developed in the first half, we show that the leading order evolution of the coarse-grained cells is given by so-called 'discrete Euler dynamics'. Iterating this procedure we find that particle systems with gamma-law potential $V(r) = Cr^{1-\gamma}$ are invariant under the coarse-graining procedure; we refer to this, for reasons that will become clear later, as the 'ideal gas fixed point'. The main novelty of this work, however, is the derivation of a dissipative stochastic correction to the discrete Euler dynamics which take into account small scale fluctuations. This dissipative fluid-particle model can be viewed as a discretization of the equations of non-linear fluctuating hydrodynamics; they conserve volume, momentum and energy, with dissipative terms modeling the effects of viscosity, thermal conductivity and thermal fluctuations in the fluid. We give conditions under which this system is well-posed, meaning the energy or volume of a coarse-grained particle cannot (with probability one) collapse to zero in finite time. We reduce the derivation of the dissipative fluid-particle model to two key approximations; the first is a relaxation approximation and is strongly related to ergodicity of the underlying system; the second is a Markovian approximation which removes certain memory effects under the assumption of sufficient decay of various auto-correlation functions.

Part II, concerns the study of kinetic equations with stochastic external forcing and the theory of renormalized solutions to transport equations. We study two related problems. The first deals with existence and uniqueness of stochastic continuity equations of the form

$$\partial_t f + \operatorname{div}(uf) - \operatorname{div}\operatorname{div}(af) + \sum_k \operatorname{div}(\sigma_k f)\dot{W}_k = 0,$$

where u is the drift $\{\sigma_k\}$ are the noise coefficients $a = \frac{1}{2}\sum_k \sigma_k \otimes \sigma_k$, and $\{W_k\}$ are a collection of independent Brownian motions. Here the main contribution is to prove the existence of renormalized (hence unique) solutions in L^p, $p > 2$ for such equations with general initial data and rough (Sobolev regular) noise coefficients $\{\sigma_k\}$. This approach is rather general is consistent with analogous results for existence and uniqueness of SDE's with rough noise coefficients [122] as well as Kolmogorov equations [84] (see also [25]).

The second problem, a joint work with collaborator Scott Smith, concerns the Boltzmann equation with Stratonovich stochastic forcing

$$\partial_t f + v \cdot \nabla_x f + \sum_k \sigma_k \cdot \nabla_v f \circ \dot{W}_k = \mathcal{B}(f, f),$$

where $\operatorname{div}_v \sigma_k = 0$. Such an equation is a kinetic theory analogue to the stochastically forced equations of fluid mechanics, which have received significant attention in recent years. Our main result is to prove global existence of renormalized martingale solutions for a general class of initial data and noise coefficients and, and obtain certain local and global averaged balance laws and global entropy dissipation. To out knowledge this is the first rigorous result regarding the stochastic perturbations of the non-linear Boltzmann equation. The result is obtained using compactness and martingale tools from the theory of stochastic partial differential equations.

Part I

Coarse-Graining

If I have had any success in mathematical physics, it is, I think, because I have been able to dodge mathematical difficulties.

Josiah Willard Gibbs

Introduction to Part I

The large-scale behavior of many-body systems is of central interest in many disciplines. Such systems typically have simple rules governing their constituents at small scales (*microscopic laws*), but exhibit complicated patterns and rules on larger scales (*macroscopic laws*). These systems often contain several distinct scales which exhibit drastically different behaviors, so-called *multi-scale phenomena*. Naturally, there is broad interest among disciplines in obtaining models that govern the effective evolution of a system at a given scale. There are a vast number of models available to describes the effective behavior of many body systems at a variety of scales. The Euler equations of fluid mechanics are a classic example of such a model, along with the myriad of other macroscopic model in continuum mechanics and kinetic theory. Sometimes when the separation between scales is not strong enough, small scale structures can couple to the behavior of the large scales and have a non-negligible effect. Most notably, when studying a fluid at the mesoscale (between micro and macro), small scale fluctuations about equilibrium become important and their non-trivial correlations are responsible for the emergence of transport phenomena like viscosity and thermal conductivity, which play an important role even at the macroscopic scale. In general, this weak coupling between scales is not fully understood and there appears to be no agreed upon way to include its effect in a macroscopic model. The work in this part of the dissertation is an attempt to understand the influence of fluctuations on the behavior of a fluid at meso-scopic scales.

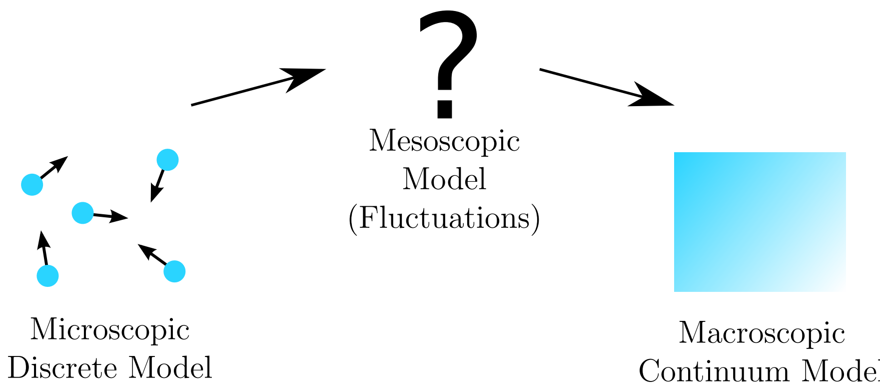

Figure 2.1: Multiscale models and the role of fluctuations

Part I one of this dissertation addresses the mathematical structure behind the theory of coarse-graining, namely the procedure of representing a system with many degrees of freedom by one with significantly fewer degrees of freedom. One of the more standard frameworks for coarse-graining is the *Mori-Zwanzig formalism*, named after its pioneers R. Zwanzig [123, 124] and H. Mori [95]. It has proven to be a tremendously powerful tool for obtaining the form for coarse-grained models at a variety of scales, although it suffers from a lack of a rigorous foundation. At its core, the Mori-Zwanzig formalism requires two main ingredients: a linear evolution equation

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = \mathcal{L} f(t)$$

and a projection operator $\mathcal{P}$. The projection operator $\mathcal{P}$ acts on f (where f takes values, perhaps, in some Banach space) and represents the action of coarse-graining, selecting certain 'relevant' variables of interest, and averaging out the 'irrelevant' degrees of freedom. A typical example of projection is the average with respect to some equilibrium measure conditioned on the value of a relevant variable. Another example is the s particle marginal of an N particle distribution with the distributions of the other particles replaced by an equilibrium measure. In general, the so-called 'projected dynamics' $\hat{f}(t) = \mathcal{P}f(t)$ will have non-Markovian memory effects on it's evolution, meaning that the future evolution of $f(t + \mathrm{d}t)$ may depend on the the entire history $\{f(s)\}_{s<t}$ as opposed to just the value at time t. However, when there is some time-scale separation, namely if the projected evolve on a time scale much slower than the persistence of the memory, then memory effects are assumed to be delta correlated in time and may be neglected; This is the so-called 'Markov approximation'. The mathematical justifications for such an approximation are in general not clear, and the precise definition of time-scale separation can be hard to define. Nevertheless, we will be interested in the mathematical structure behind various Markov approximations.

Specifically, in Chapter 3, we explore Markov approximations in the Mori-Zwanzig theory in more detail. Here we propose two perturbative approaches for obtaining dissipative corrections to the Galerkin truncated system

$$\frac{\mathrm{d}}{\mathrm{d}t}f(t) = \mathcal{P}\mathcal{L}\mathcal{P}f(t) + \text{ ``dissipative corrections''}.$$

Both approaches are formal and are meant to serve as a tool to guide in the construction of coarse-grained models.

The first approach is more applicable to systems with little or no dissipation and will be the main approach used in Chapter 4. For simplicity we suppose there exists an explicit time scale separation through the decomposition

$$\mathcal{L} = \frac{1}{\epsilon}\mathcal{L}_0 + \mathcal{L}_1, \tag{2.1}$$

with $\mathcal{P}\mathcal{L}_0 = \mathcal{L}_0\mathcal{P} = 0$. The parameter ϵ controls the scale separation between the relevant and irrelevant variables. In the limit as $\epsilon \to 0$ one can make the Markov assumption more precise and obtain an asymptotic series of corrections to the Galerkin truncation

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = \mathcal{P}\mathcal{L}\mathcal{P} f(t) + \sum_{n\geq 0} \epsilon^n \bar{\Phi}^n \mathcal{P} f(t),$$

where $\{\bar{\Phi}^n\}$ are operators encoding higher order time correlations. A similar approach has been taken in [71, 92] in a different setting. The value of this approach over the more pedestrian approach usually considered in the Mori-Zwanzig theory is that the operators $\{\bar{\Phi}^n\}$ can be computed explicitly in terms of the dynamics of known objects. Moreover, we show that the first in this asymptotic series is dissipative.

The second approach explores the process of coarse-graining systems that already have some dissipation and assumes that the dissipation dominates at the small scales. Specifically, we assume that the operator $\widetilde{\mathcal{L}} = (I - \mathcal{P})\mathcal{L}(I - \mathcal{P})$, has the explicit decomposition,

$$\widetilde{\mathcal{L}} = \widetilde{\mathcal{A}} + \frac{1}{\epsilon}\widetilde{\mathcal{S}},$$

where $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{S}}$ denote the symmetric and skew symmetric parts of $\widetilde{\mathcal{L}}$ respectively. Similar to the decomposition (2.1) one can make the Markov assumption more precise and obtain an asymptotic series of corrections to the Galerkin truncation

$$\frac{\mathrm{d}}{\mathrm{d}t}f(t) = \mathcal{PLP}f(t) + \sum_{n\geq 0} \epsilon^n \bar{\Psi}^n \mathcal{P} f(t)$$

where $\{\bar{\Psi}^n\}$ are another a collection of operators encoding information about higher order correlations. What's interesting in this setting is that not only is the first term in the series dissipative, but every $4m+1$ term is also. This is analogous to the Chapman-Enskog expansion in kinetic theory where certain terms in the truncation can be shown to lead to fluid equations that don't dissipate.

In Chapter 4 we consider a concrete example of coarse-graining; Specifically the coarse-graining of the conserved quantities of a one dimensional classical Hamiltonian particle system with nearest neighbor interactions periodically arranged on the Torus. The positions $\mathbf{x} = (x_1, \ldots, x_N)$ and $\mathbf{v} = (v_1, \ldots, v_N)$ and are governed by the Hamiltonian

$$H(\mathbf{x}, \mathbf{v}) = \sum_i \frac{1}{2} v_i^2 + V(x_i - x_{i-1}),$$

where $V(r)$ is an interparticle potential which is repulsive and singular at 0. The coarse-graining procedure involves dividing the N particles into M mesoscopic cells of size K, where $1 \ll K \ll N$, and averaging the inter-particle spacing $r_i = x_i - x_{i-1}$, momentum v_i, and energy $e_i = \frac{1}{2}v_i^2 + V(r_i)$ of the particles inside each cell, we refer to these averages as the *coarse variables*. Our goal will be to obtain a closed set of equations for the coarse variables when N and K are large. In Section 4.5.2, we show that if the microscopic particles are in equilibrium, then the coarse-grained quantities are equilibrium solutions to the so-called 'discrete Euler dynamics'

$$\dot{\ell}_i = p_i - p_{i-1}$$

$$\dot{p}_i = -P(\ell_{i+1}, e_{i+1} - \tfrac{1}{2}p_i^2) + P(\ell_i, e - \tfrac{1}{2}p_{i+1}^2)$$

$$\dot{e}_i = -p_i P(\ell_{i+1}, e_{i+1} - \tfrac{1}{2}p_{i+1}^2) + p_{i-1}P(\ell_i, e_i - \tfrac{1}{2}p_i^2),$$

where (ℓ_i, p_i, e_i) are to be interpreted as the length, momentum and energy of the ith coarse particle and P is the thermodynamic pressure function. The discrete Euler equations are a Hamiltonian discretization of the 1-D Euler equations in Lagrangian coordinates and conserve length, momentum, energy and entropy. Alternatively we may view this through the Mori-Zwanzig framework described above, working at the level of distribution functions on N particle phase space. The discrete Euler equations can then be seen as the leading order Galerkin truncation associated to a certain projection on N particle distributions.

Treating this coarse-graining procedure as map, which produces a coarse-grained entropy function $S(\ell, e)$ to govern the discrete Euler dynamics from a given potential V, we may repeatedly apply the coarse-graining operation to produce a mapping between entropy functions. Following the approach of renormalization group theory, we show that the ideal gas equation of state

$$S(\ell, e) = (c_P - c_V)\log(\ell) + c_V \log(e), \quad c_V > 1,$$

is a fixed point of this map. In this, case the discrete Euler dynamics reduce to those of a classical particle system with gamma-law potential $V(r) = Cr^{1-\gamma}$, $\gamma = c_P/c_V$, thereby justifying the use of power law potentials for mesoscopic descriptions.

When the system is not in equilibrium, we seek to obtain dissipative corrections to the discrete Euler dynamics. Here, we look at time scales of order K, and study the fluctuations about the discrete Euler dynamics. After the application of two key approximations related to convergence to equilibrium and decay of certain auto-correlation functions for large N and K we derive stochastic and dissipative corrections to the discrete Euler equation which model the effects of viscosity, thermal conductivity and transport in the volume variables, with coefficients given by analogues of the Green-Kubo formula. This is the main contribution of this chapter. The resulting dissipative fluid-particle model can be viewed as a discrete version of the *non-linear* equations of fluctuating hydrodynamics. The stochastic fluctuations are in 'fluctuation-dissipation' balance with the dissipation terms and they both conserve volume, momentum and energy. A more detailed presentation of this model can be found in the overview to Chapter 4, equation (4.6).

In Section 4.9 we present a simplification of the dissipative fluid-particle model, assuming that the transport coefficients are constant and studying it in more detail. We give a proof of well-posedness for the finite N stochastic system using the total entropy as a Lyapunov function. This implies that the volume and energy of a cell cannot collapse to zero in finite time. Indeed, the well-posedness is significant due to the difficult nature of proving well-posedness (even existence) for the corresponding non-linear fluctuating hydrodynamic equations that they discretize.

Background and Historical Remarks

Frameworks for understanding the connections between microscopic and macroscopic phenomena began development in the mid-to-late 19th century when the foundations of statistical mechanics were layed down by Gibbs, Boltzmann, Maxwell, and others. Here, fundamental concepts of equilibrium ensembles, microscopic foundations of thermodynamics and entropy, and kinetic theory were developed to make connections between microscopic and macroscopic systems, and to understand the nature of the irreversibility arising through randomness in the initial conditions. Later, in the mid 20th century, more modern theoretical foundations for statistical mechanics emerged, particularly for non-equilibrium statistical mechanics, we developed by Green, Kirkwood, Kubo, Mori, Onsager, Zwanzig and many others. The development of local equilibrium, the Green-Kubo formula, fluctuation-dissipation theorems, the theory of stochastic processes, and the Mori-Zwanzig formalism introduced a new set of machinery for understanding the emergence of irreversibility, as well the origins of transport phenomena like thermal conductivity and viscosity.

Of course, with the advent of modern scientific computing, there emerged yet another way to model macroscopic systems by directly simulating the dynamics of the microscopic system. This is the approach, for instance, taken in molecular dynamics (MD) simulations. However, while this might work in some simple situations, MD tends to be exceedingly expensive for systems of true macroscopic scales, and typically requires time-steps roughly proportion to one over the number of particles, making computations for any reasonable macroscopic length of time impractical. Of course, if one desires to model even larger systems like the climate, or the behavior of stars or galaxies, direct simulation is out of the question (and will likely never be an option).

Needless to say, it seems rather foolish to disregard convenient machinery of statistical mechanics in favor of a computational approach. In fact, it seems that much computational effort is wasted on 'irrelevant' chaotic dynamics at the small scales whose exact evolution seems to have very little effect on the large scale dynamics. It appears that what's needed is a synthesis of the methods of statistical mechanics and computational approaches. Namely a systematic theory of *coarse-graining* for the microscopic system. That is, method for producing a lower dimensional *coarse-grained* model that captures the large scale behavior at the expense of exact knowledge about the microscopic behavior. In fact, coarse-grained molecular models play a fundamental role in modern material simulations and allow the methods of molecular dynamics to by applied to larger systems and on longer time scales than typical microscopic models would. Examples of such models for studying hydrodynamic behavior are dissipative particle dynamics [72, 78], and smooth particle hydrodynamics [63].

The idea of coarse-graining, however, is as old as the foundations of statistical mechanics, originating from the ideas of Boltzmann in the equilibrium setting. Indeed, Boltzmann's original argument for the form of the microscopic entropy dividing phase space up into cells and counting particles in each cell, is essentially a coarse-graining argument. Perhaps, one of the first modern approaches to coarse-graining for non-equilibrium systems was an adaptation of Boltzmann's original argument introduced by P. Ehrenfest and T. Ehrenfest in the 1911 [40]. Here the dynamics of an classical N-body system is coarse-grained through the Liouville equation by periodically 'projecting the density' onto a maximal entropy state subject to certain constraints on its average over a family of cells. In between projections, the dynamics evolves again according to the Liouville equation (see [68] for a more in-depth discussion of the so-called 'Ehrenfest chain').

Since then, coarse-graining has become a central idea in statistical mechanics and other fields. Examples include block averaging in lattice dynamics [77], Wilson's renormalization group method [121], hydrodynamic and kinetic limits [111], optimal prediction methods [26], averaging in Hamiltonian systems [6], homogenization theory [15], heterogeneous multi-scale methods [120], and filtering methods in turbulence [85].

Outline of Part I

To summarize, Part I of the dissertation will be organized as follows:

In Chapter 3, we give an outline of the Mori-Zwanzig formalism. We outline a perturbative approach and give an example of its application to an ODE system. We present a scheme for obtaining higher order dissipative approximations to the coarse variables when the dissipation is large and show dissipativity of corrections of order $4m + 1$.

Chapter 4 deals with coarse-graining of a one-dimensional classical particle system. Here we present the one-dimension model, introduce the canonical and micro-canonical ensembles and discuss the thermodynamic structures associated with each ensemble. A general scheme for conservative coarse-graining is introduced, assigning different weights to each particle. In the particular case of coarse-graining by lumping we show the discrete Euler equations are satisfied in the equilibrium setting. In the non-equilibrium setting we make several approximations to derive the discrete Euler equations and a next order dissipative correction. We study a simplified version of the dissipative model in more detail and a well-posedness result is obtained using a Lyapunov function argument.

Mori-Zwanzig Formalism

Overview

In this chapter we discuss the Mori-Zwanzig formalism. In Section 3.2 we introduce the basic elements of the formal theory. We outline a perturbative approach in Section 3.3 based on a special decomposition of the generator into fast and slow modes. We then consider the problem of coarse-graining dissipative operators. Several approximations are discussed that preserve the dissipativity of the coarse-grained system.

The Formalism

The Mori-Zwanzig formalism, also referred to as the projection operator formalism, is one of staples of modern statistical mechanics and can be found in many modern physics textbooks [69, 102, 113]. It is named after H. Mori [95] and R. Zwanzig [123, 124] who were its early champions. The early approach by Mori was essentially a linear (or close to equilibrium) version of the later work by Zwanzig.

The basic building blocks of the Mori-Zwanzig theory are 1) a linear evolution equation,

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = \mathcal{L} f(t), \tag{3.1}$$

and 2) a projection $\mathcal{P}$ operator and its complement projection $\widetilde{\mathcal{P}} = I - \mathcal{P}$. The original application of Mori-Zwanzig was for a Hamiltonian systems, where $\mathcal{L}$ is the Liouville operator and $\mathcal{P}$ is a conditional average, although the formalism can be applied to the case when $\mathcal{L}$ is the generator of a Markov process, a C_0 semi-group on a Banach space, or a quantum Liouville equation in density matrix framework.

There appear to be essentially two approaches to the Mori-Zwanzig formalism, which are, roughly speaking dual to each other. One approach is to work directly with observables and make use of the so-called Dyson operator identity for the semi-group $e^{t\mathcal{L}}$

$$e^{t\mathcal{L}} = e^{t\widetilde{\mathcal{P}}\mathcal{L}} + \int_0^t e^{(t-s)\mathcal{L}} \mathcal{P} \mathcal{L} e^{s\widetilde{\mathcal{P}}\mathcal{L}} \mathrm{d}s, \tag{3.2}$$

which is just the usual perturbation formula in semi-group theory. However, it should be noted the if $\mathcal{L}$ and $\widetilde{\mathcal{P}}\mathcal{L}$ are unbounded operators, then the validity of (3.2) is far from obvious. Validity aside, one can use (3.2) to obtain the so-called generalized Langevin equation

$$\frac{\mathrm{d}}{\mathrm{d}t} x(t) = v(x(t)) + \int_0^t \operatorname{div} \gamma(x(s), t-s) \mathrm{d}s + \dot{\xi}(t). \tag{3.3}$$

where $x(t) = \langle X(t) \rangle_{f_0}$ is the averaged evolution of some observable of the Hamiltonian evolution $X(t)$, and the average is taken over the initial data with respect to an arbitrary initial distribution f_0 in phase space. The function $\dot{\xi}$ is interpreted as a noise term, and has correlation length

$$\mathbf{E}\left[\dot{\xi}(t) \otimes \dot{\xi}(s) \,\middle|\, A = x \right] = \gamma(x, t-s).$$

The matrix $\gamma(x,t)$ is sometimes referred to as the 'memory kernel', as it is responsible for the introduction of memory terms into the equation.

Another approach is to work directly with the distribution function on the phase space through the Liouville equation. This approach is more general, as is can be generalized to a broader class of evolution equations of the form (3.1) beyond the Liouville equation. Indeed this is the main approach that we will adopt for the rest of this chapter.

In this setting, the goal is to obtain a closed equation for the projected dynamics $\hat{f}(t) = \mathcal{P}f(t)$. Projecting both sides of (3.1) we obtain the non-closed equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{f}(t) = \widehat{\mathcal{L}}\hat{f}(t) + \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\tilde{f}(t), \tag{3.4}$$

where $\widehat{\mathcal{L}} = \mathcal{P}\mathcal{L}\mathcal{P}$ and $\tilde{f} = \widetilde{\mathcal{P}}f$. If one assumes that the initial data f_0 satisfies $\widetilde{\mathcal{P}}f_0 = 0$, then formally we have following equation for $\tilde{f}$

$$\tilde{f}(t) = \int_0^t e^{(t-s)\widetilde{\mathcal{L}}}\widetilde{\mathcal{P}}\mathcal{L}\hat{f}(s)\mathrm{d}s, \tag{3.5}$$

where $\widetilde{\mathcal{L}} = \widetilde{\mathcal{P}}\mathcal{L}\widetilde{\mathcal{P}}$. It should be noted that equation (3.5) is the analogue of the identity (3.2). It can be justified, for instance, if $\widetilde{\mathcal{L}}$ generates a C_0 semi-group and $\mathcal{P}\mathcal{L}$ is of *Desch-Schappacher* class with respect to $\widetilde{\mathcal{L}}$ (see [41]).

Substituting this into equation (3.4), we obtain the 'Nakajima-Zwanzig' equation,

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{f}(t) = \widehat{\mathcal{L}}\hat{f}(t) + \int_0^t \Psi(t-s)\hat{f}(s)\mathrm{d}s, \quad \Psi(t) = \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}e^{t\widetilde{\mathcal{L}}}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{P}, \tag{3.6}$$

which is the analogue of the generalized Langevin equation (3.3), with the operator $\Psi(t)$ playing the role of the 'non-Markovian' memory effects in the evolution of $\hat{f}$.

Equation (3.6) was first derived independently by Nakajima [99] and Zwanzig [123] in the context of quantum and classical systems respectively. Indeed, in the case of the Liouville equation, (3.3) can be obtained from (3.6) by integrating against a suitable choice of test function. In fact, equation (3.6) is equivalent to the formal operator identity

$$\mathcal{P}e^{\mathcal{L}t}\mathcal{P} = \mathcal{P} + \int_0^t \mathcal{P}\mathcal{L}\mathcal{P}e^{s\mathcal{L}}\mathrm{d}s + \int_0^t \int_0^s \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}e^{(s-r)\widetilde{\mathcal{L}}}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{P}e^{r\mathcal{L}}\mathrm{d}r\mathrm{d}s.$$

The utility of equation (3.6) is somewhat limited due to the memory effects introduced by $\Psi(t)$ as well as the intractability of the operator $\widetilde{\mathcal{L}}$. In general, and in specific examples, it is not clear that $\widetilde{\mathcal{L}}$ is a suitable generator for a semi-group, and such dynamics can be very tricky to compute. This makes the memory operator $\Psi(t)$ rather difficult to study. Several works, [28, 82], attempt to understand the behavior $\widetilde{\mathcal{L}}$ and the operator $\Psi(t)$ in a more rigorous fashion, but success is limited to very strong assumptions on the generator $\mathcal{L}$ and $\mathcal{P}$.

To circumvent these difficulties two approximations are typically made:

The first is an assumption of a time scale separation between $\Psi(t)$ and $f(t)$, that is, that $\Psi(t)$ decays suitably fast so that the following *Markov approximation* holds true,

$$\Psi(t) \approx \bar{\Psi}_T \delta(t), \quad \bar{\Psi}_T := \int_0^T \Psi(s)\mathrm{d}s, \tag{3.7}$$

Note that the integral is truncated at a finite time T, rather than taken over all of $\mathbb{R}_+$. This is typically done to avoid potential divergence of the integral, as well as to aid in computation. This approximation serves to remove the memory effect in equation (3.6) and render the dynamics 'Markovian'

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{f} = \mathcal{PLP}\hat{f} + \bar{\Psi}_T\hat{f}.$$

The Markovian assumption seems reasonable in many cases and is mainly an assumption on time-scale separation. Indeed, if one makes the right choice of relevant variables, then one typically observes correlations in the orthogonal dynamics decaying on a time-scale much faster than the evolution of the relevant dynamics. It should be noted here that this approximation breaks the equivalence of projected dynamics to the original evolution equation and potentially introduces some *dissipation* into the dynamics where there may have been none previously. Consequentially, one only expects such an approximation to valid in some appropriate limit where the scale separation becomes more pronounced.

The second assumption (which we will avoid), is that dynamics generated by $\widetilde{\mathcal{L}}$ are equivalent to $\mathcal{L}$, at least in the form that it arises in $\Psi(t)$, namely

$$\mathcal{PL}\widetilde{\mathcal{P}}e^{t\widetilde{\mathcal{L}}}\widetilde{\mathcal{P}}\mathcal{LP} \approx \mathcal{PL}\widetilde{\mathcal{P}}e^{t\mathcal{L}}\widetilde{\mathcal{P}}\mathcal{LP}.$$

This assumption is much harder to justify and is usually done as a technique to make Ψ_T computable. However, such an approximation, while convenient, can suffer from various deficiencies, among them the so-called *plateau-problem*, where Ψ_T has only a small range of value for which it is accurate before decaying to zero for large T (see [71]).

The Mori-Zwanzig formalism has had tremendous success in non-equilibrium statistical mechanics and has been applied successfully to countless problems. Indeed it is one of the standard methods used to derive the 'generalized Fokker-Planck' and 'generalized Langevin' equations from deterministic process. It has given great insight into the emergence of non-Markovian behavior through time-correlations and the emergence of dissipative and irreversible behavior through decay of correlations.

However, in part due to its extremely general nature, the Mori-Zwanzig formalism suffers from several undesirable features. The first is the reliance on the operator $\widetilde{\mathcal{L}}$ to generate the orthogonal dynamics. Indeed, except for a few special cases, $\widetilde{\mathcal{L}}$ cannot easily be shown to generate a good dynamics, and from a computational standpoint simulating such dynamics is an intractable problem. In addition, this intractability of $\widetilde{\mathcal{L}}$ makes any attempt to justify the Markov approximation (3.7) all the more difficult since any rigorous justification of the Markovian approximation will likely involve an ergodicity property of the operator $\widetilde{\mathcal{L}}$.

A Perturbative Approach

In order to avoid the complications present with the definition of $\widetilde{\mathcal{L}}$, we present here a more practical perturbative approach that allows for more explicit computations and construction of approximations. An similar approach can be found in [105], and is close (up to a time rescaling) to the work of Davies [29–31] on the so-called 'weak coupling limit'.

In this section, we will suppose that we have an explicit scale separation expressed through the decomposition

$$\mathcal{L} = \frac{1}{\epsilon}\mathcal{L}_0 + \mathcal{L}_1,$$

with ϵ playing the role of the scaling parameter. Systems exhibiting such a decomposition are often called 'fast-slow' systems with $\mathcal{L}_0$ generating the 'fast motion' and $\mathcal{L}_1$ generating the 'slow motion'. Fast-slow systems are abundant in the theory for averaging for Hamiltonian systems (see [17, 57, 58]). The projections $\mathcal{P}$ and $\widetilde{\mathcal{P}}$ can be viewed as projections onto slow and fast manifolds respectively.

In what follows, we will assume that $\mathcal{L}_0$ generates a strongly continuous semigroup $e^{t\mathcal{L}_0}$ and satisfies

$$\mathcal{P}\mathcal{L}_0 = \mathcal{L}_0\mathcal{P} = 0,$$

meaning that the 'fast motion' generated by $\mathcal{L}_0$ is constrained to the null space of $\widetilde{\mathcal{P}}$,

$$\mathcal{P}e^{t\mathcal{L}_0} = \mathcal{P}, \quad \text{and} \quad \widetilde{\mathcal{P}}\mathrm{e}^{t\mathcal{L}_0} = \mathrm{e}^{t\mathcal{L}_0}\widetilde{\mathcal{P}}.$$

The 'slow motion' generated by $\mathcal{L}_1$ need not be constrained to the null space of $\widetilde{\mathcal{P}}$, and may have a nontrivial projection under $\widetilde{\mathcal{P}}$. However, contrary to the non-perturbative approach, we will not need to assume that $\widetilde{\mathcal{P}}\mathcal{L}_1\widetilde{\mathcal{P}}$ generates a semigroup.

Here we have chosen to make the scaling ϵ explicit so as to have an explicit scale separation, however, in practice, it may also be embedded in $\mathcal{P}$ and the operator $\mathcal{L}$ (this is the case, for instance in the problem considered in Section 4.5).

The equation for $\hat{f} = \mathcal{P}f$ now reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{f} = \widehat{\mathcal{L}}\hat{f} + \mathcal{P}\mathcal{L}_1\tilde{f}, \tag{3.8}$$

while the orthogonal dynamics are given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{f} = \frac{1}{\epsilon}\mathcal{L}_0\tilde{f} + \widetilde{\mathcal{P}}\mathcal{L}_1 f.$$

Assuming $\widetilde{\mathcal{P}}f(0) = 0$, (formally) we may write

$$\tilde{f}(t) = \int_0^t e^{\epsilon^{-1}(t-s)\mathcal{L}_0}\widetilde{\mathcal{P}}\mathcal{L}_1 f_s \mathrm{d}s. \tag{3.9}$$

This Volterra-like operation on $f(s)$ can be justified if, for instance, $\widetilde{\mathcal{P}}\mathcal{L}_1$ is of *Desch-Schappacher* class with respect to $\mathcal{L}_0$ (see [41]). The main difference between this and equation (3.5) is that we have now written the orthogonal dynamics in terms of the more manageable evolution $e^{t\mathcal{L}_0}$ instead of $e^{t\widetilde{\mathcal{L}}}$. Of course, in doing this, we have paid the price that we are unable to close the dynamics. Indeed, substituting this into equation (3.8) we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{f}(t) = \widehat{\mathcal{L}}\hat{f}(t) + \int_0^t \mathcal{P}\mathcal{L}_1 e^{\epsilon^{-1}(t-s)\mathcal{L}_0}\widetilde{\mathcal{P}}\mathcal{L}_1 f_s \mathrm{d}s.$$

Keeping true to the Mori-Zwanzig philosophy, we write $f_s = \hat{f}_s + \tilde{f}_s$ on the right-hand side above and again apply (3.9). Iterating this procedure, we obtain the following formal series

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{f}(t) = \widehat{\mathcal{L}}\hat{f}(t) + \sum_{n=1}^{\infty}\int_{\Delta_n(t)} \Phi^n(\epsilon^{-1}(t-t_1), \epsilon^{-1}(t_1-t_2), \ldots, \epsilon^{-1}(t_{n-1}-t_n))\hat{f}_{t_n}\mathrm{d}t_1 \ldots \mathrm{d}t_n, \tag{3.10}$$

where $\Delta_n(t) = \{(t_1, \ldots, t_n) : 0 < t_n < \ldots < t_1 < t\}$ and the operator $\Phi^n(t_1, \ldots, t_n)$ is defined by

$$\Phi^n(t_1, \ldots, t_n) = \mathcal{P}\mathcal{L}_1 e^{t_1\mathcal{L}_0}\widetilde{\mathcal{L}}_1 e^{t_2\mathcal{L}_0}\widetilde{\mathcal{L}}_1 \ldots \widetilde{\mathcal{L}}_1 e^{t_n\mathcal{L}_0}\widetilde{\mathcal{P}}\mathcal{L}_1\mathcal{P}.$$

In a sense equation (3.10) is a generalization of (3.6), since $\mathcal{L}_0 = \widetilde{\mathcal{L}}$ implies that $\widetilde{\mathcal{L}}_1 = 0$ and then above series collapses to one term $n = 1$ with $\Phi^1(t) = \Psi(t)$. However, it is far from clear whether the series (3.10) is well defined and converges.

The operators $\{\Phi^n(t_1,\dots,t_n)\}_n$ encode a more complicated memory structure and are related nth order correlation functions, and assuming that $\mathcal{L}_0$ and $\mathcal{P}$ were chosen prudently, should contribute less and less for large n and epsilon. In general, one should interpret the series (3.10) as an asymptotic series in ϵ.

Indeed, assuming $\Phi^n(t_1,\dots,t_n)$ has enough decay, as $\epsilon \to 0$ we may regard $\epsilon^{-n}\Phi^n(\epsilon^{-1}t_1,\epsilon^{-1}t_2,\dots,\epsilon^{-1}t_n)$ as an approximation of the identity and make the following Markov approximation

$$\epsilon^{-n}\Phi^n(\epsilon^{-1}t_1,\epsilon^{-1}t_2,\dots,\epsilon^{-1}t_n) \sim \bar{\Phi}^n_\epsilon\,\delta(t_1,t_2,\dots,t_n),$$

where

$$\bar{\Phi}^n_\epsilon = \int_0^{T_\epsilon}\cdots\int_0^{T_\epsilon}\Phi_n(t_1,\dots,t_n)\mathrm{d}t_1,\dots\mathrm{d}t_n.$$

for some time $T_\epsilon \to \infty$ as $\epsilon \to 0$. This approximation then produces a Markovian equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{f} = \widehat{\mathcal{L}}\hat{f} + \sum_{n=1}^{\infty}\epsilon^n\bar{\Phi}^n_\epsilon\hat{f}. \tag{3.11}$$

Again, this series in (3.11) should be interpreted as an asymptotic series in ϵ and any approximation should truncate the series. Primarily we will be interested in the first order truncation to (3.11) governed by the operator

$$\overline{\mathcal{L}} = \widehat{\mathcal{L}} + \epsilon\mathcal{P}\mathcal{L}_1\int_0^T e^{t\mathcal{L}_0}\widetilde{\mathcal{P}}\mathcal{L}_1\mathrm{d}t$$

Dissipative Approximations

It is natural to wonder whether one might have better success justifying the Mori-Zwanzig formalism starting from a system that has some dissipation. Indeed this is the approach taken for rigorous deriving the equations of fluid mechanics, either from a Hamiltonian system that has some noise added (for instance [101]), or from the Boltzmann equation. Such approaches usually succeed where the pure Hamiltonian one fails, since the dissipation usually provides some form of ergodicity and a mechanism for equilibration.

Moreover, if one desires to further coarse-grain a system which has already been coarse-grained, then it natural to start with a system that has some dissipation. In this context, the Mori-Zwanzig formalism applied to dissipative, particularly diffusion processes, has been addressed by several authors ([44, 46, 103]).

In this section, we will suppose that the generator $\mathcal{L}$ acts on a Hilbert space H, so that it comes equipped with an inner product $\langle \cdot , \cdot \rangle$, and that $\mathcal{L}$ is dissipative

$$\langle f \, , \mathcal{L} f \rangle \leq 0, \quad \text{for all} \quad f \in D(\mathcal{L}).$$

We will mostly have in mind the case that $\mathcal{L}$ is the generator of a Markov process on a state space $\mathcal{X}$, and H is the space $L^2(\mu)$ where μ is an invariant measure for $\mathcal{L}$. Denote by $\mathcal{L}^*$ the formal adjoint of $\mathcal{L}$ under $\langle \cdot , \cdot \rangle$ and write

$$\mathcal{A} = \frac{1}{2}(\mathcal{L} - \mathcal{L}^*), \quad \mathcal{S} = \frac{1}{2}(\mathcal{L} + \mathcal{L}^*),$$

as its symmetric and antisymetric parts. We will also assume, for simplicity, that $\mathcal{L}$ $\widehat{\mathcal{L}}$ and $\widetilde{\mathcal{L}}$ all generate well defined (strongly continuous) semi-groups on H.

When the operator $\mathcal{L}$ has sufficient mixing properties and $\mathcal{S} \neq 0$ one can take the limit as $T \to \infty$ in the integral (3.7) and obtain

$$\bar{\Psi} := \int_0^\infty \Psi(s) \mathrm{d}s = \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}(-\widetilde{\mathcal{L}})^{-1}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{P},$$

where $\widetilde{\mathcal{L}}^{-1}$ is interpreted as a pseudo-inverse of $\widetilde{\mathcal{L}}$, to be interpreted through the resolvent limit

$$(-\widetilde{\mathcal{L}})^{-1} := \lim_{\lambda \to 0^+} (\lambda - \widetilde{\mathcal{L}})^{-1},$$

provided it exists.

In general, such a limit will produce a new operator

$$\overline{\mathcal{L}} = \widehat{\mathcal{L}} + \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}(-\widetilde{\mathcal{L}})^{-1}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{P}, \tag{3.12}$$

which, rather remarkably, will still be dissipative operator. This can be seen from the following identity reminiscent of the Shur complement

Lemma 3.4.1. *Suppose $\mathcal{L}^{-1}$ is invertible and let $\widetilde{\mathcal{L}}^{-1}$ be the a pseudo-inverse of $\widetilde{\mathcal{L}}$. Then $(\mathcal{P}\mathcal{L}^{-1}\mathcal{P})$ has a pseudo inverse and is given by*

$$\overline{\mathcal{L}} = (\mathcal{P}\mathcal{L}^{-1}\mathcal{P})^{-1}. \tag{3.13}$$

Proof. This can be checked by direct computation,

$$\begin{aligned}
\overline{\mathcal{L}}(\mathcal{P}\mathcal{L}^{-1}\mathcal{P}) &= \mathcal{P}\mathcal{L}\mathcal{P}\mathcal{L}^{-1}\mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\tilde{\mathcal{L}}^{-1}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{P}\mathcal{L}^{-1}\mathcal{P} \\
&= \mathcal{P}\mathcal{L}\mathcal{L}^{-1}\mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\mathcal{L}^{-1}\mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\tilde{\mathcal{L}}^{-1}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{L}^{-1}\mathcal{P} + \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\tilde{\mathcal{L}}^{-1}\widetilde{\mathcal{P}}\mathcal{L}\widetilde{\mathcal{P}}\mathcal{L}^{-1}\mathcal{P} \\
&= \mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\mathcal{L}^{-1}\mathcal{P} + \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\mathcal{L}^{-1}\mathcal{P} \\
&= \mathcal{P}
\end{aligned}$$

The same identity can easily be verified for left-multiplication $(\mathcal{P}\mathcal{L}^{-1}\mathcal{P})\overline{\mathcal{L}}$. It readily follows that $\overline{\mathcal{L}}$ is a pseudo-inverse for $(\mathcal{P}\mathcal{L}^{-1}\mathcal{P})$. □

A useful consequence of the above identity is that $\overline{\mathcal{L}}$ is a non-positive (dissipative) operator on $\widehat{H}$. Indeed, Lemma 3.4.1 immediately gives the following identity,

$$\overline{\mathcal{L}} = \overline{\mathcal{L}}^*(\mathcal{L}^*)^{-1}\overline{\mathcal{L}},$$

where * denote the adjoint. This, in turn, implies that

$$\langle f, \overline{\mathcal{L}} f \rangle = \langle \overline{\mathcal{L}} f, (\mathcal{L}^*)^{-1} \overline{\mathcal{L}} f \rangle \leq 0, \quad \text{for all } f \in H,$$

since the inverse of a non-positive operator $\mathcal{L}^*$ is also non-positive. There is another identity, similar to (3.13), which also proves useful for showing the dissipativity of $\overline{\mathcal{L}}$.

Lemma 3.4.2. *The following identity holds*

$$\overline{\mathcal{L}} = \mathcal{P}(I - \mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{L}}^{-1}\widetilde{\mathcal{P}})\mathcal{L}(I - \widetilde{\mathcal{P}}(\widetilde{\mathcal{L}}^*)^{-1}\widetilde{\mathcal{P}}\mathcal{L}^*)\mathcal{P}. \tag{3.14}$$

Proof. Again we check by direct computation,

$$\begin{aligned}
&(I - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{L}}^{-1}\widetilde{\mathcal{P}})\mathcal{L}(I - \widetilde{\mathcal{P}}(\widetilde{\mathcal{L}}^*)^{-1}\widetilde{\mathcal{P}}\mathcal{L}^*\mathcal{P}) \\
&\qquad = \mathcal{P}\mathcal{L}\mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{L}}^{-1}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}(\widetilde{\mathcal{L}}^*)^{-1}\widetilde{\mathcal{P}}\mathcal{L}^*\mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{L}}^{-1}\widetilde{\mathcal{P}}\mathcal{L}\widetilde{\mathcal{P}}(\widetilde{\mathcal{L}}^*)^{-1}\widetilde{\mathcal{P}}\mathcal{L}^*\mathcal{P} \\
&\qquad = \mathcal{P}\mathcal{L}\mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{L}}^{-1}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}(\widetilde{\mathcal{L}}^*)^{-1}\widetilde{\mathcal{P}}\mathcal{L}^*\mathcal{P} + \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}(\widetilde{\mathcal{L}}^*)^{-1}\widetilde{\mathcal{P}}\mathcal{L}^*\mathcal{P} \\
&\qquad = \overline{\mathcal{L}}.
\end{aligned}$$

□

The identity (3.14), of course, means that $\overline{\mathcal{L}}$ has the form $\mathcal{C}\mathcal{L}\mathcal{C}^*$, where $\mathcal{C} = \mathcal{P} - \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{L}}^{-1}\widetilde{\mathcal{P}}$ which means that $\overline{\mathcal{L}}$ is dissipative whenever $\mathcal{L}$ is.

It is not clear, however, that this dissipation property will be preserved upon making any approximations to $\widetilde{\mathcal{L}}$ as is usually done in the Mori-Zwanzig literature. Indeed, making the approximation $\widetilde{\mathcal{L}} \approx \mathcal{L}$ in the operator (3.12) above does not appear to preserve dissipativity, nor does the perturbative approach taken in Section 3.3.

In pursuit of dissipative approximations, we will again take a perturbative approach, and assume a decomposition of $\widetilde{\mathcal{L}}$ of the form

$$\widetilde{\mathcal{L}} = \widetilde{\mathcal{A}} + \epsilon^{-1}\widetilde{\mathcal{S}},$$

where $\widetilde{\mathcal{S}}$ is a scaling parameter. This decomposition amounts to the assumption that the dissipative part dominates the orthogonal dynamics, or in other words the dissipation dominates the small scales. For instance in kinetic theory, ϵ might be the knudsen number and $\mathcal{P}$ the projection onto the hydrodynamics fields.

Under such a scaling, one may may expand $(\widetilde{\mathcal{A}} + \epsilon^{-1}\widetilde{\mathcal{S}})^{-1}$ in a Neumann series to obtain

$$\overline{\mathcal{L}} = \widehat{\mathcal{L}} + \sum_{k=0}^{\infty} \epsilon^{k+1}\mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}(-\widetilde{\mathcal{S}})^{-1}(\widetilde{\mathcal{A}}(-\widetilde{\mathcal{S}})^{-1})^{k}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{P}, \tag{3.15}$$

which is formally equivalent to (3.12). As in Section 3.3, we will interpret the series in (3.15) as an asymptotic series and truncate to obtain approximations. Such truncations are defined by

$$\overline{\mathcal{L}}^{(n)} = \widehat{\mathcal{L}} + \sum_{k=0}^{n-1} \epsilon^{k+1}\mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}(-\widetilde{\mathcal{S}})^{-1}(\widetilde{\mathcal{A}}(-\widetilde{\mathcal{S}})^{-1})^{k}\widetilde{\mathcal{P}}\mathcal{L}\mathcal{P}.$$

The lowest order approximation $\mathcal{L}^{(0)} = \widehat{\mathcal{L}}$ is clearly dissipative since $\mathcal{L}$ is. However, not every truncation of (3.15) will lead to a generator $\overline{\mathcal{L}}^{(n)}$ which is dissipative. Interestingly, we will find that if $m \in \mathbb{N}$, then

$$\langle\, f \,,\, \overline{\mathcal{L}}^{(4m+1)} f \,\rangle \le 0.$$

This is analogous to the Chapman Enskog expansion of the Boltzmann equation, where certain truncations of the expansion lead to ill-posed equations that *do not* dissipate. This can be proven rigorously in the case the operator $\mathcal{L}$ is bounded and all pseudo-inverses are well-defined.

It is not clear whether the perturbative approach given in 3.3 can be combined with this method to produce dissipative approximations that do not rely on the operators $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{S}}$.

Our goal is now to find approximations to $\overline{\mathcal{L}}$ in the case that there is strong dissipative present in the system. We decompose $\mathcal{L}$ into an anti-symmetric part $\mathcal{A}$ and symmetric part $\mathcal{S}$, $\mathcal{L} = \mathcal{S} + \mathcal{A}$ and denote $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{S}}$ the same decomposition for $\widetilde{\mathcal{L}}$. We are interested in the case when $\widetilde{\mathcal{S}}$ is large relative to $\widetilde{\mathcal{A}}$. The key feature of the approximations that we would like to preserve here is the dissipativity. With this in mind, we formally expand $\widetilde{\mathcal{L}}^{-1} = (\widetilde{\mathcal{S}} + \widetilde{\mathcal{A}})^{-1}$ in a Neumann series,

$$\widetilde{\mathcal{L}}^{-1} = \sum_{k \geq 0} (-1)^k \widetilde{\mathcal{S}}^{-1} (\widetilde{\mathcal{A}} \widetilde{\mathcal{S}}^{-1})^k.$$

Substituting this into the expression for $\overline{\mathcal{L}}$ and truncating at the $n-1$th term, we define a sequence of approximations $\{\overline{\mathcal{L}}^{(n)} : n \in \mathbb{N}\}$, defined by

$$\overline{\mathcal{L}}^{(n)} = \widehat{\mathcal{L}} - \sum_{k \geq 0}^{n-1} (-1)^k \mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1} (\widetilde{\mathcal{A}} \widetilde{\mathcal{S}}^{-1})^k \widetilde{\mathcal{P}}\mathcal{L}\mathcal{P}, \tag{3.16}$$

where the sum is empty in the case that $n = 0$. The primary objective here is to study which of the truncations $\overline{\mathcal{L}}^{(n)}$ are dissipative operators. Our main result is the following:

Theorem 3.4.3. *The truncated approximation $\overline{\mathcal{L}}^{(n)}$, defined by (3.16), is dissipative when $n = 0$ and when $n = 4m + 1$ for each $m \in \mathbb{N}$.*

The cases $n = 0$ and $n = 1$ are fairly straight forward, and follow easily from the earlier discussion and identities. The result for larger values of n is far from obvious and requires a few algebraic ‘tricks’ to obtain. Indeed, as we will see from one of the following Lemmas, when $n = 4m + 2$, the summation term which subtracted from the right-hand side of (3.16) is actually dissipative, and therefore more work must be done to obtain dissipativity of the whole approximation $\overline{\mathcal{L}}^{(n)}$.

In order to prove this we will need a few Lemmas. The first is a very important identity, which will allow us the reduce the proof to showing the positivity of a certain sum. The identity can be see as a truncation of a formal expansion of identity (3.14).

Lemma 3.4.4. *For $n \geq 1$, the following identity holds*

$$\overline{\mathcal{L}}^{(n)} = \mathcal{P}\mathcal{H}\mathcal{P} - (-1)^{n-1}\mathcal{P}\mathcal{L}\widetilde{\mathcal{P}}\Big[\sum_{k\geq 0}^{n-1}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\Big](\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^n\widetilde{\mathcal{P}}\mathcal{L}^*\mathcal{P} \tag{3.17}$$

where $\mathcal{H}$ is a dissipative operator given by

$$\mathcal{H} = \Big(I - \sum_{k\geq 0}^{n-1}(-1)^k\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\widetilde{\mathcal{P}}\Big)\mathcal{L}\Big(I - \sum_{k\geq 0}^{n-1}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\widetilde{\mathcal{P}}\mathcal{L}^*\Big).$$

Proof. We begin by considering the dissipative operator $\mathcal{H}$, which is simply a truncation of a formal expansion of identity (3.14). Multiplying out the expression, we obtain

$$\begin{aligned}\mathcal{H} = \mathcal{L} - \sum_{k\geq 0}^{n-1}(-1)^k\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\widetilde{\mathcal{P}}\mathcal{L} - \sum_{k\geq 0}^{n-1}\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\widetilde{\mathcal{P}}\mathcal{L}^* \\ + \sum_{k,j\geq 0}^{n-1}(-1)^k\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\widetilde{\mathcal{L}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^j\widetilde{\mathcal{P}}\mathcal{L}^*.\end{aligned} \tag{3.18}$$

Writing $\widetilde{\mathcal{L}} = \widetilde{\mathcal{S}} + \widetilde{\mathcal{A}}$, the last term with the double summation on the right-hand side in (3.18) can be written as

$$\sum_{k,j\geq 0}^{n-1}(-1)^k\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^{k+j}\widetilde{\mathcal{P}}\mathcal{L}^* + \sum_{k,j\geq 0}^{n-1}(-1)^k\mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^{k+j+1}\widetilde{\mathcal{P}}\mathcal{L}^*$$

Clearly the above sums are telescoping in k and can be simplified to

$$\sum_{j\geq 0}^{n-1} \mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^j\widetilde{\mathcal{P}}\mathcal{L}^* + (-1)^{n-1}\sum_{j\geq 0}^{n-1} \mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^j(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^n\widetilde{\mathcal{P}}\mathcal{L}^*.$$

Substituting this expression back into (3.18) we obtain

$$\begin{aligned}\mathcal{H} = \mathcal{L} - &\sum_{k\geq 0}^{n-1}(-1)^k \mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\widetilde{\mathcal{P}}\mathcal{L} \\ &+ (-1)^{n-1}\sum_{j\geq 0}^{n-1} \mathcal{L}\widetilde{\mathcal{P}}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^j(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^n\widetilde{\mathcal{P}}\mathcal{L}^*.\end{aligned}$$

Using this to compute the product $\mathcal{PHP}$, and recognizing the appearance of $\overline{\mathcal{L}}^{(n)}$ from the first two terms, gives the main identity (3.14). □

The next Lemma regards dissipativity of truncations of the Neumann series expansion for $(\widetilde{\mathcal{L}}^*)^{-1}$.

Lemma 3.4.5. *The finite sums*

$$\sum_{k\geq 0}^{n} \widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k, \quad \textit{and} \quad -\sum_{k\geq 1}^{n} \widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k \tag{3.19}$$

are dissipative if $n = 4m$ *or* $n = 4m + 1$ *for some* $m \in \mathbb{N}$.

Proof. We begin by proving a simpler result, that is, for some symmetric, non-negative operator $\mathcal{B}$, the finite sum

$$\sum_{k\geq 0}^{n}(-1)^k\mathcal{B}^k$$

is a symmetric non-negative operator whenever n is even. Indeed this result easily follows from the following formula,

$$\sum_{k\geq 0}^{n}(-1)^k\mathcal{B}^k = (I - \mathcal{B}^{1/2})^{-1}(I + (-1)^n\mathcal{B}^{n+1})(I - \mathcal{B}^{1/2})^{-1},$$

which can be obtained by a simple computation. Indeed this fact now easily proves that

$$\sum_{k\geq 1}^{n}(-1)^k\mathcal{B}^k$$

is dissipative if n is even.

In order to prove that the series in (3.19) are dissipative, it suffices to prove it only for the symmetric part, which involves only the even terms in the sum. This also means that we may, without loss of generality assume that n is even, since proving dissipativity for any even n even will also imply dissipativity $n+1$ through the addition of an inconsequential anti-symmetric term.

The symmetric parts of the sums in (3.19) are given by,

$$\left(\sum_{k\geq 0}^{n}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\right)_{\text{sym}} = \sum_{k\geq 0}^{n/2}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^{2k},$$
$$\left(-\sum_{k\geq 1}^{n}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\right)_{\text{sym}} = -\sum_{k\geq 0}^{n/2}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^{2k}$$

Since $-\widetilde{\mathcal{S}}$ is symmetric and non-negative we may define the operator

$$\mathcal{B} = -\left[(-\widetilde{\mathcal{S}})^{-1/2}\widetilde{\mathcal{A}}(-\widetilde{\mathcal{S}})^{-1/2}\right]^2,$$

which is also symmetric and non-negative. We may then rewrite the sums above in terms of $\mathcal{B}$,

$$\sum_{k\geq 0}^{n/2}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^{2k} = -(-\widetilde{\mathcal{S}})^{-1/2}\left(\sum_{k\geq 0}^{n/2}(-1)^k\mathcal{B}^k\right)(-\widetilde{\mathcal{S}})^{-1/2},$$
$$-\sum_{k\geq 1}^{n/2}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^{2k} = (-\widetilde{\mathcal{S}})^{-1/2}\left(\sum_{k\geq 1}^{n/2}(-1)^k\mathcal{B}^k\right)(-\widetilde{\mathcal{S}})^{-1/2},$$

Clearly, by the results at the beginning of the proof, both quantities are dissipative only when $n/2$ is even. □

We can now use these Lemmas to prove the main Theorem.

Proof of Theorem 3.4.3. We will being by using identity (3.17). Clearly $\mathcal{PHP}$ is dissipative, so it suffices to show that the remaining series is also dissipative. Therefore to prove the theorem we simply need to prove dissipativity of

$$-(-1)^{n-1}\sum_{k\geq 0}^{n-1}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^n.$$

Assuming n is even, we may rewrite the expression above as

$$(\widetilde{\mathcal{S}}^{-1}\widetilde{\mathcal{A}})^{n/2}\left(\sum_{k\geq 0}^{n-1}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\right)(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^{n/2}.$$

By Lemma 3.4.5, if $n = 4m + 2$, then the above operator is positive, and therefore has the wrong sign.

Assuming n is odd, we may rewrite the expression instead as

$$-(\widetilde{\mathcal{S}}^{-1}\widetilde{\mathcal{A}})^{(n-1)/2}\left(\sum_{k\geq 1}^{n}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\right)(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^{(n-1)/2}. \tag{3.20}$$

We note that by Lemma 3.4.5

$$\sum_{k\geq 1}^{n}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k$$

is positive when $n = 4m + 1$. Therefore, since $(n-1)/2 = 2m$ the quantity (3.20) can be written as

$$-(\widetilde{\mathcal{S}}^{-1}\widetilde{\mathcal{A}}^*)^{(n-1)/2}\left(\sum_{k\geq 1}^{n}\widetilde{\mathcal{S}}^{-1}(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^k\right)(\widetilde{\mathcal{A}}\widetilde{\mathcal{S}}^{-1})^{(n-1)/2}.$$

which is clearly dissipative. □

Coarse-Graining of ODE's

Of course, the Mori-Zwanzig theory was originally studied for Hamiltonian systems. In this section we will narrow the discussion to the more concrete setting of ordinary differential equations (ODE's). In this setting, the nature of the Mori-Zwanzig formalism becomes more transparent and several approximations can be made more explicit.

We begin by considering the following ODE system

$$\dot{X}_t = b(X_t), \tag{3.21}$$

where $b : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth vector field. Suppose we have a smooth map $a : \mathbb{R}^n \to \mathbb{R}^m$ with $m < n$, which designates some interesting quality of the dynamics of (3.21), and suppose that it is non-degenerate, meaning that the matrix

$$G_{ij}(x) = \sum_k \partial_k a_i(x) \partial_k a_j(x)$$

is invertible for all $x \in \mathbb{R}^n$. We will refer to the map a as the *coarse-graining map*, and we will be interested in the behavior of the coarse dynamics $Y_t = a(X_t)$. Easily, Y_t satisfies the equation

$$\dot{Y}_t = \partial a(X_t) b(X_t),$$

where $(\partial a)_{ij} = \partial_j a_i$ denotes the Jacobian matrix. It is not surprising that this is not a closed equation in terms of Y_t, since Y_t is lower dimensional than X_t, and should not be determined in terms of Y_t unless X_t evolves transversely to the level sets of a. Our goal will be to obtain approximate closures for the evolution of Y_t.

We will find it useful to work in a probabilistic setting. Namely if X_0 is initially distributed according to a probability density $f_0(x)$, then the density $f(t,x)$ at later times $t>0$ is governed by the 'Liouville' equation

$$\partial_t f = \mathcal{L}^* f, \quad f|_{t=0} = f_0,$$

where $\mathcal{L} = b\cdot\nabla$, and $\mathcal{L}^* = \operatorname{div}(b\,\cdot\,)$ denotes the formal adjoint of $\mathcal{L}$. Suppose that (3.21) admits an invariant measure $\mu(\mathrm{d}x) = g(x)\mathrm{d}x$ (not necessarily probability) satisfying $\mathcal{L}^* g = 0$. Following the conventions of statistical physics, we will denote $\langle\,\cdot\,\rangle_\mu$ the average with respect to μ,

$$\langle u\rangle_\mu = \int_{\mathbb{R}^n} u\,\mathrm{d}\mu.$$

The coarse-graining map a naturally induces a coarse measure $\hat\mu = a_\#\mu$ and a fluctuation probability measure $\mu(\mathrm{d}x\,|\,y)$ obtain by conditioning μ on the event that $a(x)=y$. We will denote by $\langle\,\cdot\,\rangle_{\hat\mu}$ and $\langle\,\cdot\,|\,y\rangle_\mu$ the expectations with respect to $\hat\mu$ and $\mu(\cdot\,|\,y)$ respectively. Note that $\mu(\,\cdot\,|\,y)$ is a probability measure concentrated on the manifold $\Sigma_y = \{\,x\,:\,a(x)=y\}$, while $\hat\mu$ might not be (if μ isn't). These measures give rise to the decomposition

$$\mu(\mathrm{d}x) = \mu(\mathrm{d}x\,|\,y)\,\hat\mu(\mathrm{d}y),$$

which is to be interpreted by its action on test functions $\varphi(y)$ and $\psi(x)$

$$\int_{\mathbb{R}^n} \varphi(a(x))\psi(x)\,\mu(\mathrm{d}x) = \int_{\mathbb{R}^m}\left(\int_{\Sigma_y}\psi(x)\mu(\mathrm{d}x\,|\,y)\right)\hat\mu(\mathrm{d}y) \tag{3.22}$$

Define the operator $\mathcal{R}$ and its formal adjoint $\mathcal{R}^*$ (with respect to μ) by the action on a continuous bounded functions $\varphi(y)$, $\psi(x)$

$$\mathcal{R}\varphi(x) = \varphi(a(x)), \quad \mathcal{R}^*\psi(y) = \langle\psi\,|\,y\rangle_\mu\,.$$

The operators are adjoints in the sense that equation (3.22) can be rewritten as

$$\langle (\mathcal{R}\varphi)\, \psi \rangle_{\hat{\mu}} = \langle \varphi\, (\mathcal{R}^* \psi) \rangle_{\mu} .$$

Note that $\mathcal{R}^*\mathcal{R} = I$, so that

$$\mathcal{P} = \mathcal{R}\mathcal{R}^*, \quad \widetilde{\mathcal{P}} = I - \mathcal{P}$$

define projections.

We will find it useful to describe things in terms of the relative density $h(t, x) = f(t, x)/g(x)$, which solves

$$\partial_t h = \mathcal{L} h, \quad h|_{t=0} = h_0 = f_0/g, \tag{3.23}$$

whose solution is given by the action of the semi-group $e^{t\mathcal{L}}$

$$h(t, x) = e^{t\mathcal{L}} h_0(x) = h_0(\phi_t(x)).$$

where $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ is the flow of homeomorphisms associated to (3.21), defined by

$$\partial_t \phi_t(x) = b(\phi_t(x)), \quad \phi_0(x) = x.$$

We are interested in the distribution $\hat{f}(t, y)\mathrm{d}y = \hat{h}(t, y)\hat{\mu}(\mathrm{d}y)$ of Y_t defined by pushforward $\hat{f}(t, y)\mathrm{d}y = a_{\#}(f(t, x)\mathrm{d}x)$. From this we may deduce that $\hat{h}$ is given by

$$\hat{h}(y) = \mathcal{R}^* h(y) = \langle h \,|\, y \rangle_{\mu} .$$

Note that this framework lends itself to working in 'weak form' of (3.23),

$$\partial_t \langle \psi\, h \rangle_{\mu} = \langle \mathcal{L}\psi\, h \rangle_{\mu} .$$

where $\psi(x)$ is a suitably smooth test function. The process of coarse-graining then corresponds to choosing a test function of the form $\psi(x) = \mathcal{R}\varphi(x) = \varphi(a(x))$. Such a choice of test function yields

$$\partial_t \langle \varphi \, \hat{h} \rangle_{\hat{\mu}} = \langle (\mathcal{L}\mathcal{R}\varphi) \, h \rangle_{\mu} = \langle \mathcal{R}^* [(b \cdot \nabla a) h] \cdot \nabla \varphi \rangle_{\hat{\mu}}$$

We now have all the components for the Mori-Zwanzig formalism, namely a projection $\mathcal{P} = \mathcal{R}\mathcal{R}^*$ and an evolution equation (3.23). Lets now apply the perturbative framework of Section 3.3 and assume that the vector field b can be written as

$$b = \epsilon^{-1} b_0 + b_1,$$

where b_0 satisfies $b_0 \cdot \nabla a = 0$, and epsilon is an explicit scaling parameter identifying the speed of the fast and slow time scales. This in turn induces the decomposition of $\mathcal{L}$

$$\mathcal{L} = \epsilon^{-1} \mathcal{L}_0 + \mathcal{L}_1, \quad \mathcal{L}_0 = b_0 \cdot \nabla, \quad \mathcal{L}_1 = b_1 \cdot \nabla.$$

We will also assume that μ is an invariant measure for both $\mathcal{L}_0$ and $\mathcal{L}_1$ separately. Note that the fact that $\mathcal{L}_0 a = 0$ implies that $\mu(\,\cdot\,|\,y)$ is an invariant measure for $\mathcal{L}_0$ for each y.

The approximate Markovian equation (3.11) truncated at $n = 1$ is equivalent to the equation

$$\partial_t \hat{h}^1 = \mathcal{R}^* \mathcal{L}_1 \mathcal{R} \hat{h}^1 + \epsilon \mathcal{R}^* \bar{\Phi}^1_\epsilon \mathcal{R} \hat{h}^1. \tag{3.24}$$

The operator $\mathcal{R}\mathcal{L}_1\mathcal{R}^*$ can be easily shown to satisfy

$$\mathcal{R}\mathcal{L}_1\mathcal{R}^* = \hat{b} \cdot \nabla, \quad \hat{b}(y) = \langle b_1 \cdot \nabla a \,|\, y \rangle_\mu$$

and similarly the operator $\mathcal{R}\bar{\Phi}^1_\epsilon\mathcal{R}^*$ satisfies

$$\left\langle \varphi\,\mathcal{R}\bar{\Phi}^1_\epsilon\mathcal{R}^*\psi\right\rangle_{\hat{\mu}} = \langle (M\nabla\psi)\cdot\nabla\varphi\rangle_{\hat{\mu}}$$

with

$$M(y) = \int_0^{T_\epsilon} \left\langle \tilde{b}\otimes e^{t\mathcal{L}_0}\tilde{b}\,\middle|\,y\right\rangle_\mu \mathrm{d}t, \quad \tilde{b}(x) = b(x)\cdot\nabla a(x) - \hat{b}(a(x)). \tag{3.25}$$

The formula (3.25) for the matrix $M(y)$ is an analogue of the famous *Green Kubo* formula, and is usually called the friction matrix. Note that M it is not necessary a symmetric matrix, because of the potential lack of time-symmetry of the operator $\mathcal{L}_0$ and parity of the flux. However for any $\xi\in\mathbb{R}^m$, we have

$$(M(y)\xi)\cdot\xi \geq 0,$$

since it is a time integral of an auto-correlation function, and therefore the Wiener-Kinchin theorem implies that it is positive for large enough T_ϵ.

If $\hat{h}^1$ satisfies (3.24), then the measure $\hat{\nu}^1 = \hat{h}^1\hat{\mu}$ satisfies a Kolmogorov (Fokker-Planck) equation

$$\partial_t\hat{\nu}^1 + \mathrm{div}\big(\hat{b}\hat{\nu}^1\big) = \epsilon\,\mathrm{div}\Big(\hat{\mu}M\nabla\hat{h}^1\Big)\,.$$

If $\hat{\nu}$ has a density $\hat{g}$ with respect to Lebesgue measure, then the above Kolmogorov equation corresponds to an Itô diffusion process

$$\dot{Y}_t = \bar{b}(Y_t) + \sqrt{2\epsilon D(Y_t)}\dot{W}_t, \quad \bar{b} = \hat{b} + \epsilon M\nabla\log\hat{g} + \epsilon\,\mathrm{div}\,M,$$

where D denotes the symmetric part of M and $\sqrt{D}$ denotes the square root matrix.

In general, there is no clear strategy on how to pick the vector field b_0 as does above, and, in general, it is not clear that such a decomposition even exists for any

a. However, the basic perturbative strategy above will be applied in Chapter 4 to the coarse-graining of a one-dimensional particle system, in this case there is a very clear choice for b_0.

Taking higher order truncations of the Markovian equation (3.11) will lead to higher-order derivatives in the equation for $\hat{h}$ and contain coefficients contain higher order time-correlations functions. It is not clear what the utility of such a higher order approximation might be as there does not appear to be any stochastic process associated with such an equation. Nevertheless, such an approximation may be useful for computing higher order corrections to the evolution of the distribution $\hat{\mu}$.

In addition, the perturbative framework does not play well with stochastic differential equations, where the generator $\mathcal{L}$ is a second order operator. Indeed, directly applying the first order truncation of (3.11) to this example produces a fourth order differential equation, which again does not appear to correspond to any stochastic process.

Coarse-Graining of a One-dimensional Particle System

Overview

In this chapter we turn to a more concrete example of coarse-graining of classical particle systems. Coarse-graining classical N-particle Hamiltonian systems is of fundamental interest in statistical mechanics and many related fields. Continuum equations in fluid mechanics and kinetic theory can be viewed as coarse-grained models of such a system. However, it is often desirable, from the perspective of computations, to obtain certain coarse-grained descriptions that allow the coarse-grained model to be 'tuned' to the regime of interest, and will need to incorporate both macroscopic and micro-scopic fluctuations. In general, this is a difficult task, especially if one has any hope of obtaining rigorous results. Indeed, even in the case of simple fluids, it is not even clear how to properly incorporate the fluctuations and dissipation into a macroscopic model.

In order to simplify the picture, we will consider a Hamiltonian system of N particles in *one dimension* with positions $\mathbf{x} = (x_1, \ldots, x_N) \in \mathbb{T}^N$ and $\mathbf{v} = (v_1, \ldots, v_N) \in \mathbb{R}^N$, satisfying periodic boundary conditions and interacting through nearest neighbors. The particles are governed by the Hamiltonian

$$H(\mathbf{x},\mathbf{u}) = \sum_{i=1}^{N} \left(\frac{1}{2} v_i^2 + V(x_i - x_{i-1}) \right),$$

and the potential V is singular enough at the origin so that particles cannot cross. Here, it is useful to introduce the deformation coordinates $r_i = x_i - x_{i-1}$ and view the particle system $(\mathbf{r}, \mathbf{v})$ as a lattice system on $\mathbb{Z}_N = \mathbb{Z} \backslash N\mathbb{Z}$. Indeed, in these coordinates $(\mathbf{r}, \mathbf{v})$, the particle system now takes the form of a one-dimensional *anharmonic chain*, which has been widely studied in the literature. The equations of motion are

$$\begin{aligned} \dot{r}_i &= v_i - v_{i-1} \\ \dot{u}_i &= V'(r_{i+1}) - V'(r_i) \\ \dot{e}_i &= (v_i V'(r_{i+1}) - v_{i-1} V'(r_i)). \end{aligned} \tag{4.1}$$

Typically, if interested in the large scale hydrodynamic behavior of the system, one studies the empirical measure η_N on $\mathbb{T}$, defined by

$$\eta_N(t) = \frac{1}{N} \sum_{i=1}^{N} w_i(Nt) \delta_{i/N},$$

where $w_i(Nt) = (r_i(Nt), v_i(Nt), \frac{1}{2} v_i^2(Nt) + V(r_i(Nt)))$ is the Hamiltonian evolution of the locally conserved quantities sped up by a factor of N. In this scaling, as $N \to \infty$, one typically expects $\eta_N(t)$ to converges weakly to the fluid densities $(\ell(x), p(x), e(x))$ satisfying the one dimensional Euler equations in Lagrangian form,

$$\begin{aligned} \partial_t \ell &= \partial_x p \\ \partial_t p &= -\partial_x P(\ell, e - \tfrac{1}{2} p^2) \\ \partial_t e &= -\partial_x \left(p P(\tau, e - \tfrac{1}{2} p^2) \right), \end{aligned}$$

where (ℓ, p, e) are the volume, momentum and energy densities, and $P(\ell, e)$ is the thermodynamic pressure function obtained from a concave entropy function $S(\ell, e)$ satisfying the first law of thermodynamics

$$\partial_\ell S(\ell, e) = \beta(\ell, e)P(\ell, e) > 0, \quad \partial_e S(\ell, e) = \beta(\ell, e). \tag{4.2}$$

It should be mentioned that the hydrodynamic limit cannot be rigorously proven directly from the underlying Hamiltonian system without assumptions of ergodicity of the deterministic Hamiltonian system, a fact that is notoriously difficult to prove. Typically, to get around this, one introduces certain momentum and energy conserving stochastic perturbations to the system to obtain the required mixing. In this setting such a limit can be proven rigorously using relative entropy methods (see [14] for a proof).

Often, one is interested in higher order corrections to the system above, taking into account diffusive (or super diffusive) transport effects that might appear on times scales of order N^α, $\alpha > 1$. Since we are in dimension 1, and the particle system has no pinning potential, the corrections are expected to be super-diffusive (see [11, 75]) and therefore the typical Naiver-Stokes corrections are not expected to hold. This, of course does not stop one from studying the one-dimensional Navier-Stokes equations, which can be instead thought of as a model for a higher dimensional fluid with a large degree of symmetry (slab symmetry).

Of course, one can not simply look at times scales of order N^α, since the 'Euler' part of the dynamics will blow up in such a scaling. Often, this can be studied by looking at the fluctuation measure $\xi_N(t)$

$$\xi_N(t) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left\{ w_i(N^\alpha t) - \langle w_i(N^\alpha t) \rangle \right\} \delta_{i/N},$$

where $\langle \cdot \rangle$ denotes an ensemble average. In general, one expects that ξ_N converges to a stochastic process which is governed by the linearized Euler system as well as by a dissipative and stochastic part satisfying a fluctuation dissipation relation. Such linearized stochastic evolution is often referred to as *fluctuating hydrodynamics* (see [111]).

It is important to note that it is very difficult to capture both the nonlinear Euler dynamics and any *nonlinear* dissipative corrections as an exact scaling limit due to lack of scale invariance of the target equations. Therefore, in order to capture both the Euler and Navier-Stokes behavior, one must forego any attempt to obtain exact scaling limits and instead find approximations which, in some sense, asymptotically describe the hydrodynamic behavior of the system in a certain regime.

Since we are in one dimension, and the particle system has an interpretation as a lattice system, we may approach the problem of coarse graining by lumping conserved quantities into certain cells of mesoscopic size, that is, cells which contain a large number of particles, but a small number relative to N. Specifically, partition T_N into M cells $\Lambda = \{\Lambda_i\}_{i \in T_N}$ of equal size $K = N/M$ and define a local averaging map

$$\widehat{\mathbf{w}}(\mathbf{r}, \mathbf{v})_i = \frac{1}{K} \sum_{j \in \Lambda_i} w_j. \tag{4.3}$$

If $\mathbf{r}(t)$ and $\mathbf{r}(t)$ satisfy the original Hamiltonian dynamics, then, analogous to the hydrodynamic limit, we expect that $\widehat{\mathbf{w}}(\mathbf{r}(Kt), \mathbf{v}(Kt))$ will converge (in a statistical sense) as $K \to \infty$ with $N/K \to \infty$ to an infinite particle system $\mathbf{U}(t) \in (\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R})^{\mathbb{Z}}$, $U_i = (\ell_i, p_i, e_i)$ satisfying the *discrete Euler equations*,

$$\dot{\ell}_i = p_i - p_{i-1}$$

$$\dot{p}_i = -P(\ell_{i+1}, e_{i+1} - \tfrac{1}{2}p_i^2) + P(\ell_i, e - \tfrac{1}{2}p_{i+1}^2)$$

$$\dot{e}_i = -p_i P(\ell_{i+1}, e_{i+1} - \tfrac{1}{2}p_{i+1}^2) + p_{i-1} P(\ell_i, e_i - \tfrac{1}{2}p_i^2)$$

where $P(\ell, e)$ is the same thermodynamic pressure function obtained for the continuous Euler system from the entropy function $S(\ell, e)$. It is easy to check that the discrete Euler system is a Poisson manifold and that for each i, the entropy $S(\ell_i, e_i - \frac{1}{2}p_i^2)$ is a constant of the motion. Furthermore, one can produce a family of invariant probability measures $\{\mu_\alpha : \alpha \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+\}$ on $(\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+)^{\mathbb{Z}}$

$$\mathrm{d}\mu_\alpha = \prod_{i \in \mathbb{Z}} \frac{1}{Z(\alpha)} \exp\Big\{ -\alpha \cdot U_i + S(\ell_i, e_i - \tfrac{1}{2}p_i^2) \Big\} \beta(\ell_i, e_i - \tfrac{1}{2}p_i^2)\, \mathrm{d}\ell_i \mathrm{d}p_i \mathrm{d}e_i, \qquad (4.4)$$

analogous to the grand-canonical measures of classical statistical mechanics.

The benefit of this approach is that the limit system is still a particle system, but with a fluid character, and that the limiting dynamics has an explicit (Gibbs) invariant probability measure. Indeed, this allows us to re-apply the same lumping procedure to this discrete Euler system, scaling the cell size in the same way as time. We, again, expect such a procedure to produce the same discrete Euler system as above, just with a different thermodynamic structure. In the language of the renormalization group, this means that the discrete Euler equations lie on an invariant set with respect to the coarse-graining procedure. Seeking fixed points of the thermodynamic functions, one can show that the *ideal gas* entropy (up to an additive constant)

$$S(\ell, e) = (c_P - c_V)\log(\ell) + c_V \log(e), \quad c_V > 1,$$

remains invariant under coarse-graining. Where c_V and c_P are the specific heats at constant volume and pressure. In this case, the Discrete Euler equations simplify to the so-called *gamma-law*

$$\begin{aligned}
\dot{\ell}_i &= p_i - p_{i-1} \\
\dot{p}_i &= (1-\gamma)e^{S_0/R}\left(\frac{1}{(\ell_{i+1})^\gamma} - \frac{1}{(\ell_i)^\gamma}\right),
\end{aligned}$$

where S_0 is the initial entropy, $R = c_p - c_V$ is the gas constant and $\gamma = c_P/c_v > 1$ is the heat capacity ratio. What's interesting is that this system is again a one-dimensional particle chain with Hamiltonian

$$H_{\text{Ideal}} = \sum_i \frac{1}{2}p_i^2 + e^{S_0/R}(\ell_i)^{1-\gamma}$$

Of course, just as with the hydrodynamic limit, rigorously proving such results is well out of reach due to lack of ergodicity of the underlying Hamiltonian system. Again, one approach to remedy this is to add certain energy and momentum conserving stochastic perturbations to the dynamics.

Corrections to Discrete Euler

We would like to consider corrections to the discrete Euler dynamics that take into account longer-time dissipative phenomena. As discussed, such effects are not easy to obtain in conjunction with discrete Euler dynamics in any sort of limiting regime. Indeed if one scales so that the dissipative effects are of order one, the discrete Euler part of the dynamics will blow up. If instead one subtracts off the Euler dynamics, and studies the fluctuations on the right time scale, the limiting stochastic equation will be linear.

Our goal is to try to capture both the leading order Euler dynamics as well as the 'second order' dissipative stochastic dynamics through the coarse-graining procedure outlined above. To do this, we will follow the strategy of the Mori-Zwanzig perturbative approach described in Section 3.3 applied the Liouville equation

$$\partial_t f^N + \mathcal{A}_N f^N = 0,$$

where $\mathcal{A}_N$ is the Liouville operator associated to (4.1), and the solution f_t^N is the density of particles in phase space at time $t > 0$. Let K be the size of the cell and $M = N/K$ be the number of cells, with $\widehat{\mathbf{w}}$ be corresponding local averaging map defined in (4.3). We choose a Gibbs measure μ^N as a reference invariant measure for $\mathcal{A}_N$ and denote $\hat{\mu}^M = \widehat{\mathbf{w}}_{\#}\mu^N$ the push forward and $\mu_N(\,\cdot\,|\,\mathbf{y}_M)$ the measure conditioned on $\{\widehat{\mathbf{w}} = y\}$, whose expectation we denote by $\langle\,\cdot\,|\,\mathbf{y}_M\,\rangle_N$.

Following the perturbative Mori-Zwanzig approach, we decompose $\mathcal{A}_N$ into

$$\mathcal{A}_N = \mathring{\mathcal{A}}_M + \overline{\mathcal{A}}_M,$$

where $\mathring{\mathcal{A}}_M$ is the operator corresponding Hamiltonian motion inside each cell suitably periodized so that the cells do not interact, and $\overline{\mathcal{A}}_M$ corresponds to boundary interactions between the cells. The operator $\mathring{\mathcal{A}}_M$ plays the role of the operator $\mathcal{L}_0$ since

$$\mathring{\mathcal{A}}_M \widehat{\mathbf{w}} = 0,$$

and $\mu_N(\,\cdot\,|\,\mathbf{y}_M)$ is an invariant measure for $\mathring{\mathcal{A}}_M$ for each $\mathbf{y}_M$. The density of the coarse-particles is given by push forward $\hat{f}_t^M = \widehat{\mathbf{w}}_\# f_t^N$. If f_t^N is initially distributed according μ^N, then the system is in equilibrium, namely $\mathcal{A}_N f^N = 0$. In this case, one case show that *for any* K, N, $\hat{f}^M$ solves

$$\widehat{\mathcal{A}}_M^* \hat{f}^M = 0$$

where $\widehat{\mathcal{A}}_M$ is the generator of a finite M discrete Euler dynamics.

When f_t^N is not in equilibrium, after rescaling in time $t \to Kt$, we make two approximations under the assumption large K, N. The first is a *relaxation approximation*

$$\langle\,\cdot\,|\,\mathbf{y}_M\rangle_{f_{Kt}^N} \approx \langle\,\cdot\,|\,\mathbf{y}_M\rangle_N,$$

where $\langle\,\cdot\,|\,\mathbf{y}_M\rangle_{f_{Kt}^N}$ corresponds measure obtained by conditioning the distribution f_{Kt}^N on $\{\widehat{\mathbf{w}} = \mathbf{y}_M\}$. This approximation is essentially a statement of local equilibrium implying that the measure f_t^N equilibrates within the cells faster than the cells do. Indeed, one expects this to be valid in a regime where N and K are large, but N is much larger than K. In comparison to the perturbative Mori-Zwanzig approach shown earlier, the *relaxation approximation* is simply a more precise justification of truncation of the series (3.11). The second approximation is a *Markovian assumption*, which is expected to be valid in the large K (long time) limit.

After these approximations, we obtain a Fokker-Planck equation for $\hat{f}_t^M$,

$$\partial_t \hat{f}_t^M - \widehat{\mathcal{A}}_M^* \hat{f}_t^M = K^{-1} \sum_{i \in \mathbb{Z}_M} \operatorname{div}_{i-1,i} \left(g_K^M \, \mathrm{d}_i \nabla_{i-1,i} \left(\frac{\hat{f}_t^M}{g_K^M} \right) \right). \tag{4.5}$$

$\operatorname{div}_{i-1,i} := \operatorname{div}_{y_i} - \operatorname{div}_{y_{i-1}}$, $\nabla_{i-1,i} := \nabla_{y_i} - \nabla_{y_{i-1}}$, and $g_K^M \mathrm{d}\mathbf{y}_M = \widehat{\mathbf{w}}_\# \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}$ is the tensor product of the density of states inside each cell. The matrix $\mathrm{d}_i = \mathrm{d}(y_{i-1}, y_i)$

is the diffusion matrix and is defined by

$$\mathrm{d}(y_{i-1},y_i) =$$

$$\begin{pmatrix} T(y_{i-1})\bar\theta(y_{i-1}) & 0 & 0 \\ 0 & T(y_i)\bar\eta(y_i) & p_{i-1}T(y_i)\bar\eta(y_i) \\ 0 & p_{i-1}T(y_i)\bar\eta(y_i) & T(y_{i-1})T(y_i)\bar\kappa(y_{i-1},y_i) + T(y_i)\bar\eta(y_i)p_{i-1}^2 \end{pmatrix}$$

with $T(y) = \beta(y)^{-1}$ and $P(y)$ are *micro-canonical* temperature and pressure functions associated to the so-called *volume entropy*

$$S_V(\ell, e) := \log\left(\int_{V(\ell)}^{e} g_K(\ell, 0, u)\mathrm{d}u\right),$$

where $g_K(\ell, p, e)$ is the micro-canonical density of states and $T(y), P(y)$ are related to $S_V(y)$ through the first law (4.2). The functions $\bar\theta(y)$, $\bar\eta(y)$ given by time integrals of auto-correlation functions with respect to the micro-canonical measure $\langle\,\cdot\,|\,y\rangle_K$ on $(\mathbb{R}_+ \times \mathbb{R})^K$, analogous to the *Green-Kubo* formula,

$$\bar\theta(y) = \frac{1}{T(y)}\int_0^K \left\langle \frac{1}{K}\sum_{j=1}^K \left(e^{t\mathcal{A}_K}v_j - p_i\right)\left(v_j - p_i\right)\,\middle|\, y\right\rangle_K \mathrm{d}t$$

$$\bar\eta(y) = \frac{1}{T(y)}\int_0^K \left\langle \frac{1}{K}\sum_{j=1}^K \left(e^{t\mathcal{A}_K}V'(r_j) + P(y_i)\right)\left(V'(r_j) + P(y_i)\right)\,\middle|\, y\right\rangle_K \mathrm{d}t.$$

and $\bar\kappa(y_{i-1}, y_i)$ is given in terms of $\bar\theta$ and $\bar\eta$,

$$\bar\kappa(y_{i-1},y_i) = \bar\theta(y_{i-1})\bar\eta(y_i) + \bar\theta(y_{i-1})\beta(y_i)P(y_i)^2.$$

If K is large enough we can ensure that

$$\bar\theta, \bar\eta, \bar\kappa \geq 0.$$

The SDE system associated with (4.5) can be written as

$$
\begin{aligned}
\dot{\ell}_i &= (p_i - p_{i-1}) + K^{-1}\left(\mathcal{J}^\ell(y_{i+1}, y_i) - \mathcal{J}^\ell(y_i, y_{i-1})\right) + K^{-1}(\dot{\mathcal{M}}^\ell_{i+1} - \dot{\mathcal{M}}^\ell_i) \\
\dot{p}_i &= -(P(y_{i+1}) - P(y_i)) + K^{-1}\left(\mathcal{J}^p(y_{i+1}, y_i) - \mathcal{J}^p(y_i, y_{i-1})\right) + K^{-1}(\dot{\mathcal{M}}^p_{i+1} - \dot{\mathcal{M}}^p_i) \\
\dot{e}_i &= -(p_i P(y_{i+1}) - p_{i-1} P(y_i)) \\
&\qquad + K^{-1}\left(\mathcal{J}^e(y_{i+1}, y_i) - \mathcal{J}^e(y_i, y_{i-1})\right) + K^{-1}(\dot{\mathcal{M}}^e_{i+1,i} - \dot{\mathcal{M}}^e_{i,i-1})
\end{aligned}
\tag{4.6}
$$

where $(\mathcal{J}^\ell_{i,i-1}, \mathcal{J}^p_{i,i-1}, \mathcal{J}^e_{i,i-1})$ are the dissipative fluxes given by

$$
\begin{aligned}
\mathcal{J}^\ell_{i,i-1} &= T_{i-1}\bar{\theta}_{i-1}(\beta_i P_i - \beta_{i-1} P_{i-1}) + \beta_i \partial_\ell \bar{\theta}_{i-1} + \bar{\theta}_{i-1}\partial_\ell \bar{\eta}_i - P_i \partial_{e_i} \bar{\kappa}_{i-1,i} \\
\mathcal{J}^p_{i,i-1} &= (\bar{\eta}_i + T_i \partial_e \bar{\eta}_i)(p_i - p_{i-1}) \\
\mathcal{J}^e_{i,i-1} &= p_{i-1}\mathcal{J}^p_{i,i-1} + T_i \bar{\eta}_i + \bar{\kappa}_{i,i-1}(T_i - T_{i-1}) - T_i T_{i-1}(\partial_{e_i}\bar{\kappa}_{i,i-1} - \partial_{e_{i-1}}\bar{\kappa}_{i,i-1}),
\end{aligned}
\tag{4.7}
$$

and $(\mathcal{M}^\ell_i, \mathcal{M}^p_i, \mathcal{M}^e_{i,i-1})$ are mean-zero martingales, given by stochastic integration against a collection of independent Wiener processes $\{W^\ell_i\}, \{W^p_i\}, \{W^e_i\}$

$$
\begin{aligned}
\dot{\mathcal{M}}^\ell_i &= \sqrt{2T_{i-1}\bar{\theta}_{i-1}}\,\dot{W}^\ell_i \\
\dot{\mathcal{M}}^p_i &= \sqrt{2T_i \bar{\eta}_i}\,\dot{W}^p_i \\
\dot{\mathcal{M}}^e_{i,i-1} &= u_{i-1}\dot{\mathcal{M}}^p_i + \sqrt{2\bar{\kappa}_{i,i-1}T_i T_{i-1}}\,\dot{W}^e_i.
\end{aligned}
\tag{4.8}
$$

In equations (4.7) and (4.8) we have used subscripts to denote dependence certain coarse particle, for instance $P_i = P(y_i)$.

The system (4.6) is a discrete model for the Landau-Lifshitz-Navier-Stokes equations in Lagrangian form and is derived in Chapter 4. The quantity $\bar{\eta}$ plays the role of the bulk viscosity, while $\bar{\kappa}$ plays the role of the thermal conductivity. There are, however, some additional terms in the equation that don't usually appear in the Navier stokes equations. Indeed, the quantity $\bar{\theta}$ does not directly have an analogue in Navier stokes equations, as typically the density equation doesn't dissipate. However, in this setting, dissipation terms are due to correlations of the fluxes between cells and resemble the auto-correlations present in a tracer particle rather than average correlations between all the particles. It is interesting to notice as well that here, the thermal conductivity has an exact expression in terms of the $\bar{\eta}$, $\bar{\theta}$ and some thermodynamic quantities. This is a consequence of the fact that fluxes between cells are solely determined by the fluxes on the boundaries. In addition, we observe the emergence of terms that depend on derivatives of the quantities $\bar{\theta}, \bar{\eta}$ and $\bar{\kappa}$.

It should be noted that this system conserves total length, momentum and energy, and the measure

$$\mathrm{d}\mu_\alpha^M = \prod_{i\in\mathbb{Z}_M} \frac{1}{Z(\alpha)} \exp\Big\{ -\alpha \cdot U_i + S(\ell_i, e_i - \tfrac{1}{2}p_i^2)\Big\} \beta(\ell_i, e_i - \tfrac{1}{2}p_i^2)\, \mathrm{d}\ell_i \mathrm{d}p_i \mathrm{d}e_i, \quad (4.9)$$

is an invariant measure for both the Euler and the dissipative part of the dynamics separately.

The equations (4.6) are very similar a popular model called 'Dissipative Particle Dynamics' (DPD). The DPD model was initially developed by Hoogerbrugge and Koelman [72, 78] as model to simulate complex fluids, it has since been generalized [42, 43, 91, 108] to produce consistent equilibrium behavior and to conserve energy. Generally speaking, DPD consists of a collection of 'fluid parcels' that have, volume, momentum, and internal energy, interacting with various friction terms that corresponds to viscosity and thermal conductivity, and perturbed by stochastic 'fluctuations' which are in fluctuation-dissipation balance with the friction. There have also been several other attempts to derive dissipative particle dynamics from Hamiltonian mechanics in the literature [45, 54, 55].

At the moment, without more knowledge of the behavior of the functions $\bar{\theta}, \bar{\eta}, \bar{\kappa}$ it is not clear that the SDE has a global in time solution, indeed the coefficients are not locally Lipschitz. However, since the length, energy, and momentum are all conserved, the only possible blow-up that could occur is if one of the cells attains zero volume or zero energy in finite time.

In order to simplify matters, in Section 4.9 we introduce a simplified version of the model (4.6) by assuming constant transport coefficients $\bar{\theta}$, $\bar{\eta}$ and $\bar{\kappa}$. The model takes the form

$$\begin{aligned}
\dot{\ell}_i &= (p_i - p_{i-1}) + T_{i-1}\bar{\theta}(\beta_i P_i - \beta_{i-1} P_{i-1}) + \dot{\mathcal{M}}^{\ell}_{i+1} - \dot{\mathcal{M}}^{\ell}_i \\
\dot{p}_i &= (P_i - P_{i+1}) + [\bar{\eta}(p_{i+1} - p_i) - \bar{\eta}(p_i - p_{i-1})] + \dot{\mathcal{M}}^{p}_{i+1} - \dot{\mathcal{M}}^{p}_i \\
\dot{e}_i &= (p_{i-1} P_i - p_i P_{i+1}) + \bar{\eta}\left[u_i(p_{i+1} - p_i) - p_{i-1}(p_i - p_{i-1})\right] \\
&\quad + \bar{\kappa}\left[(T_{i+1} - T_i) - (T_i - T_{i-1})\right] + \bar{\eta}(T_i - T_{i-1}) \\
&\quad + p_i \dot{\mathcal{M}}^{p}_{i+1} - p_{i-1}\dot{\mathcal{M}}^{p}_i + \dot{\mathcal{M}}^{e}_{i+1} - \dot{\mathcal{M}}^{e}_i .
\end{aligned}$$

Such a model has a clearer structure and the local entropy dissipation becomes more apparent. In this setting, one can show that global strong solutions exist, we prove this in Theorem 4.9.1.

This chapter is organized as follows:

In Section 4.2 we introduce the particle system and discuss in detail the invariant measure and thermodynamic structure. In Section 4.5 we introduce a conservative coarse-graining scheme by lumping the lattice points into cells and discuss the coarse-graining in equilibrium giving rise to Discrete Euler dynamics. In Section 4.6 we discuss more detail about the discrete Euler dynamics and discuss its invariant measure and thermodynamics structures. In Section 4.8 we address the problem of non-equilibrium coarse-graining by lumping. Under a relaxation assumption and a Markovian approximation, we obtain a stochastic particle system for the coarse-grained cells which resembles a discretization of the non-linear Landau-Lifshitz-Navier-Stokes equations of fluctuating hydrodynamics in Lagrangian form. In Section 4.9 we introduce a simplified version of this stochastic particle system and study its behavior. In particular, we show that the system is well-posed under certain conditions on the entropy function.

A Classical Particle System in 1-D

In this section, we discuss properties of the one-dimensional particle model we wish to coarse-grain. We give a precise formulation of the system and give a detailed discussion of its invariant measure and limiting thermodynamic structure.

Suppose that we have a collection of N particles with unit mass, periodically arranged on the torus $\mathbb{T}_L$ of size L. The positions are given by $\mathbf{x} = \{x_i\}_{i\in\mathbb{Z}_N} \in \mathbb{T}_L^{\mathbb{Z}_N}$ and the velocities are $\mathbf{v} = \{v_i\}_{i\in\mathbb{Z}_N} \in \mathbb{R}^{\mathbb{Z}_N}$ where $\mathbb{Z}_N = \mathbb{Z}\backslash N\mathbb{Z}$ denotes the N-periodic one dimensional lattice. We will assume that the positions $\mathbf{x}$ arranged on $\mathbb{T}_L$ in an ordered configuration in the space $\mathbb{O}_L^N$, where $\mathbb{O}_L^N \subseteq \mathbb{T}_L^{\mathbb{Z}_N}$ denotes the set of all ordered configurations on $\mathbb{T}_L$. More precisely, given an identification of $\mathbb{T}_L$ with the interval $[0, L]$ then we say $\mathbf{x} \in \mathbb{O}_L^N$ if there exists a cyclic permutation of

$\{1, \ldots, N\}$, call it σ, such that

$$0 \leq x_{\sigma(1)} < x_{\sigma(2)} < \ldots < x_{\sigma(N)} < L.$$

We assume that the particles interact only with their neighbors through a pair potential function $V(r)$ and governing the evolution of the N particles is a Hamiltonian H_N, taking the form

$$\mathsf{H}_N(\mathbf{x}, \mathbf{v}) = \sum_{i \in \mathbb{Z}_N} \left(\frac{1}{2} v_i^2 + V(x_i - x_{i-1}) \right).$$

The particles then evolve according to Hamilton's equations

$$\begin{aligned} \dot{x}_i &= v_i \\ \dot{v}_i &= -V'(x_i - x_{i-1}) + V'(x_{i+1} - x_i), \end{aligned}$$

and are initially arranged on $\mathbb{T}_L$ in an ordered configuration in $\mathbb{O}_L^N$. We will make the following assumptions on the potential

Hypothesis 4.2.1. *The potential $V : \mathbb{R}_+ \to \mathbb{R}$ is a non-negative, smooth, non-increasing, convex function on the interior of $\mathbb{R}_+$, and satisfies*

$$\lim_{r \to 0} V(r) = +\infty,$$

while, $V'(r)$ is a smooth concave function on the interior of $\mathbb{R}_+$ and satisfies

$$\lim_{r \to \infty} V'(r) = 0.$$

The singularity of the potential V implies that the particles cannot cross. This ensures that any initial configuration in $\mathbb{O}_L^N$ remains in $\mathbb{O}_L^N$ under the evolution of the dynamics.

As the phase space $\mathbb{O}_L^N$ of ordered configurations is rather painful to work with, we will find it convenient to change coordinates to deformation variables

$$r_i = x_i - x_{i-1} \in \mathbb{R}_+, \quad i \in \mathbb{Z}_N,$$

describing the relative distance between neighboring particles. The deformation variables $\mathbf{r} = \{r_i\}_{i\in\mathbb{Z}_N}$ take values in the simplex Δ_L^{N-1} defined by the *total length* constraint

$$\mathsf{L}_N(\mathbf{r}) \equiv \sum_{i\in\mathbb{Z}_N} r_i = L.$$

Of course, such a change of variables is *not* one-to-one, since the coordinates $\mathbf{r} = \{r_i\}_{i\in\mathbb{Z}_N}$ are invariant under translations of $\mathbb{T}$ and are constrained to the simplex Δ_L^{N-1}. However, given the position of one particle, say $x_1 \in \mathbb{T}$, one can reconstruct the positions $\mathbf{x}$ uniquely from from $\mathbf{r}$, by the formula

$$x_i = x_1 + \sum_{j=1}^{i} r_j.$$

Indeed, it is not difficult to see that the mapping

$$\mathbb{O}_L^N \ni (x_1, x_2, \ldots, x_N) \mapsto (x_1, r_1, \ldots, r_N) \in \mathbb{T}_L \times \Delta_L^{N-1}$$

is a volume preserving diffeomorphism.

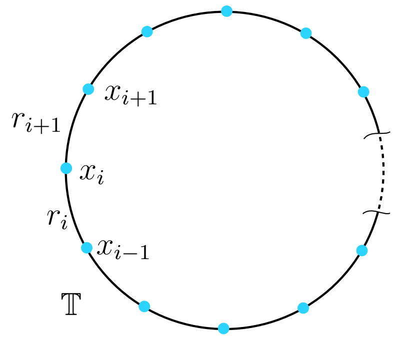

Figure 4.1: The periodic arragement of particles on a circle

Under these new coordinates we define the *phase space* $\Omega^N = \mathbb{R}_+^{\mathbb{Z}_N} \times \mathbb{R}^{\mathbb{Z}_N}$ and obtain the following evolution equation

$$\begin{aligned} \dot{r}_i &= v_i - v_{i-1} \\ \dot{v}_i &= V'(r_{i+1}) - V'(r_i), \end{aligned} \tag{4.10}$$

with the new Hamiltonian

$$\mathsf{H}_N(\mathbf{r}, \mathbf{v}) = \sum_{i \in \mathbb{Z}_N} E(r_i, v_i), \quad E(r, v) = \frac{1}{2} v^2 + V(r).$$

The above system has *three* conserved quantities, the Hamiltonian, or *total energy* H_N, the *total momentum* P_N, and the *total length* L_N, where

$$\mathsf{P}_N(\mathbf{v}) = \sum_{i \in \mathbb{Z}_N} v_i.$$

Remark 4.2.2. It is important to the equations (4.10) are no longer canonically Hamiltonian, due to the degeneracy associated with the conserved quantity L_N. In fact, the dynamics in (4.10) define a Poisson structure with Poisson bracket

$$\{f, g\} = \sum_{i \in \mathbb{Z}_N} \left[\partial_{r_i} f \left(\partial_{v_i} g - \partial_{v_{i-1}} g \right) - \partial_{r_i} g \left(\partial_{v_i} f - \partial_{v_{i-1}} f \right) \right],$$

acting on smooth functions. In this setting Hamiltonian governing the evolution is still H_N, while L_N is a Casimir invariant, that is

$$\{\mathsf{L}_N, f\} = 0,$$

for all suitably smooth functions f.

Associated with the dynamics is the *Liouville operator*

$$\mathcal{A}_N = \sum_{i\in\mathbb{Z}_N} -v_{i-1}(\partial_{r_i} - \partial_{r_{i-1}}) - V'(r_i)(\partial_{v_i} - \partial_{v_{i-1}}) = \{\,\cdot\,, \mathsf{H}_N\}\,,$$

which governs the evolution of observables and distributions of particles over Ω^N. The fact that $\mathsf{L}_N, \mathsf{P}_N$ and H_N are conserved is expressed by the fact that they belong to the null space of $\mathcal{A}_{\mathsf{H}_N}$,

$$\mathcal{A}_N\mathsf{L}_N = \mathcal{A}_N\mathsf{P}_N = \mathcal{A}_N\mathsf{H}_N = 0.$$

In particular, if one is only interested in statistical properties of the particle $(\mathbf{r}, \mathbf{v})$, then the probability density $f^N(t, \mathbf{r}, \mathbf{v})$, describing the density of particles with positions and velocities $(\mathbf{r}, \mathbf{v})$ in Ω^N at time t is given by the *Liouville equation*

$$\partial_t f^N + \mathcal{A}_N f^N = 0. \tag{4.11}$$

Grand-Canonical Ensemble

Associated with the conserved quantities $\mathsf{H}_N, \mathsf{P}_N, \mathsf{L}_N$, is the *grand canonical ensemble*, that is, a measure $\mu^N_{\tau,\beta,\lambda}(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v})$ on the phase space Ω^N with parameters $(\tau, \lambda, \beta) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$. It is defined by

$$\mu^N_{\tau,\lambda,\beta}(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}) = \frac{1}{\mathcal{Z}^N(\tau, \lambda, \beta)}\mathrm{e}^{-\tau\mathsf{L}_N(\mathbf{r})-\lambda\mathsf{P}_N(\mathbf{v})-\beta\mathsf{H}_N(\mathbf{r},\mathbf{v})}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v},$$

where

$$\mathcal{Z}^N(\tau,\lambda,\beta) = \int_{\Omega^N} \mathrm{e}^{-\tau \mathsf{L}_N(\mathbf{r}) - \lambda \mathsf{P}_N(\mathbf{v}) - \beta \mathsf{H}_N(\mathbf{x},\mathbf{v})} \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v},$$

is the grand canonical partition function, which serves as a normalization constant for $\mu^N_{\ell,\beta,\lambda}$. Because the Hamiltonian H_N is just a sum of single particle energies E_i, it is easy to see that $\mu^N_{\tau,\beta,\lambda}(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v})$ is just a product measure

$$\mu^N_{\tau,\lambda,\beta}(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}) = \prod_{i \in \mathbb{Z}_N} \mu_{\tau,\lambda,\beta}(\mathrm{d}r_i \mathrm{d}v_i),$$

where $\mu_{\tau,\lambda,\beta}(\mathrm{d}r\mathrm{d}v)$ is the single particle Gibbs measure

$$\mu_{\tau,\lambda,\beta}(\mathrm{d}r\mathrm{d}v) = \frac{1}{\mathcal{Z}(\tau,\lambda,\beta)} \mathrm{e}^{-\tau r - \lambda v - \beta(\frac{1}{2}v^2 + V(r_i))} \mathrm{d}r\mathrm{d}v$$

and $\mathcal{Z}(\tau,\lambda,\beta)$ is the single particle partition function

$$\mathcal{Z}(\tau,\lambda,\beta) = \int_{\mathbb{R}_+ \times \mathbb{R}} \mathrm{e}^{-\tau r - \lambda v - \beta(\frac{1}{2}v^2 + V(r_i))} \mathrm{d}r\mathrm{d}v.$$

The measure $\mu_{\tau,\lambda,\beta}(\mathrm{d}r\mathrm{d}v)$ can also be written as a product of a Gaussian measure in velocity and another measure in r, namely

$$\mu_{\tau,\lambda,\beta}(\mathrm{d}r\mathrm{d}v) = \left[\frac{\mathrm{e}^{-\frac{1}{2}\beta(v_i - \beta^{-1}\lambda)^2}}{\sqrt{2\pi\beta^{-1}}} \mathrm{d}v \right] \left[\frac{1}{Z(\tau,\beta)} \mathrm{e}^{-\tau r - \beta V(r)} \mathrm{d}r \right]$$

where

$$Z(\tau,\beta) = \int_{\mathbb{R}_+} \mathrm{e}^{-\tau r - \beta V(r)} \mathrm{d}\mathbf{r}.$$

It is important to note that under the assumption that since $V(r)$ is non-increasing, in order for $Z(\tau,\beta)$ to be finite, we need $\tau > 0$.

It is easy to see that $\mu^N_{\tau,\lambda,\beta}$ is a stationary measure for the dynamics. In fact, it is a consequence of the more general skew-symmetry property $\mathcal{A}_N$ with respect to $\mu^N_{\tau,\lambda,\beta}$.

Lemma 4.2.3. *The operator $\mathcal{A}_N$ is skew-symmetric with respect to $\mu^N_{\tau,\lambda,\beta}$. That is, for each $F, G \in C^1_b(\Omega^N)$, we have*

$$\int_{\Omega^N} F \mathcal{A}_N G \, \mathrm{d}\mu^N_{\tau,\lambda,\beta} = - \int_{\Omega^N} \mathcal{A}_N G \, F \, \mathrm{d}\mu^N_{\tau,\lambda,\beta}.$$

Proof. With some abuse of notation, we write $\mu^N_{\tau,\lambda,\beta}(\mathbf{r}, \mathbf{v})$ as the density of the measure $\mu^N_{\tau,\lambda,\beta}(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v})$. Note that since $\mu^N_{\tau,\lambda,\beta}$ is a function of the conserved quantities and $\mathcal{A}_N$ is a first order differential operator, we have $\mathcal{A}_N \mu^N_{\tau,\lambda,\beta} = 0$. The proof then follows from the fact that $\mathcal{A}_N$ is skew-symmetric with respect to the Lebesgue measure, since Hamiltonian vector fields are divergence free. □

The quantities $\beta^{-1}\lambda$ and β play the usual role of mean velocity and inverse temperature for the measure $\mu_{\tau,\beta,\lambda}$, as can be see by computing the Gaussian integrals,

$$\begin{aligned} &\int_{\mathbb{R}_+ \times \mathbb{R}} v \, \mu_{\tau,\lambda,\beta}(\mathrm{d}r\mathrm{d}v) = \beta^{-1}\lambda, \\ &\int_{\mathbb{R}_+ \times \mathbb{R}} \frac{1}{2}(v - \beta^{-1}\lambda)^2 \, \mu_{\tau,\lambda,\beta}(\mathrm{d}r\mathrm{d}v) = \frac{1}{2}\beta^{-1}, \end{aligned} \tag{4.12}$$

where the second identity for $\frac{1}{2}(v - \beta^{-1}\lambda)^2$ is a manifestation of the equipartition theorem of statistical mechanics. The quantity $\beta^{-1}\tau$ also has a physical interpretation. In fact, a special feature of plays the role of the pressure (or tension) of the segments between particles, as it follows from a simple integration by parts and an appeal to the behavior of $V(r)$ at 0 and ∞ that

$$\int_{\mathbb{R}_+ \times \mathbb{R}} -V'(r) \, \mu_{\tau,\lambda,\beta}(\mathrm{d}r) = \beta^{-1}\tau.$$

Grand-Canonical Thermodynamic Structure

The thermodynamic free energy F associated with the grand-canonical measure is defined by taking the logarithm of the one-particle partition function

$$F(\tau, \lambda, \beta) = \log Z(\tau, \beta) + \tfrac{1}{2}\beta^{-1}\lambda^2 - \tfrac{1}{2}\log\beta + \tfrac{1}{2}\log 2\pi.$$

The corresponding *thermodynamic entropy* S is given by Legendre-Fenchel transform

$$S(\ell, p, e) = \inf_{\tau,\lambda,\beta} \left[\tau\ell + \lambda p + \beta e + F(\tau, \lambda, \beta)\right]$$

where the infimum is taken over all $(\tau, \lambda, \beta) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$. Note that we have altered the definition of the entropy from that of section A.2 to match the physical notion of entropy, and to think of the parameters (ℓ, p, e) as the physical values of average length, momentum and energy, respectively. Indeed, if $\tilde{S}(\ell, p, e)$ represents the entropy as defined in Section A.2, then S and $\tilde{S}$ are related by

$$\tilde{S}(\ell, p, e) = -S(-\ell, -p, -e).$$

It follows, by Lemma A.2.1 and Lemma A.2.2 that F is a smooth strictly convex function on $(\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+)$ and S is a smooth strictly concave function on its domain. Moreover, dual pairs of Legendre variables $\alpha = (\tau, \lambda, \beta)$ and $y = (\ell, p, e)$ satisfy

$$\alpha = \nabla S(y), \quad y = -\nabla F(\alpha).$$

The entropy can computed more explicitly using the structure of F. Indeed, taking the infimum over λ first, we find

$$S(\ell, p, e) = \inf_{\tau,\beta} \left[\tau\ell + \beta\left(e - \tfrac{1}{2}p^2\right) - \tfrac{1}{2}\log\beta + \log Z(\tau, \beta)\right] + \frac{1}{2}\log 2\pi. \tag{4.13}$$

In particular, this implies that $S(\ell, p, e)$ only depends on the average momentum p and average energy e through the internal energy $u = e - \frac{1}{2}p^2$. This is a consequence of the *Galilean invariance* of the system (4.10). Particularly this property can be written as

$$S(\ell, p, e) = S(\ell, 0, e - \tfrac{1}{2}p^2).$$

We will find it convenient to define for each (ℓ, e) the inverse temperature function

$$\beta(\ell, u) = \partial_e S(\ell, 0, u) > 0$$

and the pressure function

$$P(\ell, u) = \partial_\ell S(\ell, 0, u)/\partial_e S(\ell, 0, u).$$

The fact that $\beta(\ell, u)$ is strictly positive follows from Gaussian nature of the measure $\mu_{\tau,\lambda,\beta}$ in velocity and the formulas (4.12). Then it is seen that the function $S_0(\ell, u) = S(\ell, 0, u)$ satisfies the *first law of thermodynamics*

$$\mathrm{d}S_0 = \beta P \mathrm{d}\ell + \beta \mathrm{d}u.$$

It is important to remark, that because of the exclusion effects of V, the domain of S

$$D_S = \{(\ell, p, e) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+ \,:\, |S(\ell, p, e)| < \infty\}$$

will be a non-trivial subset $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$. Indeed, small values of ℓ will restrict how small e can be. In fact, the convexity assumption on V allows form a precise definition of D_S.

Lemma 4.2.4. *The under Hypothesis 4.2.1 on V, the domain D_S is given by*

$$D_S = \left\{(\ell, p, e) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+ \,:\, e \geq \tfrac{1}{2}p^2 + V(\ell)\right\}. \tag{4.14}$$

Proof. It suffices to show that the domain of $S(\ell, 0, e)$ are the values of $(\ell, e) \in \mathbb{R}_+^2$ such that $e \geq V(\ell)$. Since F is smooth an convex, this is equivalent to showing that

$$\ell = -\partial_\tau F(\tau, 0, \beta), \quad e = -\partial_\beta F(\tau, 0, \beta), \tag{4.15}$$

is uniquely invertible for $\ell > 0$ and $e > V(\ell)$.

First, we remark that for fixed $\beta > 0$, the following limits hold

$$\lim_{\tau\to\infty} \frac{1}{Z(\tau,\beta)} \int_{\mathbb{R}_+} r\,\mathrm{e}^{-\tau r - \beta V(r)} \mathrm{d}r = 0, \quad \lim_{\tau\to 0} \frac{1}{Z(\tau,\beta)} \int_{\mathbb{R}_+} r\,\mathrm{e}^{-\tau r - \beta V(r)} \mathrm{d}r = \infty.$$

Therefore, by the monotonicity of $\partial_\tau \log(Z(\tau, \beta))$, for each $\ell > 0$ and $\beta > 0$, there exits a unique $\tau_{\ell,\beta}$ such that

$$\ell = -\partial_\tau \log(Z(\tau_{\ell,\beta}, \beta)) = \frac{1}{Z(\tau,\beta)} \int_{\mathbb{R}_+} r\,\mathrm{e}^{-\tau r - \beta V(r)} \mathrm{d}r.$$

Next, we claim that for $\tau_{\ell,\beta}$, we have

$$\lim_{\beta\to\infty} \frac{1}{\mathcal{Z}(\tau_{\ell,\beta}, 0, \beta)} \int_{\mathbb{R}_+\times\mathbb{R}} \left(\frac{1}{2}v^2 + V(r)\right) \mathrm{e}^{-\tau_{\ell,\beta} r - \beta\left(\frac{1}{2}v^2 + V(r)\right)} \mathrm{d}r\mathrm{d}v = V(\ell).$$

Indeed the fact that

$$\lim_{\beta\to\infty} \frac{1}{\sqrt{2\pi\beta^{-1}}} \int_{\mathbb{R}} \frac{1}{2} v^2 \mathrm{e}^{-\beta\frac{1}{2}v^2} \mathrm{d}v = 0.$$

follows from simple Gaussian integration, while the fact that

$$\lim_{\beta\to 0} \frac{1}{Z(\tau_{\beta,\ell}, \beta)} \int_{\mathbb{R}_+} V(r) \mathrm{e}^{-\tau_{\ell,\beta} r - \beta V(r)} \mathrm{d}r = V(\ell),$$

follows in a straight forward manner from the fact that the measure $Z(\tau,\beta)^{-1}\mathrm{e}^{-\tau_{\ell,\beta}r-\beta V(r)}$ concentrates at its mean value $r=\ell$, as $\beta\to\infty$. Moreover, it is easy to see that $\tau_{\ell,\beta}\to\ell^{-1/2}$ as $\beta\to 0$ and conclude that

$$\lim_{\beta\to 0}\frac{1}{\mathcal{Z}(\tau_{\ell,\beta},0,\beta)}\int_{\mathbb{R}_+\times\mathbb{R}}\left(\frac{1}{2}v^2+V(r)\right)\mathrm{e}^{-\tau_{\ell,\beta}r-\beta\left(\frac{1}{2}v^2+V(r)\right)}\mathrm{d}r\mathrm{d}v=\infty,$$

Also, since V is convex, we have by Jensen's inequality

$$\frac{1}{\mathcal{Z}(\tau,0,\beta)}\int_{\mathbb{R}_+\times\mathbb{R}}\left(\frac{1}{2}v^2+V(r)\right)\mathrm{e}^{-\tau_{\ell,\beta}r-\beta\left(\frac{1}{2}v^2+V(r)\right)}\mathrm{d}r\mathrm{d}v\geq V(\ell).$$

It follows, again from the monotonicity of $\beta\mapsto\partial_\beta F(\tau_{\ell,\beta},\beta)$ that for each $\ell>0$ and $e>V(\ell)$ there exists a unique $\beta_{\ell,e}$ that satisfies

$$e=-\partial_\beta F(\tau_{\ell,\beta_{\ell,e}},0,\beta_{\ell,e}).$$

In addition, the above limits show that the interior of $\{(\ell,e)\in\mathbb{R}_+\times\mathbb{R}_+ : e\geq V(\ell)\}$ are the only values for which (4.15) have a solution. □

Micro-Canonical Ensemble

While the grand-canonical ensemble is rather convenient to work with, being a product measure, it does disregard the fact that the particle evolution associated to (4.10) is actually constrained to certain lower dimensional submanifolds of Ω^N. Indeed, the evolution takes place on the manifold defined by energy, momentum, and length conservation.

To be more precise, suppose that we fix values $(\ell,p,e)\in D_S\subseteq\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+$ and define the manifold

$$\Sigma^N_{\ell,p,e}=\left\{(\mathbf{r},\mathbf{v})\in\Omega^N : \mathsf{L}_N(\mathbf{r})=N\ell,\ \mathsf{P}_N(\mathbf{v})=Np,\ \mathsf{H}_N(\mathbf{r},\mathbf{v})=Ne\right\}.$$

It is important to note that this manifold is only non-empty for certain values of ℓ and e. Indeed, under the constraints

$$\frac{1}{N}\sum_{i=1}^{N} r_i = \ell, \quad \frac{1}{N}\sum_{i=1}^{N} v_i = p$$

the energy has the sharp lower bound

$$\frac{1}{N}\sum_{i=1}^{N} \frac{1}{2}v_i^2 + V(r_i) \geq \frac{1}{2}p^2 + V(\ell).$$

Therefore in order for $\Sigma_{\ell,p,e}^{K}$ to be non-empty, we will need $e \geq \frac{1}{2}p^2 + V(\ell)$ for any given $\ell > 0$. As it turns out, this condition is precisely the one that defines the domain of the thermodynamic entropy D_S defined in (4.14). Specifically, we have

$$\left\{(\ell, p, e) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \,:\, \Sigma_{\ell,p,e}^{N} \neq \emptyset\right\} = D_S.$$

We refer to any minimizing state $(\mathbf{r}, \mathbf{v})$ of the Hamiltonian H_N under length and momentum constraints a *ground state*. It is clear such a minimum is achieve when all of the particles have constant deformation ℓ and momentum p. In fact, if the potential V is non-negative for all $r > 0$ then this state is the unique ground state. If, however, the potential has finite range, then depending on ℓ there many minimizers corresponding to non-interacting configurations.

If the dynamics of (4.10) start on $\Sigma_{\ell,p,e}^{N}$, they will stay on $\Sigma_{\ell,p,e}^{N}$ due to the fact that $\mathsf{L}_N, \mathsf{P}_N, \mathsf{H}_N$ are conserved. Moreover, if the choice of V is generic enough, and there are no other conserved quantities, then one expects that the dynamics become uniformly mixed on $\Sigma_{\ell,p,e}^{N}$ after a long time and can be described by a uniform distribution on $\Sigma_{\ell,p,e}^{N}$ (often referred to as Boltzmann's ergodic hypothesis).

Naturally this motivates the study of the *micro canonical ensemble* with parameters $(\ell, p, e) \in D_S$ to be the measure

$$\mu_N^{\ell,p,e}(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}) \equiv \mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} \,|\, \ell, p, e)$$

obtained by conditioning the grand-canonical measure $\mu^N_{\tau,\lambda,\beta}(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v})$ with respect to the map

$$\widehat{w}^N(\mathbf{r}, \mathbf{v}) := \frac{1}{N}\left(\mathsf{L}_N(\mathbf{r}), \mathsf{P}_N(\mathbf{v}), \mathsf{H}_N(\mathbf{r}, \mathbf{v})\right). \tag{4.16}$$

Moreover, since the density of $\mu^N_{\tau,\lambda,\beta}(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v})$ depends explicitly on the quantities $\mathsf{L}_N, \mathsf{P}_N, \mathsf{H}_N$, then $\mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} \,|\, \ell, p, e)$ does not depend on (τ, λ, β) (c.f. Lemma A.3.4) and can be understood through the decomposition

$$\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} = \mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} \,|\, \ell, p, e)\, \widehat{\gamma}_N(\mathrm{d}\ell \mathrm{d}p \mathrm{d}e), \tag{4.17}$$

where $\hat{\gamma}_N$ is the pushforward

$$\hat{\gamma}_N(\mathrm{d}\ell \mathrm{d}p \mathrm{d}e) = \widehat{w}^N_{\#}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}.$$

Then for each $(\ell, p, e) \in D_S$, the measure $\mu^{\ell,p,e}$ concentrated on $\Sigma^N_{\ell,p,e}$ and since $\Sigma^N_{\ell,p,e} = \{\widehat{w}_N = (\ell, p, e)\}$ is a bounded subset of Ω^N, this measure is well-defined.

Using this decomposition, it is easy to see that $\mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} \,|\, \ell, p, e)$ is also an invariant measure for (4.10) and that, just as we had for the grand-canonical measure $\mu^N_{\tau,\lambda,\beta}$, we have the following anti-symmetry property

Lemma 4.2.5. *The operator $\mathcal{A}_N$ is skew-symmetric with respect to $\mu_N^{\ell,p,e}$, that is, for each $F, G \in C^1_b(\Omega^N)$ we have*

$$\int_{\Sigma^N_{\ell,p,e}} F\, \mathcal{A}_N G \,\mathrm{d}\mu_N^{\ell,p,e} = -\int_{\Sigma^N_{\ell,p,e}} \mathcal{A}_N F\, G \,\mathrm{d}\mu_N^{\ell,p,e}$$

Proof. Let $\varphi \in C_b^1(\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+)$, then the decomposition (4.17) implies that

$$\int_{\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+} \left(\varphi(\ell, p, e) \int_{\Sigma^N_{\ell,p,e}} F\, \mathcal{A}_N G \,\mathrm{d}\mu_N^{\ell,p,e} \right) \widehat{\gamma}_N(\mathrm{d}\ell \mathrm{d}p \mathrm{d}e) \\ = \int_{\Omega^N} \varphi(\widehat{w}^N)\, F\, \mathcal{A}_N G \,\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}.$$

Using the fact that $\mathcal{A}_{\mathsf{H}_N}\varphi(\widehat{w}^N) = 0$ and that $\mathcal{A}_{\mathsf{H}_N}$ is skew-symmetric with respect to Lebesgue measure, we obtain

$$\int_{\Omega^N} \varphi(\widehat{w}^N)\, F\, \mathcal{A}_N G \,\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} = -\int_{\Omega^N} \varphi(\widehat{w}^N)\, \mathcal{A}_N F\, G \,\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}.$$

This completes the proof. □

To continue we will need a further hypothesis on $V(r)$

Hypothesis 4.2.6. *Let $w(r, v) = (r, v, \frac{1}{2}v^2 + V(r))$, then the function*

$$\phi(\xi) = \int_{\mathbb{R}_+\times\mathbb{R}} \mathrm{e}^{(i\xi-\alpha)\cdot w(z)} \mathrm{d}z$$

belongs to $L^\nu(\mathbb{R}^3)$ for some $\nu \geq 1$ and satisfies the non lattice condition

$$|\phi(\xi)| < 1, \quad \textit{for} \quad |\xi| > 0.$$

Remark 4.2.7. Hypothesis 4.2.6 is equivalent to requiring that the push forward measure $w_\# \mathrm{e}^{-\alpha\cdot w(z)}\mathrm{d}z$, for $\alpha \in D_S$, satisfies the first condition of Hypothesis A.3.1, in fact, the second condition of Hypothesis A.3.1 is also satisfied, since $-w(z)$ has compact super-level sets. Of course, this hypothesis ensures that the ν^{th} convolution of the push-forward of $\mathrm{d}z$ under $w(z)$ has a density with respect to Lebesgue measure on $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$ even though $w_\#\mathrm{d}z$ is only supported on a sub-manifold of $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$.

It is an interesting question as to which class of singular potentials $V(r)$ satisfy Hypothesis 4.2.6.

Hypothesis 4.2.6, ensures that when N is large enough $N > N_0$, $\hat{\gamma}_N(\mathrm{d}\ell\mathrm{d}p\mathrm{d}e)$ has a density

$$\hat{\gamma}_N(\mathrm{d}\ell\mathrm{d}p\mathrm{d}e) = g_N(\ell,p,e)\,\mathrm{d}\ell\mathrm{d}p\mathrm{d}e.$$

The density $g_N(\ell,p,e)$ is the so-called *density of states*, and is formally written as,

$$g_N(\ell,p,e) = \int_{\mathbb{R}_+^N\times\mathbb{R}^N} \delta\big(N^{-1}\mathsf{L}_N(\mathbf{r})-\ell\big)\delta\big(N^{-1}\mathsf{P}_N(\mathbf{v})-p\big)\delta\big(N^{-1}\mathsf{H}_N(\mathbf{r},\mathbf{v})-e\big)\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}.$$

In this setting, the micro-canonical measure $\mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}\,|\,\ell,p,e)$ can also be written as

$$\begin{aligned}\mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}\,|\,\ell,p,e)&\\ &= \frac{1}{g_N(\ell,p,e)}\delta\big(N^{-1}\mathsf{L}_N(\mathbf{r})-\ell\big)\delta\big(N^{-1}\mathsf{P}_N(\mathbf{v})-p\big)\delta\big(N^{-1}\mathsf{H}_N(\mathbf{r},\mathbf{v})-e\big)\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v},\end{aligned}$$

We can give a more explicit representation of the function $g_N(\ell,p,e)$ on D_S.

Lemma 4.2.8. *Let $(\ell,p,e)\in D_S$, then we have the representation*

$$g_N(\ell,p,e) = N^{3/2}\int_{\Sigma^N_{\ell,p,e}} [G_N(\mathbf{r},\mathbf{v})]^{-1/2}\,\mathrm{d}\mathcal{H}^{2N-3}(\mathbf{r},\mathbf{v}),$$

where $\mathrm{d}\mathcal{H}^{2N-3}$ is the $2N-3$ dimensional Hausdorff measure and

$$G_N(\mathbf{r},\mathbf{v}) = \frac{1}{N}\sum_{i=1}^N\left(v_i - \frac{1}{N}\sum_{j=1}^N v_i\right)^2 + \frac{1}{N}\sum_{i=1}^N\left(V'(r_i)-\frac{1}{N}\sum_{i=1}^N V'(r_i)\right)^2.$$

Proof. We will use the co-area formula applied to the function $\widehat{w}^N(\mathbf{r},\mathbf{v})$, which states that for any function φ on $\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+$,

$$\begin{aligned}\int_{\mathbb{R}_+^N\times\mathbb{R}^N}&\varphi(\widehat{w}^N(\mathbf{r},\mathbf{v}))\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}\\ &= \int_{\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}_+}\varphi(\ell,p,e)\left(\int_{\Sigma^N_{\ell,p,e}}\left|\det\left(\partial\widehat{w}^N[\partial\widehat{w}^N]^\top\right)\right|^{-1/2}\mathrm{d}\mathcal{H}^{2N-3}(\mathbf{r},\mathbf{v})\right)\mathrm{d}\ell\mathrm{d}p\mathrm{d}e.\end{aligned}$$

By direct computation

$$\partial_{r_j}\widehat{w}_\ell^N = N^{-1}, \quad \partial_{r_j}\widehat{w}_p^N = 0, \quad \partial_{r_j}\widehat{w}_e^N = N^{-1}V'(r_i),$$

and

$$\partial_{v_j}\widehat{w}_\ell^N = 0, \quad \partial_{v_j}\widehat{w}_p^N = N^{-1}, \quad \partial_{v_j}\widehat{w}_e^N = N^{-1}v_j,$$

Therefore

$$\partial\widehat{w}^N[\partial\widehat{w}^N]^\top = N^{-2}\sum_{j=1}^N \begin{pmatrix} 1 & 0 & V'(r_i) \\ 0 & 1 & v_j \\ V'(r_j) & v_j & v_j^2 + (V'(r_j))^2. \end{pmatrix}$$

Taking the determinant yields

$$\begin{aligned} &\det\left(\partial\widehat{w}^N[\partial\widehat{w}^N]^\top\right) \\ &\quad = N^{-3}\left[\frac{1}{N}\sum_{j=1}^N v_j^2 + (V'(r_j))^2 - \left(\frac{1}{N}\sum_{j=1}^N v_j\right)^2 - \left(\frac{1}{N}\sum_{j=1}^N V'(r_j)\right)^2\right] \\ &\quad = N^{-3}G_N(\mathbf{r},\mathbf{v}). \end{aligned}$$

Therefore, using the definition of g_N, we obtain

$$\int_{D_S}\varphi\, g_N\, \mathrm{d}\ell\mathrm{d}p\mathrm{d}e = \int_{D_S}\varphi\, N^{3/2}\left(\int_{\Sigma_{\ell,p,e}^N}[G_N(\mathbf{r},\mathbf{v})]^{-1/2}\mathrm{d}\mathcal{H}^{2N-3}\right)\mathrm{d}\ell\mathrm{d}p\mathrm{d}e.$$

□

As a consequence of Galilean invariance, if we write $y = (\ell, p, e)$, we will see that g_N is just a function of ℓ and the internal energy $u = e - \frac{1}{2}p^2$. In fact we will say that a function $f(\ell, p, e)$ has the *Galilean shifty property* if it satisfies

$$f(\ell, p, e) = f(\ell, 0, e - \tfrac{1}{2}p^2).$$

Indeed we show that g_N has the Galilean shift property.

Lemma 4.2.9. *Let $y = (\ell, p, e) \in D_S$ and $N > \nu$, then $g_N(\ell, p, e)$ has the Galilean shift property, that is*

$$g_N(\ell, p, e) = g_N(\ell, 0, e - \tfrac{1}{2}p^2).$$

Proof. Formally this can be seen using delta function notation, and writing

$$g_N(\ell, p, e) = \int \delta\left(N^{-1}\mathsf{L}_N(\mathbf{r}) - \ell\right) \delta\left(N^{-1}\mathsf{P}_N(\mathbf{v}) - p\right) \delta\left(N^{-1}\mathsf{H}_N(\mathbf{r}, \mathbf{v}) - e\right) \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}.$$

Changing coordinates from $\mathbf{v} \to \mathbf{v} + p$ and using the fact that when $\mathsf{P}(\mathbf{v}) = 0$,

$$\mathsf{H}_N(\mathbf{r}, \mathbf{v} + p) = \mathsf{H}_N(\mathbf{r}, \mathbf{v}) + \frac{N}{2}p^2,$$

we find

$$g_N(\ell, p, e) = \int \delta\left(N^{-1}\mathsf{L}_N(\mathbf{r}) - \ell\right) \delta\left(N^{-1}\mathsf{P}_N(\mathbf{v})\right) \delta\left(N^{-1}\mathsf{H}_N(\mathbf{r}, \mathbf{v}) - (e - \tfrac{1}{2}p^2)\right) \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}.$$

However, this can be shown more rigorously using the representation for g_N given in Lemma 4.2.8 and noting that $\Sigma^N_{\ell,p,e}$ satisfies following property with respect to a shift in velocity

$$\Sigma^N_{\ell,p,e} - (0, p\mathbf{1}) = \Sigma^N_{\ell,0,e-\frac{1}{2}p^2}, \tag{4.18}$$

where $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^N$. The proof is complete upon changing variables from $\mathbf{v} \to \mathbf{v} + p$, using the shift property (4.18) and the translation invariance of Hausdorff measure. □

The Galilean shift property also arise with respect to micro-canonical averages. For any bounded continuous function G on $\mathbb{R}^N_+ \times \mathbb{R}^N$, denote the average with respect to $\mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} \,|\, \ell, p, e)$ by

$$\widehat{G}_N(\ell, p, e) = \int_{\mathbb{R}^N_+ \times \mathbb{R}^N} G(\mathbf{r}, \mathbf{v})\, \mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} \,|\, \ell, p, e).$$

Then we have the following Galilean shift property, analogous to Lemma 4.2.9.

Lemma 4.2.10. *Let G be a bounded continuous function on $\mathbb{R}_+^N \times \mathbb{R}^N$, that doesn't depend on velocity $\mathbf{v}$ (i.e. it is Galilean invariant) and let $(\ell, p, e) \in D_\Sigma$. Then $\widehat{G}_N(\ell, p, e)$ satisfies the Galilean shift property,*

$$\widehat{G}_N(\ell, p, e) = \widehat{G}_N(\ell, 0, e - \tfrac{1}{2}p^2).$$

Proof. The proof is the same as that of Lemma 4.2.9. □

As it turns out, the grand-canonical ensemble $\mu^N_{\tau,\lambda,\beta}(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v})$ is a good approximation of $\mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} \,|\, \ell, p, e)$ as $N \to \infty$, where (τ, λ, β) are related to (ℓ, p, e) through the entropy function $S(\ell, p, e)$, specifically for $(\ell, p, e) \in D_S$

$$\tau = \partial_\ell S(\ell, p, e), \quad \lambda = \partial_p S(\ell, p, e), \quad \beta = \partial_e S(\ell, p, e). \tag{4.19}$$

Indeed if one follows the formalism of Section A.2, then the result of Theorem A.3.7 (and Hypothesis 4.2.6) can be restated in to give the following *equivalence of ensembles* between the grand-canonical and micro-canonical ensembles.

Theorem 4.2.11. *For each $(\ell, p, e) \in D_S$ let (τ, λ, β) be given by (4.19). Then for each bounded continuous G on $\mathbb{R}_+^K \times \mathbb{R}^K$, for some K, the following limit holds*

$$\lim_{N\to\infty} \widehat{G}_N(\ell, p, e) = \int_{\mathbb{R}_+^K \times \mathbb{R}^K} G(r_1, \ldots r_K, v_1, \ldots v_K) \prod_{i=1}^K \mu_{\tau,\lambda,\beta}(\mathrm{d}r_i \mathrm{d}v_i).$$

Micro-canonical Thermodynamic Structure

We would now like to define a micro-canonical thermodynamic structure for finite, but large, N. Namely we would like to identify a pressure $P_N(\ell, p, e)$ an entropy $S_N(\ell,p,e)$, and an inverse temperature $\beta_N(\ell,p,e)$ which satisfy the Galilean shift property, are related by the first law of thermodynamics

$$\mathrm{d}S_N(\ell,0,e)=\beta_N(\ell,0,e)\mathrm{d}e+\beta_N P_N(\ell,0,e)\mathrm{d}\ell.$$

Moreover we would like each function P_N, S_N and β_N to converge as $N\to\infty$ to the corresponding thermodynamic functions P, S and β.

As recognized by Gibbs in [62], at the level of the micro-canonical ensemble, there are several notions of entropy, or so-called 'thermodynamic analogies', that give rise to the first law, each one with its own drawbacks. In our approach, we will find it desirable to have the pressure P_N in that arises the first law to be the micro-canonical averaged force

$$P_N(\ell,p,e)=-\int_{\Sigma^N_{\ell,p,e}}\left(\frac{1}{N}\sum_{i=1}^{N}V'(r_i)\right)\mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}\,|\,\ell,p,e). \tag{4.20}$$

Indeed as a consequence of the equivalence of ensembles (Theorem 4.2.11), we have

$$\lim_{N\to\infty}P_N(\ell,0,u)=P(\ell,u)$$

so that P_N and P agree for large N. In order to ensure that the first law is satisfied, then define the micro-canonical entropy S_N to be the so-called *volume entropy*,

$$S_N(\ell,p,e)=\frac{1}{N}\log\left(\int_{\frac{1}{2}p^2+V(\ell)}^{e}g_N(\ell,p,\mathrm{e}')\mathrm{d}\mathrm{e}'\right), \tag{4.21}$$

and then define the corresponding inverse temperature β_N by

$$\beta_N(\ell,p,e)=\partial_e S_N(\ell,p,e)=\frac{1}{N}g_N(\ell,p,e)\mathrm{e}^{-NS_N(\ell,p,e)}>0. \tag{4.22}$$

Remark 4.2.12. Of course, taking a hint from Boltzmann, one might expect that the entropy S_N to be given by the logarithm of the density of states $\log g_N$. Indeed from Lemma A.3.5, we know that

$$\lim_{N\to\infty}\frac{1}{N}\log g_N(y) = S(y),$$

on D_S, where $S(y)$ is the thermodynamic entropy, defined by (4.13). Therefore, it seems natural that the quantity

$$\bar{S}_N(y) = \frac{1}{N}\log g_N(y)$$

would make a good candidate for the entropy. This version of the entropy we will refer to as the *surface entropy.* However, as we will see, the forthcoming Lemma 4.2.13 implies that $\bar{S}(\ell, p, e)$ satisfies the relation

$$\partial_\ell \bar{S}_N(\ell, p, e) = \frac{1}{N}\partial_e P_N(\ell, p, e) + P_N(\ell, p, e)\partial_e \bar{S}_N(\ell, p, e),$$

and therefore $\bar{S}_N$ *does not* satisfy the first law with respect to P_N as defined in (4.20), and it therefore undesirable for our considerations. This discrepancy between the notion of 'volume entropy' (i.e. entropy of all states less than a certain energy) and 'surface entropy' (i.e. the entropy of all states with a certain prescribed energy) was introduced by Gibbs in [62] while studying the micro-canonical ensemble. One of the major downfalls of surface entropy as it's defined is that the pressure it gives rise to is a complicated quantity and not clearly related to the averaged pressure P_N defined above. Moreover, in certain circumstances, the inverse temperature that arises from the surface entropy can give rise to negative temperatures (see [38]), which is again undesirable.

We have the following relation between the coarse-grained pressure function P_N and the density of states g_N.

Lemma 4.2.13. *The following identity holds*

$$\partial_\ell g_N = \partial_e(P_N\, g_N)$$

Proof. Let $y = (\ell, p, e) \in D_S$ and let $\varphi(y)$ be a C^1 function on D_S which vanishes at ∞, then it suffices to show for all such φ,

$$\int_{D_S} \partial_\ell \varphi(y)\, g_N(y)\,\mathrm{d}y = \int_{D_S} \partial_e \varphi(y)\, P_N(y)\, g_N(y)\,\mathrm{d}y.$$

To this end, let $\widehat{w}^N$ be as in (4.16) then for each $1 \le i \le N$, we have the identity

$$N\partial_{r_i}\left[\varphi\left(\widehat{w}^N(\mathbf{r},\mathbf{v})\right)\right] = \partial_\ell\varphi\left(\widehat{w}^N(\mathbf{r},\mathbf{v})\right) + V'(r_i)\partial_e\varphi\left(\widehat{w}^N(\mathbf{r},\mathbf{v})\right).$$

Integrating both sides over Ω^N, we obtain

$$\int_{\Omega^N} \partial_\ell\varphi\left(\widehat{w}^N(\mathbf{r},\mathbf{v})\right)\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} = -\int_{\Omega^N} V'(r_i)\partial_e\varphi\left(\widehat{w}^N(\mathbf{r},\mathbf{v})\right)\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}. \tag{4.23}$$

Using the permutation symmetry of $\widehat{w}^N(\mathbf{r},\mathbf{v})$, and the definition of $\mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}\,|\,y)$, the right-hand side of (4.23) becomes

$$\begin{aligned}
-\int_{\mathbb{R}_+^N\times\mathbb{R}^N} & V'(r_i)\partial_e\varphi\left(\widehat{w}^N(\mathbf{r},\mathbf{v})\right)\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} \\
&= -\int_{D_S} \partial_e\varphi(y)\left(\int_{\Sigma_y^N}\left(\frac{1}{N}\sum_{i=1}^N V'(r_i)\right)\mu_N(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}\,|\,y)\right) g_N(y)\mathrm{d}y \\
&= \int_{D_S} \partial_e\varphi(y) P_N(y)\, g_N(y)\mathrm{d}y.
\end{aligned}$$

Similarly using the definition of $g_N(y)$, the left-hand side of (4.23) becomes

$$\int_{\mathbb{R}_+^N\times\mathbb{R}^N} \partial_\ell\varphi\left(\widehat{w}^N(\mathbf{r},\mathbf{v})\right)\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v} = \int_{D_S} \partial_\ell\varphi(y)\, g_N(y)\,\mathrm{d}y.$$

□

We will need the following limits as the energy e approaches the ground state.

Lemma 4.2.14. *Suppose, in addition to Hypothesis 4.2.1, the potential V is strictly convex. Then following limit holds*

$$\lim_{e \to V(\ell)} P_N(\ell, 0, e) = -V'(\ell).$$

Proof. First, note that since by assumption, $-V'$ is convex, by Jensen's inequality, we have the lower bound

$$-V'(\ell) \leq P_N(\ell, 0, e).$$

Furthermore, under the constraint $\sum_{i=1}^N r_i = N\ell$, by Taylor's theorem and the fact the $V''(r)$ is decreasing

$$\sum_{i=1}^N (V(r_i) - V(\ell)) \geq \sum_{i=1}^N V''(\max\{r_i, \ell\})(r_i - \ell)^2 \geq V''(N\ell) \sum_{i=1}^N (r_i - \ell)^2.$$

Since $V(r)$ is strictly convex, we define $V''(N\ell) = C > 0$. It follows that if $0 < e - V(\ell) < \delta$, then on the manifold $\Sigma^N_{\ell,0,e}$, we have

$$\sum_{i=1}^N (r_i - \ell)^2 \leq NC^{-1}\delta.$$

Indeed, this implies that for each $\epsilon > 0$, we may choose δ small enough so that on $\Sigma^N_{\ell,0,e}$, we have $|r_i - \ell| < \epsilon$. This implies, by the fact that $-V'(r)$ is decreasing, that on $\Sigma^N_{\ell,0,e}$ and for small enough $\epsilon > 0$,

$$-V'(r_i) \leq -V'(\ell - \epsilon).$$

Therefore, when $e - V(ell)$ is small enough, we have the bound

$$-V'(\ell) \leq P_N(\ell, 0, e) \leq -V'(\ell - \epsilon).$$

Sending $\epsilon \to 0$ gives the proof. □

Lemma 4.2.15. *The following limit holds*

$$\lim_{e\to V(\ell)} g_N(\ell,0,e) = 0.$$

First, we observe that, with these definitions, the first law is satisfied.

Theorem 4.2.16. *Let P_N, S_N and β_N be defined through equations (4.20),(4.21) and (4.22) respectively. Then they satisfy the first law, i.e.*

$$\partial_e S_N = \beta_N, \quad \partial_\ell S_N = \beta_N P_N.$$

Moreover, P_N, S_N and β_N satisfy the Galilean shift property, and the following limits hold for $(\ell,p,e)\in D_S$

$$\lim_{N\to\infty} S_N(\ell,p,e) = S(\ell,p,e), \quad \lim_{N\to\infty} P_N(\ell,p,e) = P(\ell,p,e).$$

Proof. First we prove the Galilean shift property. This follows for P_N from Lemma 4.2.10. For S_N it follows from the fact that g_N has the property and a change of variables,

$$\int_{\frac{1}{2}p^2+V(\ell)}^{e} g(\ell,p,e')\mathrm{d}e' = \int_{\frac{1}{2}p^2+V(\ell)}^{e} g(\ell,0,e'-\tfrac{1}{2}p^2)\mathrm{d}e' = \int_{V(\ell)}^{e-\frac{1}{2}p^2} g(\ell,0,e')\mathrm{d}e'$$

Finally Galilean shift property for β_N follows from the fact that S_N has it.

To verify the first law, note that $\partial_e S_N = \beta_N$ is satisfied by definition, therefore we simply need to check that $\partial_\ell S_N = \beta_N P_N$. Moreover, using the Galilean shift property it suffices to check for $p = 0$. Using Lemma 4.2.13 and the fact that

$\lim_{e\to V(\ell)} P_N(\ell,0,e) = -V'(\ell)$ gives

$$\begin{aligned}
\partial_\ell S_N(\ell,0,e) &= \frac{1}{N}\mathrm{e}^{-NS_N(\ell,0,e)}\left(\int_{V(\ell)}^{e} \partial_\ell g_N(\ell,0,e')\mathrm{d}e' - V'(\ell)\lim_{e\to V(\ell)} g_N(\ell,0,e)\right)\\
&= \frac{1}{N}\mathrm{e}^{-NS_N(\ell,0,e)}\left(\int_{0}^{e} \partial_e(P_N(\ell,0,e')g_N(\ell,0,e'))\mathrm{d}e' - V'(\ell)\lim_{e\to V(\ell)} g_N(\ell,0,e)\right)\\
&= \frac{1}{N}\mathrm{e}^{-NS_N(\ell,0,e)}\Bigg(P_N(\ell,0,e)g_N(\ell,0,e) - \lim_{e\to V(\ell)} P_N(\ell,0,e)g_N(\ell,0,e)\\
&\qquad\qquad - V'(\ell)\lim_{e\to V(\ell)} g_N(\ell,0,e)\Bigg)\\
&= \beta_N(\ell,0,e)P_N(\ell,0,e)
\end{aligned}$$

Next, we show the limits of S_N and P_N as $N\to\infty$. Note the limit for P_N already follows from the equivalence of ensembles (Theorem A.3.7). While for S_N, we will need the following locally uniform asymptotic

$$\hat{g}_N(\ell,p,e) = \frac{\mathrm{e}^{NS(\ell,p,e)}}{(2\pi)^{3/2}}\sqrt{N^3\det\left(-\nabla^2 S(\ell,p,e)\right)}\left(1+\mathcal{O}(N^{-1/2})\right).$$

Then a straight forward application of Laplace's method yields

$$\lim_{n\to\infty}\frac{1}{N}\log\left(\int_{\frac{1}{2}p^2+V(\ell)}^{e} g_N(\ell,p,\mathrm{e}')\mathrm{d}e'\right) = \sup_{e'\in(\frac{1}{2}p^2+V(\ell),e)} S(\ell,p,e') = S(\ell,p,e).$$

where in the last equality, we used the fact that $e\mapsto S(\ell,p,e)$ is an increasing function. □

Stochastic Regularizations

If V is sufficiently nonlinear, one expects that for long times and large enough N the dynamics (4.10) becomes suitably mixed on the micro-canonical manifold $\Sigma^N_{\ell,p,e}$. While this is a natural conjecture, establishing this is an incredibly difficult mathematical problem. Indeed, to obtain such mixing, one must ensure that the obvious conserved quantities of length, momentum and energy are the *only* conserved quantities, so that the dynamics are not constrained to any proper submanifolds of $\Sigma^N_{\ell,p,e}$. As is well known (see [21, 96]) the potential $V(r) = r^{-2}$ (which satisfies Hypothesis 4.2.1) leads to an integrable dynamical system in one dimension, and therefore has more conserved quantities than just length, momentum, and energy. Another example is the hard rod fluid, where, due to the fact that collisions swap velocities of the colliding particles, one can verify, for instance, that the number of particles with a particular velocity is a conserved quantity. Of course, if one removes the restriction that the potential has a singularity at zero then there are many examples of potentials that lead to integrable systems, the harmonic potential $V(r) = r^2$, and the Toda potential $V(r) = e^{-r}$ ([112]), are just a few.

Indeed, this problem appears to be well out of the reach current mathematical techniques. However, a common to technique to circumvent such mathematical difficulties is to introduce a stochastic perturbation to the dynamics which conserves the quantities of interest, namely the length, momentum, and energy, while introducing the necessary mixing to obtain ergodicity. Typically, these perturbations are chosen to mimic certain random collisions between neighboring particles and are introduced to simulate, in some qualitative sense, the ergodicity and mixing that one expects from the deterministic Hamiltonian dynamics. This approach, for instance, was taken in by Olla, Varadhan, Yau [101] in their pioneering paper on the hydrodynamic limit of a classical Hamiltonian system of particles in three dimensions, where is was used to obtain a local ergodic theorem (see also [59, 90]), which is a necessary step in the proof of the hydrodynamic limit.

In the following subsections we will discuss several stochastic perturbations to the deterministic dynamics (4.10) which conserve energy and momentum and act locally on the momentum variables only. Such perturbation are regularly considered in the literature on stochastic lattice systems, particularly that of Harmonic chains.

Poisson type noise

One of simplest strategies for an energy/ momentum conserving noises is one that preserves pairwise momentum $v_i + v_{i+1}$ and energy $\frac{1}{2}v_i^2 + \frac{1}{2}v_{i+1}^2$ for each $i \in \mathbb{Z}_N$.

Unfortunately, in one dimension, for a given pair of velocities (v_i, v_{i+1}) the only other pair that share the same momentum and kinetic energy is exchanged pair (v_{i+1}, v_i). As a consequence, it is not possible to construct a diffusion type noise on the manifold of pairwise momentum and energy conserving interactions, as the manifold consists of two disconnected points. Instead, we can construct a Poisson type noise that randomly swaps the momentum of adjacent particles. That is, particles i and $i+1$ exchange their velocities v_i and v_{i+1} at independent random exponentially distributed times with rate 1. This type of process can be equivalently described by a family of independent standard Poisson process $\{N_{i,i+1}(t)\}_{i\in\mathbb{Z}_N}$ with rate 1, whereby the evolution equations (4.10) become the following family of stochastic differential equations

$$
\begin{aligned}
\dot{r}_i &= v_i - v_{i-1} \\
\dot{v}_i &= V'(r_{i+1}) - V'(r_i) + (v_{i+1}^- - v_i^-)\dot{N}_{i,i+1} - (v_i^- - v_{i-1}^-)\dot{N}_{i-1,i},
\end{aligned}
\tag{4.24}
$$

where $v_i^-(t) = v_i(t-)$ denotes the velocity of the ith just before time t (its left limit at time t) and $\dot{N}_{i,i+1}$ can be represented as a train of random impulses

$$\dot{N}_{i,i+1}(t) = \sum_j \delta(t - T^j_{i,i+1}),$$

where $\{T^j_{i,i+1}\}_{j=1}^{\infty}$ is a Poisson distributed collection of random times when particle i and $i+1$ exchange velocities.

The generator $\mathcal{S}_N$ of the stochastic part of the above dynamics can be written as

$$\mathcal{S}_N = \sum_{i \in \mathbb{Z}_N} \mathcal{T}_i,$$

where $\mathcal{T}_i$ are so-called *exchange operators* $\{\mathcal{T}_i\}_{i\in\mathbb{Z}_N}$ acting on functions $\phi : \Omega^N \to \mathbb{R}$ and defined by

$$\mathcal{T}_i \phi(\mathbf{r}, \mathbf{v}) \equiv \mathcal{T}_{i-1,i}\phi(\mathbf{r}, \mathbf{v}) = \phi(\mathbf{r}, \mathbf{v}^{i-1,i}) - \phi(\mathbf{r}, \mathbf{v}),$$

where $\mathbf{v}^{i-1,i}$ denotes the velocities $\mathbf{v}$ with the velocity of the $i-1$th and ith particle swapped.

It is easy to obtain the following symmetry properties of the operator $\mathcal{S}_N$.

Lemma 4.3.1. *Let $F, G \in C_b(\Omega^N)$, and let $\nu(\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v})$ be a measure on Ω^N which is exchangeable in velocity, meaning that the measure is invariant under exchanges in the index of the velocities of neighboring particles. Then we have*

$$\int_{\Omega^N} F \mathcal{S}_N G \,\mathrm{d}\nu = \int_{\Omega^N} \mathcal{S}_N F \, G \,\mathrm{d}\nu.$$

Note that $\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{v}$, $\nu^N_{\tau,\lambda,\beta}$ and $\nu^{\ell,p,e}_N$ are all measure that are exchangeable in velocity, and therefore $\mathcal{S}_N$ is symmetric with respect to each of these measures. Since

$\mathcal{S}_N$ vanishes on constants (it is the generator of a Markov process), this clearly implies that any measure which is exchangeable with respect to velocity is an invariant measure for $\mathcal{S}_N$.

The generator of the full process (4.24) is now given by

$$\mathcal{L}_N = \mathcal{A}_N + \mathcal{S}_N,$$

and instead of the Liouville equation, the distribution of particles f^N is given by the forward Kolmogorov equation

$$\partial_t f^N + \mathcal{A}_N f^N - \mathcal{S}_N f^N = 0.$$

Diffusion Type Noise

If one allows for interactions between more than two consecutive particles, one can consider noises which are of diffusion type. This has been done, for instance, in [10], while studying the divergence of thermal conductivity in a momentum conserving anharmonic chain.

For any three indices $(i-1, i, i+1)$, the set of velocities (v_{i-1}, v_i, v_{i+1}) satisfying $v_{i-1} + v_i + v_{i+1} = c_1$ and $v_{i-1}^2 + v_i^2 + v_{i+1}^2 = c_2$ is a one dimensional manifold. It is not hard to see that this set is a just the intersection of a 2-dimensional sphere and a plane, and therefore is just a circle S_{c_1,c_2} embedded in $\mathbb{R}^3$. We aim to construct a Brownian motion on this circle. To do this, note that the following vector field

$$\mathcal{Y}_i = (v_i - v_{i+1})\partial_{v_{i-1}} + (v_{i+1} - v_{i-1})\partial_{v_i} + (v_{i-1} - v_i)\partial_{v_{i+1}}$$

is tangent the manifold of three particle energy and momentum conserving interactions since

$$\mathcal{Y}_i(v_{i-1}+v_i+v_{i+1}) = \mathcal{Y}_i(v_{i-1}^2+v_i^2+v_{i+1}^2) = 0.$$

Therefore the operator $\mathcal{Y}_i^2$ is proportional to the Laplace Beltrami operator on S_{c_1,c_2}. If one now takes into account *all* such consecutive three particle interactions, we can construct a generator $\mathcal{G}_N$ for a momentum and energy conserving diffusion by

$$\mathcal{G}_N = \sum_{i\in\mathbb{Z}_N} \mathcal{Y}_i^2.$$

Because of the conservation properties of $\mathcal{G}_N$ (i.e. $\mathcal{G}_N \mathsf{H}_N = \mathcal{G}_N \mathsf{P}_N = 0$), it is easy to see that $\mathcal{G}_N$ is symmetric with respect to $d\mathbf{r}d\mathbf{v}$, $\nu^N_{\tau,\lambda,\beta}$ and $\nu^{\ell,p,e}_N$.

If one adds this diffusive stochastic dynamics to the deterministic Hamiltonian dynamics (4.10), we obtain a diffusion process with generator

$$\mathcal{A}_N + \mathcal{G}_N.$$

The evolution equations for the stochastically perturbed system now become the follows system of Itô stochastic differential equations,

$$\begin{aligned}
\dot{r}_i &= v_i - v_{i-1} \\
\dot{v}_i &= V'(r_{i+1}) - V'(r_i) - (v_{i-1} - v_{i-2}) \circ \dot{W}_{i-1} \\
&\qquad + (v_{i+1} - v_{i-1}) \circ \dot{W}_i - (v_{i+2} - v_{i+1}) \circ \dot{W}_{i+1},
\end{aligned}$$

where $\{W_i\}_{i\in\mathbb{Z}_N}$ are a family of independent one dimensional Wiener processes, and $\circ$ indicates the Stratonovich product.

Note that this type of noise has the effect of adding more than just nearest neighbor interactions to the system.

General Conservative Coarse-Graining on $\mathbb{Z}_N$

In this section, we will discuss several procedures for coarse-graining the particle system of Section 4.2.

Let $\Omega^N = (\mathbb{R}_+ \times \mathbb{R})^{\mathbb{Z}_N}$, then for any collection of particles $\mathbf{z}_N = \{z_i\} \in \Omega^N$, $z_i = (r_i, v_i)$, we denote the corresponding collection of locally conserved quantities by

$$\mathbf{w}_N = \{w(z_i)\}_{i\in\mathbb{Z}_N} \in \Gamma^N, \quad w(z_i) = \left(r_i, p_i, \tfrac{1}{2}v_i^2 + V(r_i)\right),$$

where $\Gamma^N = \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+^{\mathbb{Z}_N}$. Let, w_i^δ denote the δth local conserved quantity of the ith particle, with $\delta = 0$ corresponding to length, $\delta = 1$ corresponding to velocity, and $\delta = 2$ corresponding to energy. To be clear, we have defined

$$w_i^0 = r_i, \quad w_i^1 = v_i, \quad w_i^2 = \tfrac{1}{2}v_i^2 + V(r_i).$$

Recall the Liouville operator associated to (4.10)

$$\mathcal{A}_N = -\sum_{i\in\mathbb{Z}_N} \left(v_{i-1}(\partial_{r_i} - \partial_{r_{i-1}}) + V'(r_i)(\partial_{v_i} - \partial_{v_{i-1}})\right).$$

Each collection of locally conserved quantities $\{w_i^\delta\}_{i\in\mathbb{Z}_N}$ has a corresponding collection of local currents $\{\mathrm{J}_i^\delta\}_{i\in\mathbb{Z}_N}$ which satisfy

$$\mathcal{A}_N w_i^\delta = \mathrm{J}_{i+1}^\delta - \mathrm{J}_i^\delta,$$

and are given explicitly by

$$\mathrm{J}_i^0 = v_{i-1}, \quad \mathrm{J}_i^1 = V'(r_i), \quad \mathrm{J}_i^2 = v_{i-1}V'(r_i).$$

It is useful to remark that J_i^0 is itself another locally conserved quantity, while J_i^1 and J_i^2 are typically not conserved.

The procedure of conservative coarse-graining consists of taking a configuration $\mathbf{z}_N \in \Omega^N$ and associating to it a lower dimensional quantity whose elements represent *local averages* of the locally conserved quantities $\mathbf{w}_N$. To describe such a coarse-graining procedure, we introduce the empirical measure $\eta(\mathbf{z}_N)$ which is defined by its action on any smooth test function $\varphi : \mathbb{T} \to \mathbb{R}$ by

$$\eta(\mathbf{z}_N)[\varphi] = \frac{1}{N} \sum_{i \in \mathbb{Z}_N} w(z_i)\, \varphi(i/N).$$

The empirical measure η defines a mapping from Ω^N to $\mathcal{M}(\mathbb{T};\Gamma)$, where $\mathcal{M}(\mathbb{T};\Gamma)$ is the space of finite Γ valued measures on $\mathbb{T}$. From the empirical measure, one can always recover a configuration $\mathbf{z}_N$ that produces it, and this configuration will be unique up to permutations of the indices. Given any set $A \subseteq \mathbb{T}$, the empirical measure $\eta(A)$ computes the sample average of the locally conserved quantities w_i with $i/N \in A$.

The empirical measure gives information about the hydrodynamic behavior of a system. Indeed, if one lets $\mathbf{z}_N(t)$ be a solution of (4.10) then one expects $\eta(\mathbf{z}_N(Nt))$ to be close to a solution of the Euler equations in Lagrangian form. This can be proved rigorously when stochastic collisions are added and is proved in [14] for the case of anharmonic chains.

We can use the empirical measure to construct a coarse-graining map in the following way. Begin by choosing a sampling function $\varphi : \mathbb{T} \to \mathbb{R}_+$, which is typically a function centered around zero and symmetric, with support on a proper subset of $\mathbb{T}$. From this sampling function, we may construct a partition of unity $\{\varphi^i\}_{i=1}^M$ on the Torus by

$$\varphi^i(x) = \frac{\varphi(x - i/M)}{\sum_{j=1}^{M} \varphi(x - j/M)}.$$

The functions $\{\varphi^j\}_{j=1}^M$ induce a collection $\{\eta(\mathbf{z}_N)[\varphi^j]\}_{j=1}^M$ of weighted averages which are also locally conserved in the sense that

$$\sum_{j=1}^{M} \eta(\mathbf{z}_N)[\varphi^j] = \sum_{i=1}^{N} w(z_i).$$

These averages induce a map

$$\Omega^N \ni \mathbf{z}_N \to \{\eta(\mathbf{z}_N)[\varphi^j]\}_{j=1}^M \in \Gamma^M,$$

which serves to coarse-grain the configuration $\mathbf{z}_N$ by associating groups of nearby particles with their average of length, momentum, and energy. Note that in this general framework, when a particle is summed with weight less than one, it is automatically shared with another average.

Coarse-graining by lumping in $\mathbb{Z}_N$

Our first case of a conservative coarse-graining map is what is often referred to in the theory of discrete Markov processes as "lumping" (see [76]). In this setting we will choose the sampling function ϕ as an indicator function on an interval $I = [-1/M, 1/M)$ where M is a natural number that evenly divides N, so that $N/M = K$ for some natural number K. Then the partition of unity is just

$$\phi^j = \mathbb{1}_{I+j/M},$$

and the support of each ϕ^j does no overlap the support of any other ϕ^i. Such a "hard" sampling induces a partition of the periodic lattice $\mathbb{Z}_N$ into cells $\{\Lambda_i\}_{i\in\mathbb{Z}_M}$,

$\bigcup_{i\in\mathbb{Z}_M} \Lambda_i = \mathbb{Z}_N$, defined by

$$\Lambda_i = \{j \in \mathbb{Z}_N \,:\, j/N - i/M \in I\},$$

where a subset of of $\mathbb{Z}_N$ is called a *cell* if it is *proper and connected.* It is easy to see that each cell Λ_i contains exactly K elements.

For each Λ_i, and a given configuration $\mathbf{z}_N$, we denote by $\mathbf{z}_{\Lambda_i}$ the collection of particles with indices in Λ_i. The empirical measure then introduces the following averages

$$\widehat{w}_i \equiv \widehat{w}(\mathbf{z}_{\Lambda_i}) = M\eta(\mathbf{z}_N)[\phi^j] = \frac{1}{K}\sum_{j\in\Lambda_i} w(z_i),$$

with the collection of all such averages denoted by

$$\widehat{\mathbf{w}}(\mathbf{z}_N) = \{\widehat{w}(\mathbf{z}_{\Lambda_i})\}_{i\in\mathbb{Z}_M} \in \Gamma^{\mathbb{Z}_M}.$$

The function $\mathbf{z}_N \mapsto \widehat{\mathbf{w}}(\mathbf{z}_N)$ defines a coarse-graining map from $\Omega^{\mathbb{Z}_N}$ to $\Gamma^{\mathbb{Z}_M}$. We will denote each component of $\widehat{\mathbf{w}}$ by

$$\widehat{w}_i = (\hat{\ell}_i, \hat{p}_i, \hat{e}_i) \equiv (\widehat{w}_i^0, \widehat{w}_i^1, \widehat{w}_i^2).$$

Clearly $\hat{\ell}_i, \hat{p}_i, \hat{e}_i$ are to be interpreted as the average length, momentum, and energy of the particles in the ith cell, and are given explicitly

$$\hat{\ell}_i = \frac{1}{K}\sum_{j\in\Lambda_i} r_j, \quad \hat{p}_i = \frac{1}{K}\sum_{j\in\Lambda_i} v_j, \quad \hat{e}_i = \frac{1}{K}\sum_{j\in\Lambda_i}\left(\frac{1}{2}v_j^2 + V(r_i)\right).$$

Note that for any $\mathbf{z}_N \in \Sigma^N_{\ell,p,e}$, we have

$$\frac{1}{M}\sum_{i\in\mathbb{Z}_M} \widehat{w}(\mathbf{z}_{\Lambda_i}) = \frac{1}{N}\sum_{i\in\mathbb{Z}_N} w(z_i) = (\ell, p, e),$$

so that $\{\widehat{w}_i\}_{i\in\mathbb{Z}_N}$ are also locally conserved variables for the mesoscopic system.

Decomposition into periodized operators

For any cell Λ, we will find it useful to define the *boundary elements* l^+ and l^- to be the unique elements of Λ such that

$$l^- \notin \Lambda + 1, \quad l^+ \notin \Lambda - 1.$$

Intuitively l^- is thought of as the least element of Λ, while l^+ is thought of as the largest element of Λ. Let l_i^+ and l_i^- be the boundary elements of the cell Λ_i.

We then define the *periodization* $\mathring{\Lambda}$ of a cell Λ to be the set with the elements in Λ with l^- and $l^+ + 1$ identified, so that $\mathring{\Lambda}$ is a periodic lattice with period $|\Lambda|$. Naturally, we will use the set $\mathring{\Lambda}$ to define a *periodized Liouville operator* on Λ, given by

$$\mathring{\mathcal{A}}_\Lambda = -\sum_{i \in \mathring{\Lambda}} \left(v_{i-1}(\partial_{r_i} - \partial_{r_{i-1}}) + V'(r_i)(\partial_{v_i} - \partial_{v_{i-1}}) \right). \tag{4.25}$$

Note that $\mathring{\mathcal{A}}_\Lambda$ *not* just the restriction of $\mathcal{A}_N$ the cell Λ as it ignores all interaction between neighboring cells and particles on either side of the boundary of Λ interact. In fact, if ψ_Λ is a function on Ω^N that depends only particles with indices in Λ, then we have following useful relation between $\mathcal{A}_N$ and $\mathring{\mathcal{A}}_\Lambda$

$$\mathcal{A}_N \psi_\Lambda = +\mathring{\mathcal{A}}_\Lambda \psi_\Lambda + \overline{\mathcal{A}}_\Lambda \psi_\Lambda \tag{4.26}$$

where $\overline{\mathcal{A}}_\Lambda$ is the boundary interaction operator

$$\overline{\mathcal{A}}_\Lambda = (v_{l^+} - v_{l^- -1})\partial_{r_{l^-}} + (V'(r_{l^+ +1}) - V'(r_{l^-}))\partial_{v_{l^+}}. \tag{4.27}$$

Indeed, the relation (4.26) induces a decomposition

$$\mathcal{A}_N = \mathring{\mathcal{A}}_M + \overline{\mathcal{A}}_M, \quad \mathring{\mathcal{A}}_M = \sum_{i \in \mathbb{Z}_M} \mathring{\mathcal{A}}_{\Lambda_i}, \quad \overline{\mathcal{A}}_M = \sum_{i \in \mathbb{Z}_M} \overline{\mathcal{A}}_{\Lambda_i}. \tag{4.28}$$

By the fact that we have

$$\mathring{\mathcal{A}}_M \widehat{w}_i^\delta = \mathring{\mathcal{A}}_{\Lambda_i} \widehat{w}_i^\delta = 0,$$

and therefore each $\widehat{w}_i^\delta$ satisfies

$$\mathcal{A}_N \widehat{w}_i^\delta = \overline{\mathcal{A}}_M \widehat{w}_i^\delta = \mathrm{J}^\delta_{l_{i+1}^-} - \mathrm{J}^\delta_{l_i^-}. \tag{4.29}$$

Of course, this also implies that $\widehat{\mathbf{w}}$ is locally conserved, and has local currents $\{\mathrm{J}^\delta_{l_i^-}\}$ that live on the boundary elements of each cell.

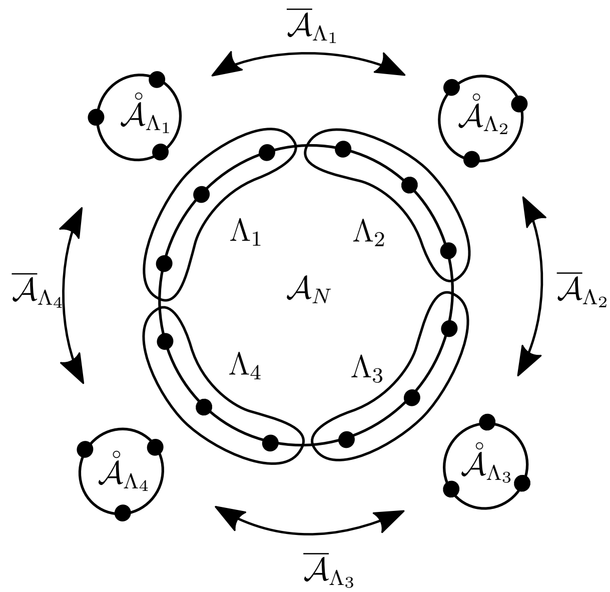

Figure 4.2: Diagram of the coarse-graining by lumping in the case that $N = 12$ and $K = 3$. The partition, the periodized operators, and the boundary interaction operators are shown on the cells on which they act.

Coarse-graining in equilibrium

The primary goal of coarse-graining is to obtain effective equations for evolution of the coarse-grained quantity $\widehat{\mathbf{w}}$. Our first step will be to coarse-grain our particle system when it is in equilibrium, namely when the system (4.10) is started with random initial data starting from a grand-canonical ensemble $\mu_\alpha^N(\mathrm{d}\mathbf{z}_N)$, where $\alpha = (\tau, \lambda, \beta) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+$. Indeed if f_0^N is distributed according to $\mu_\alpha^N(\mathrm{d}\mathbf{z}_N)$, then since the grand-canonical measure is invariant with respect to $\mathcal{A}_N$, the solution to the Liouville equation (4.11) is just $f_t^N = \mu_\alpha^N(\mathrm{d}\mathbf{z}_N)$.

Therefore, we would like to study the distribution $\hat{\mu}_{K,\alpha}^M(\mathrm{d}\mathbf{y}_M)$ of the coarse conserved variables $\widehat{\mathbf{w}}(\mathbf{z}_N)$ under the grand-canonical ensemble $\mu_\alpha^N(\mathrm{d}\mathbf{z}_N)$, where $\mathbf{y}_M = (y_1, \ldots, y_M)$ denotes an element of coarse-grained phase space $\Gamma^{\mathbb{Z}_M}$. Of course $\hat{\mu}_{K,\alpha}^M(\mathrm{d}\mathbf{y}_M)$ is just given by push forward

$$\hat{\mu}_{K,\alpha}^M(\mathrm{d}\mathbf{y}_M) = \widehat{\mathbf{w}}_{\#}\mu_\alpha^N(\mathrm{d}\mathbf{y}_M) = \prod_{i \in \mathbb{Z}_M} \mathrm{e}^{-K\alpha \cdot y_i - KF(\alpha)} \hat{\gamma}_K(\mathrm{d}y_i),$$

where

$$\hat{\gamma}_K(\mathrm{d}y) = \widehat{w}_{\#}\mathrm{d}\mathbf{z}_K.$$

Appealing to Hypothesis 4.2.6, when K is large enough $K > \nu$, $\hat{\gamma}_K(\mathrm{d}y)$ has a density

$$\hat{\gamma}_K(\mathrm{d}y) = g_K(y)\,\mathrm{d}y,$$

where $g_K(y)$ just the density of states associated to cell Λ_i. The mapping $\widehat{\mathbf{w}}$ also defines a conditional measure

$$\mu_K^M(\mathrm{d}\mathbf{z}_N \,|\, \mathbf{y}_M) = \prod_{i \in \mathbb{Z}_M} \mu_K(\mathrm{d}\mathbf{z}_{\Lambda_i} \,|\, y_i),$$

which is given by conditioning μ_α^N with respect to the event $\{\mathbf{y}_M = \widehat{\mathbf{w}}\}$. In each cell, $\mu_K(\mathrm{d}\mathbf{z}_{\Lambda_i} \,|\, y_i)$ is just the micro-canonical measure concentrated on the set $\Sigma_{y_i}^K = \{\mathbf{z} \in \Omega^{\Lambda_i} \,:\, \widehat{w}_i(\mathbf{z}) = y_i\}$.

Just as we did in Section 4.2.4, we may define the micro-canonical entropy S_K, pressure P_K and inverse temperature β_K by

$$S_K(\ell,p,e) := \frac{1}{K}\log\left(\int_{\frac{1}{2}p^2+V(\ell)} g_K(\ell,p,e')\mathrm{d}e'\right),$$

$$P_K(\ell,p,e) := \int_{\Sigma^K_{\ell,p,e}}\left(-\frac{1}{K}\sum_{i=1}^{K}V'(r_i)\right)\mu_K(\mathrm{d}\mathbf{z}_K\,|\,\ell,p,e),$$

and

$$\beta_K(\ell,p,e) = \partial_e S_K(\ell,p,e) = g_K(\ell,p,e)e^{-KS_K(\ell,p,e)}.$$

By Theorem 4.2.16, we have that S_K, P_K, and β_K satisfy the first law

$$\mathrm{d}S_K(\ell,0,e) = \beta_K P_K(\ell,0,e)\mathrm{d}\ell + \beta_K(\ell,0,e)\mathrm{d}e.$$

It follows that if $(\ell_i,p_i,e_i) = \widehat{w}_i$ are the coarse variables in the cell Λ_i, then $S_K(\ell_i,p_i,e_i)$, $P_K(\ell_i,p_i,e_i)$ and $\beta_K(\ell_i,p_i,e_i)$ denote the entropy, pressure, and inverse temperature of that cell. Furthermore, we may write $\hat{\mu}^M_{K,\alpha}$ in terms of these thermodynamic quantities by

$$\hat{\mu}^M_{K,\alpha}(\mathrm{d}\mathbf{y}_M) = \prod_{i\in\mathbb{Z}_M}\frac{1}{\widehat{Z}_K(\alpha)}\mathrm{e}^{-K\alpha\cdot y_i + KS_K(y_i)}\beta_K(y_i)\mathrm{d}y_i, = \prod_{i\in\mathbb{Z}_M}\hat{\mu}_{K,\alpha}(\mathrm{d}y_i)$$

where $\widehat{Z}_K(\alpha)$ is the normalizing constant for the measure $\mathrm{e}^{-K\alpha\cdot y+KS_K(y)}\beta_K(y)\mathrm{d}y$.

We will denote the averages with respect to $\mu^N(\mathrm{d}\mathbf{z}_N)$ and $\hat{\mu}^M_{K,\alpha}(\mathrm{d}\mathbf{y}_M)$ by $\langle\,\cdot\,\rangle^N_\alpha$, and $\langle\,\cdot\,\rangle^M_{K,\alpha}$ respectively. In addition, we will denote the averages with respect to the conditional measures $\mu_N(\mathrm{d}\mathbf{z}_N\,|\,y)$ and $\mu^M_K(\mathrm{d}\mathbf{z}_N\,|\,\mathbf{y}_M)$ by $\langle\,\cdot\,|\,y\,\rangle_N$ and $\langle\,\cdot\,|\,\mathbf{y}_M\rangle_N$ respectively.

In the equilibrium setting, the dynamics for $\widehat{\mathbf{w}}(\mathbf{z}_N)$ is statistically equivalent (in the sense of equality of time marginals), to an *exact* closed dynamics on the coarse space Γ^M. Specifically, define the coarse-grained Liouville operator by,

$$\widehat{\mathcal{A}}_M \phi(\mathbf{y}_M) := K \left\langle \mathcal{A}_N(\phi \circ \widehat{\mathbf{w}}) \,|\, \mathbf{y}_M \right\rangle_N .$$

Then $\hat{\mu}^M_{K,\alpha}$ is invariant with respect to $\widehat{\mathcal{A}}_M$, since for any $F \in C^1_b(\Gamma^M)$ we have the simple identity,

$$\big\langle \widehat{\mathcal{A}}_M F \big\rangle^M_{K,\alpha} = K \big\langle \mathcal{A}_N (F \circ \widehat{\mathbf{w}}) \big\rangle^N_\alpha = 0.$$

Indeed we also have the following anti-symmetry property for $\widehat{\mathcal{A}}_M$,

Lemma 4.5.1. *Let $F, G \in C^1_b(\Gamma^M)$, then $\widehat{\mathcal{A}}_M$ satisfies the following anti-symmetry property*

$$\big\langle \widehat{\mathcal{A}}_M F \, G \big\rangle^M_{K,\alpha} = -\big\langle F \widehat{\mathcal{A}}_M G \big\rangle^M_{K,\alpha}.$$

Proof. The definition of the conditional measure $\mu^M_K(\mathrm{d}\mathbf{z}_N \,|\, \mathbf{y}_M)$ gives the identity

$$\big\langle \widehat{\mathcal{A}}_M F \, G \big\rangle^M_{K,\alpha} = K \big\langle \mathcal{A}_N (F \circ \widehat{\mathbf{w}}) \, G \circ \widehat{\mathbf{w}} \big\rangle^N_\alpha$$

The proof then follows from Lemma 4.2.3. □

We can compute $\widehat{\mathcal{A}}_M$ explicitly. Given a coarse-grained state $\mathbf{y}_M \in \Gamma^M$, we denote the components of the ith cell by $y_i = (y_i^0, y_i^1, y_i^2) \in \Gamma$. Using property (4.29)

$$\mathcal{A}_N(\phi \circ \widehat{\mathbf{w}}) = \overline{\mathcal{A}}_N(\phi \circ \widehat{\mathbf{w}}) = -K^{-1} \sum_{\delta=0}^{2} \sum_{i \in \mathbb{Z}_M} \mathrm{J}^\delta_{l_i^-} \left(\partial_{y_i^\delta} \phi - \partial_{y_{i-1}^\delta} \phi \right) \circ \widehat{\mathbf{w}}.$$

Therefore,

$$\widehat{\mathcal{A}}_M \phi = -\sum_{\delta=0}^{2} \sum_{i \in \mathbb{Z}_M} \widehat{\mathrm{J}}^\delta_i \left(\partial_{y_i^\delta} - \partial_{y_{i-1}^\delta} \right) \phi,$$

where

$$\widehat{\mathrm{J}}^\delta_i(\mathbf{y}_M) = \big\langle \mathrm{J}^\delta_{l_i^-} \,\big|\, \mathbf{y}_M \big\rangle_N = \left\langle \frac{1}{K} \sum_{j \in \Lambda_i} \mathrm{J}^\delta_j \,\middle|\, \mathbf{y}_M \right\rangle.$$

Using the fact that $\mu_K^M(\mathrm{d}\mathbf{z}_N \,|\, \mathbf{y}_M)$ is a product of micro-canonical measures $\mu_K(\mathrm{d}\mathbf{z}_{\Lambda_i}\,|y_i)$ on each cell, and is therefore symmetric with respect to permutations of the indices inside each cell, we find

$$\begin{aligned}
\widehat{\mathrm{J}}_i^0(\mathbf{y}_M) &= \left\langle \frac{1}{K}\sum_{j\in\Lambda_{i-1}} v_j \,\middle|\, \mathbf{y}_M \right\rangle_N = y_{i-1}^1, \\
\widehat{\mathrm{J}}_i^1(\mathbf{y}_M) &= \left\langle \frac{1}{K}\sum_{j\in\Lambda_i} V'(r_j) \,\middle|\, \mathbf{y}_M \right\rangle_N = -P_K(y_i), \\
\widehat{\mathrm{J}}_i^2(\mathbf{y}_M) &= \left\langle \left(\frac{1}{K}\sum_{j\in\Lambda_{i-1}} v_j\right)\left(\frac{1}{K}\sum_{j\in\Lambda_i} V'(r_j)\right) \,\middle|\, \mathbf{y}_M \right\rangle_N = -y_{i-1}^1 P_K(y_i).
\end{aligned}$$

If one reverts back to a more transparent notation, and denotes $y_i = (\ell_i, p_i, e_i)$, then $\widehat{\mathcal{A}}_M$ takes the following form

$$\widehat{\mathcal{A}}_M = \sum_{i\in\mathbb{Z}_M} \left(-p_{i-1}(\partial_{\ell_i} - \partial_{\ell_{i-1}}) + P_K(\ell_i, p_i, e_i)\mathcal{X}_i\right),$$

where $\mathcal{X}_i = (\partial_{p_i} - \partial_{p_{i-1}}) + u_{i-1}(\partial_{e_i} - \partial_{e_i})$. The operator $\widehat{\mathcal{A}}_M$ can be seen to generate the following *discrete Euler* dynamics,

$$\begin{aligned}
\dot{\ell}_i &= p_i - p_{i-1} \\
\dot{p}_i &= P_K(\ell_i, e_i - \tfrac{1}{2}p_i^2) - P_K(\ell_{i+1}, e_{i+1} - \tfrac{1}{2}p_{i+1}^2) \\
\dot{e}_i &= u_{i-1}P_K(\ell_i, e_i - \tfrac{1}{2}p_i^2) - u_i P_K(\ell_{i+1}, e_{i+1} - \tfrac{1}{2}p_{i+1}^2),
\end{aligned} \tag{4.30}$$

which can be viewed as a discretization of the Euler equations in Lagrangian form (this system will be discussed in more detail in Section 4.6).

Remark 4.5.2. Contrary to the behavior of $\mathcal{A}_N$ and despite Lemma 4.5.1, it is important to note that $\widehat{\mathcal{A}}_M$ is *not* skew symmetric with respect to Lebesgue measure. In fact, the vector field associated with $\widehat{\mathcal{A}}_M$ (and written on the right-hand side of (4.30)) is not divergence free since the pressure depends on the energy e_i and therefore $\mathcal{X}_i P_K(\ell_i, p_i, e_i) \neq 0$.

Discrete Euler Dynamics

In this section we devote some discussion to the properties of the discrete Euler system (4.30). In general the discrete Euler system consists of a collection of 'parcels' $\{(\ell_i, p_i, e_i)\}_{i\in\mathbb{T}_N}$ on the phase space $\Gamma^N = (\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+)^{\mathbb{Z}_N}$ governed by a general concave thermodynamic entropy function $S(\ell, u)$,

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\ell_i &= p_i - p_{i-1}, \\ \tfrac{\mathrm{d}}{\mathrm{d}t}p_i &= P(\ell_i, e_i - \tfrac{1}{2}p_i^2) - P(\ell_{i+1}, e_{i+1} - \tfrac{1}{2}p_{i+1}^2) \\ \tfrac{\mathrm{d}}{\mathrm{d}t}e_i &= -p_{i-1}P(\ell_i, e_i - \tfrac{1}{2}p_i^2) + p_i P(\ell_{i+1}, u_{i+1} - \tfrac{1}{2}p_{i+1}^2) \end{aligned} \tag{4.31}$$

where $P(\ell, u)$ is thermodynamic pressure function, defined by t he first law

$$\partial_u S = \beta, \quad \partial_\ell S = \beta P \tag{4.32}$$

where $\beta(\ell, u) = T(\ell, u)^{-1} > 0$ is the inverse thermodynamic temperature function. We will find it useful change variables to internal energy variables

$$u_i := e_i - \tfrac{1}{2}p_i^2$$

whereby the equations become

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\ell_i &= p_i - p_{i-1}, \\ \tfrac{\mathrm{d}}{\mathrm{d}t}p_i &= P(\ell_i, u_i) - P(\ell_{i+1}, u_{i+1}) \\ \tfrac{\mathrm{d}}{\mathrm{d}t}u_i &= -P(\ell_i, u_i)(p_i - p_{i-1}). \end{aligned} \tag{4.33}$$

The discrete Euler equations in form (4.31) will be referred as the *conservative form* and the discrete Euler equations in form (4.33) will be referred to as *internal* *form*. Similarly to classical Euler equations, system (4.33) is a Poisson system with Hamiltonian

$$\mathsf{H}_N = \sum_{i\in\mathbb{Z}_N} \left(\frac{1}{2} p_i^2 + u_i \right),$$

and Poisson bracket $\{\,\cdot\,,\,\cdot\,\}_N$ given by

$$\{f,g\}_N = \sum_{i\in\mathbb{Z}_N} \left[(\partial_{\ell_i} f - P_i \partial_{u_i} f)(\partial_{p_i} g - \partial_{p_{i-1}} g) - (\partial_{\ell_i} g - P_i \partial_{u_i} g)(\partial_{p_i} f - \partial_{p_{i-1}} f) \right],$$

where P_i denotes $P(\ell_i, u_i)$. Note that the Hamiltonian H_N and bracket $\{\,\cdot\,,\,\cdot\,\}_N$ is a direct discretization of the Hamiltonian and Poisson brackets associated to the compressible Euler system. The Liouville operator associated to this system is

$$\begin{aligned} \mathcal{L}_{\mathsf{H}} &= \sum_{i\in\mathbb{Z}_N} \left((p_i - p_{i-1})\partial_{\ell_i} + (P_i - P_{i+1})\partial_{p_i} - P_i (p_i - p_{i-1}) \partial_{u_i} \right) \\ &= \sum_{i\in\mathbb{Z}_N} \left((p_i - p_{i-1})\partial_{\ell_i} + P_i \mathcal{X}_i \right), \end{aligned} \tag{4.34}$$

where we have introduced, for later convenience, the family of differential operators $\{\mathcal{X}_i\}_{\mathbb{Z}_N}$ defined by

$$\mathcal{X}_i = \partial_{p_i} - \partial_{p_{i-1}} - (p_i - p_{i-1})\partial_{u_i}.$$

A consequence of the periodicity implies that the total length of the chain

$$\mathsf{L}_N = \sum_{i\in\mathbb{Z}_N} \ell_i,$$

is constant under the evolution. Also, it is easy to see that the operator $\mathcal{X}_i$ vanishes on the quantities

$$\tfrac{1}{2m} p_i^2 + \tfrac{1}{2m} p_{i-1}^2 + u_i, \quad \text{and} \quad p_i + p_{i-1},$$

thereby implying the conservation of energy H_N and total momentum

$$\mathsf{P}_N = \sum_{i\in\mathbb{Z}_N} p_i$$

by the dynamics.

Being a discrete model of compressible fluid dynamics we have an additional thermodynamic structure. As in the Euler equations, the thermodynamic relations (4.32) imply that

$$\tfrac{\mathrm{d}}{\mathrm{d}t} S(\ell_i, u_i) = \partial_\ell S_i \, \tfrac{\mathrm{d}}{\mathrm{d}t} \ell_i + \partial_{u_i} S_i \, \tfrac{\mathrm{d}}{\mathrm{d}t} u_i = 0,$$

and therefore the total entropy

$$\mathsf{S}_N = \sum_{i \in \mathbb{Z}_N} S(\ell_i, u_i)$$

is conserved by the dynamics. In fact, both the total length L_N, and the entropy S_N are Casimir invariants of Poisson bracket $\{\,\cdot\,,\,\cdot\,\}_N$, meaning that for any smooth function g on Γ^N (not just the Hamiltonian), we have

$$\{\mathsf{S}_N, g\}_N = \{\mathsf{L}_N, g\}_N = 0.$$

The conserved quantities $(\mathsf{L}_N, \mathsf{P}_N, \mathsf{H}_N)$ have corresponding locally conserved fields $U_i = (\ell_i, p_i, e_i) \in \Gamma$ and fluxes

$$\widehat{\mathrm{J}}_i = (\widehat{\mathrm{J}}_i^\ell, \widehat{\mathrm{J}}_i^p, \widehat{\mathrm{J}}_i^e) = (p_{i-1}, -P_i, -p_{i-1} P_i),$$

so that the conservative form (4.31) can be written as a discrete conservation law

$$\dot{U}_i = \widehat{\mathrm{J}}_{i+1} - \widehat{\mathrm{J}}_i,$$

corresponding to conservation of length, momentum, and energy of the fluid chain.

Invariant measures and generalized canonical ensemble

Given the conserved quantities $(\mathsf{L}_N, \mathsf{P}_N, \mathsf{H}_N, \mathsf{S}_N)$, it is natural to seek invariant measures with density proportional to

$$\exp\left\{-\zeta(\mathsf{H}_N + \lambda \mathsf{P}_N) - \tau \mathsf{L}_N + \mathsf{S}_N\right\}.$$

What's important about the Poisson nature of this model is that this measure is *not* an invariant measure for the dynamics even though it is a function of the conserved quantities. The main difficultly with finding an invariant measure is due to the non-canonical Hamiltonian structure, and the fact that the Hamiltonian vector field X_{H} associated to the Poisson bracket $\{\,\cdot\,,\,\cdot\,\}_N$ and the Hamiltonian H_N, defined by

$$X_{\mathsf{H}} \cdot \nabla f = \{\mathsf{H}, f\}_N,$$

is not divergence free on Γ^N since the evolution equation for the energy evolves according to a function of the energy itself (this was mentioned as well in Remark 4.5.2)

Regardless of this difficultly, the thermodynamic structure allows us to find an invariant measure explicitly. We have the following result

Lemma 4.6.1. *Let β_i denote $\beta(\ell_i, u_i) = \partial_e S(\ell_i, u_i)$, then $\mathcal{L}_{\mathsf{H}}$ satisfies,*

$$\mathcal{L}_{\mathsf{H}}^* \prod_{i \in \mathbb{Z}_N} \beta_i = 0, \tag{4.35}$$

and therefore the product measure $\prod_{i \in \mathbb{Z}_N} \beta(\ell_i, u_i)\, \mathrm{d}\ell_i \mathrm{d}p_i \mathrm{d}u_i$ is an invariant measure for $\mathcal{L}_{\mathsf{H}}$ on Γ^N.

Proof. We being by using the Maxwell relation $\partial_\ell \beta = \partial_e(\beta P)$ to conclude,

$$\mathcal{L}_{\mathsf{H}} \log \beta_i = (p_i - p_{i-1})\beta_i^{-1}(\partial_\ell \beta_i - P_i \partial_e \beta_i) = (p_i - p_{i-1})\partial_e P_i.$$

Therefore since

$$\prod_i \beta_i = \exp\Big\{\sum_i \log \beta_i\Big\},$$

we have

$$\mathcal{L}_{\mathsf{H}} \prod_{i\in\mathbb{Z}_N} \beta_i = \sum_{i\in\mathbb{Z}_N} (\mathcal{L}_{\mathsf{H}} \log \beta_i) \prod_{j\in\mathbb{Z}_N} \beta_j = \Big(\sum_{i\in\mathbb{Z}_N} (p_i - p_{i-1})\partial_e P_i\Big) \prod_{j\in\mathbb{Z}_N} \beta_j.$$

Using the fact that the $L^2(\Gamma^N)$ adjoint of the Liouville operator $\mathcal{L}_{\mathsf{H}}$ is,

$$\mathcal{L}_{\mathsf{H}}^* = -\mathcal{L}_{\mathsf{H}} + \sum_{i\in\mathbb{Z}_N} (p_i - p_{i-1})\partial_e P_i,$$

we conclude (4.35). □

Naturally this leads us define to a version of the canonical ensemble, a probability measure ν_α^N on Γ^N defined by

$$\begin{aligned} \mathrm{d}\nu_\alpha^N &= \frac{1}{Z(\alpha)^N} \exp\big\{-\alpha_1 \mathsf{L}_N - \alpha_2 \mathsf{P}_N - \alpha_3 \mathsf{H}_N + \mathsf{S}_N\big\} \prod_{i\in\mathbb{Z}_N} \beta(\ell_i, u_i)\,\mathrm{d}\ell_i \mathrm{d}p_i \mathrm{d}u_i \\ &= \prod_{i\in\mathbb{Z}_N} \frac{1}{Z(\alpha)} \exp\big\{-\alpha_1 \ell_i - \alpha_2 p_i - \alpha_3 \tfrac{1}{2} p_i^2 - \alpha_3 u_i + S(\ell_i, u_i)\big\}\beta(\ell_i, u_i)\,\mathrm{d}\ell_i \mathrm{d}p_i \mathrm{d}u_i, \end{aligned} \tag{4.36}$$

where the normalization factor is

$$Z(\alpha) = \sqrt{2\pi\alpha_3^{-1}}\, e^{\frac{m\alpha_2^2}{2\alpha_3}}\, \overline{Z}(\alpha_1, \alpha_3)$$

and

$$\overline{Z}(\alpha_1, \alpha_3) = \iint_{\mathbb{R}_+\times\mathbb{R}_+} \exp\big\{-\alpha_1 \ell - \alpha_3 u + S(\ell, u)\big\}\beta(\ell, u)\,\mathrm{d}\ell \mathrm{d}u. \tag{4.37}$$

We will call any measure of the form ν_α^N, for some particular choice of parameters $\alpha \in \Gamma$ a *generalized canonical measure.* We will often drop the dependence on the parameters and denote the measure by ν^N. As the definition in (4.36) implies, a canonical measure ν^N can be written as a product of N *one-particle measures* ν,

$$\nu_\alpha^N = \nu_\alpha^{\otimes N},$$

where

$$\mathrm{d}\nu = \frac{e^{-\frac{m\alpha_2^2}{2\alpha_3}}}{\overline{Z}(\alpha_1,\alpha_3)\sqrt{2\pi m\alpha_3^{-1}}} \mathrm{e}^{-\alpha_1\ell-\alpha_2 p-\alpha_3(\frac{1}{2}p^2+u)+S(\ell,u)}\beta(\ell,u)\,\mathrm{d}\ell\mathrm{d}p\mathrm{d}u.$$

To ensure that the normalization constant $\overline{Z}(\alpha_1,\alpha_3)$ is finite, we will require the following assumptions on the entropy function,

Hypothesis 4.6.2. *The entropy function $S : (0,\infty)\times(0,\infty)\to\mathbb{R}$, is C^2, concave and has the following properties*

1. *Positive temperature*

$$\beta(\ell,u) = \partial_u S(\ell,u) > 0 \quad \textit{on} \quad (0,\infty)\times(0,\infty)$$

2. *For each $u \in (0,\infty)$,*

$$\lim_{\ell\to 0} S(\ell,u) = -\infty,$$

3. *For each $\ell \in (0,\infty)$,*

$$\lim_{u\to 0} S(\ell,u) = -\infty,$$

4. *Sub-linear growth*

$$\lim_{(\ell,e)\to\infty} \frac{S(\ell,e)}{\ell+e} = 0.$$

5. For each $\ell \in (0, \infty)$,

$$\lim_{u \to 0} [S(\ell, u) + \log \beta(\ell, u)] = -\infty.$$

The finiteness of $\overline{Z}(\alpha_1, \alpha_3)$ now follows from hypothesis 4.6.2. In fact we have

Proposition 4.6.3. *Let $S(\ell, e)$ be a entropy function satisfying hypothesis 4.6.2, if the parameters $\alpha_1, \alpha_3 \in (0, \infty)$, then*

$$\overline{Z}(\alpha_1, \alpha_3) < \infty.$$

Proof. Since $\beta \exp\{S\} = \partial_u \exp\{S\}$ we may use integration by parts and the growth conditions in hypothesis 4.6.2 to obtain

$$\overline{Z}(\alpha_1, \alpha_3) = \alpha_3 \iint_{\mathbb{R}_+ \times \mathbb{R}_+} \exp\Big\{ -\alpha_1 \ell - \alpha_3 u + S(\ell, u) \Big\} \mathrm{d}\ell \mathrm{d}u.$$

The function inside the exponential can be bounded by

$$-\alpha_1 \ell - \alpha_3 u + S(u, \ell) \leq -(l + e)\Big(\min\{\alpha_1, \alpha_3\} - \frac{S(\ell, u)}{\ell + u} \Big).$$

Again the growth condition in hypothesis 4.6.2 implies that there is an $R > 0$ such that on the set $\{\ell + u > R\}$,

$$\frac{S(\ell, u)}{\ell + u} < \tfrac{1}{2} \min\{\alpha_1, \alpha_3\}.$$

Since $-\alpha_1 \ell - \alpha_3 u + S(e, \ell)$ is bounded above on $\{u + \ell \leq R\}$, we only need to ensure that the integral on $\{\ell + u > R\}$ is finite. This follows since

$$\iint_{\{\ell + u > R\} \cap \mathbb{R}_+ \times \mathbb{R}_+} \exp\Big\{ -\alpha_1 \ell - \alpha_3 u + S(\ell, u) \Big\} \mathrm{d}\ell \mathrm{d}u$$

$$\leq \iint_{\mathbb{R}_+ \times \mathbb{R}_+} \exp\Big\{ -\tfrac{1}{2} \min\{\alpha_1, \alpha_3\}(\ell + u) \Big\} \mathrm{d}\ell \mathrm{d}u < \infty.$$

□

Remark 4.6.4. It's important to realize that under the conditions on S defined in 4.6.2, the measure

$$e^{\mathsf{S}_N} \prod_{i \in \mathbb{Z}_N} \beta_i \mathrm{d}\ell_i \mathrm{d}u_i$$

on $(\mathbb{R}_+ \times \mathbb{R}_+)^{\mathbb{Z}_N}$ is *not* a bounded measure. Therefore one must be careful to ensure that $\alpha_1, \alpha_2 > 0$ when defining the measure ν_α^N, $\alpha = (\alpha_1, \alpha_2, \alpha_3)$.

Remark 4.6.5. The condition that $\lim_{(\ell,u)\to 0} S(\ell, u) = -\infty$ appears to be necessary for $Z(\alpha_1, \alpha_3)$ to remain finite. In fact, as we will see, it will also be necessary to work out the correct expressions for average temperature and pressure, and is crucial for ensuring that the parcels don't collapse to zero in finite time. This type of singularity is present, for instance, in the equation for the entropy of an ideal gas, which takes the form

$$S_{\text{Ideal}}(\ell, u) = (c_P - c_V) \log(\ell) + c_V \log(u), \quad c_V > 1,$$

where c_V and c_P are the specific heats at constant volume and pressure. In, fact the same type of logarithmic singularity is present in the expression for the entropy of a Van der Walls gas.

It is now a simple consequence of the fact that $\mathsf{H}_N, \mathsf{P}_N, \mathsf{L}_N, \mathsf{S}_N$, are conserved by the dynamics, and the fact that $\prod_i \beta_i$ is invariant, that any canonical measure ν^N is an invariant probability measure for the dynamics, i.e. for every bounded and continuous $\varphi : \Gamma^N \to \mathbb{R}$,

$$\int_{\Gamma^N} \mathcal{L}_{\mathsf{H}} \varphi \, \mathrm{d}\nu^N = 0.$$

As a consequence we have the following anti-symmetry property,

Proposition 4.6.6. *Let ν^N be a canonical measure on Γ^N. Then then the operator $\mathcal{L}_{\mathsf{H}}$, defined in (4.34) is skew symmetric with respect to ν^N.*

Not only does the discrete Euler system have a local thermodynamic structure determined by $S(\ell, u)$ also has a global thermodynamic structure determined by the generalized canonical ensemble (4.36). Similarly to Section 4.2.2, we may define the global free energy

$$\widehat{F}(\alpha) = \log Z(\alpha)$$

for $\alpha \in \Gamma$ and its corresponding concave global entropy function

$$\widehat{S}(\overline{U}) = \inf_{\alpha \in \Gamma} \left(\overline{U} \cdot \alpha + F(\alpha) \right),$$

where $\overline{U} = (\bar{\ell}, \bar{p}, \bar{e}) \in \Gamma$. Of course Lemma A.2.2 implies that $\widehat{S}$ is smooth and strictly concave. Moreover, just as with the entropy defined by the grand canonical ensemble we have the Galilean shift property

$$\widehat{S}(\bar{\ell}, \bar{p}, \bar{e}) = \widehat{S}(\bar{\ell}, 0, \bar{e} - \tfrac{1}{2}).$$

With an abuse of notation sometimes denote

$$\widehat{S}(\bar{\ell}, \bar{u}) = \widehat{S}(\bar{\ell}, 0, \bar{u}).$$

We would like to determine the physical meaning of the parameters α in the canonical ensemble. To simplify matters we will define the parameters $(\tau, \lambda, \zeta) \in \Gamma$ by

$$\tau := \alpha_1, \quad \lambda := \alpha_2/\alpha_3, \quad \zeta := \alpha_3.$$

Obviously, we will assume that $\alpha_1, \alpha_3 > 0$ so that λ is well defined and so that the measure ν is a well defined probability measure. The physical meaning of the parameters (ζ, λ, τ) can be identified at the level of the one-particle measure ν. Let $\langle \cdot \rangle_\nu$ denote the expectation with respect to the one-particle measure ν,

$$\langle f \rangle_\nu := \int_\Gamma f \, \mathrm{d}\nu,$$

and similarly let $\langle \cdot \rangle_{\nu^N}$ denote the expectation with respect to generalized canonical measure ν^N.

Since ν^N is just a product of N one-particle measures ν, we see that averages of the quantities $(\mathsf{H}_N, \mathsf{P}_N, \mathsf{L}_N)$ can be expressed in terms of one-particle averages, being sums of functions over the one-particle phase space,

$$\langle \mathsf{H}_N \rangle_{\nu^N} = N \langle \tfrac{1}{2}p^2 + u \rangle_\nu, \quad \langle \mathsf{P}_N \rangle_{\nu^N} = N \langle p \rangle_\nu, \quad \langle \mathsf{L}_N \rangle_{\nu^N} = N \langle \ell \rangle_\nu$$

In fact since the ν is a Gaussian integral in p, we may explicitly compute powers of p,

$$\langle p \rangle_\nu = \lambda,$$

$$\langle \tfrac{1}{2}p^2 \rangle_\nu = \tfrac{1}{2}\lambda^2 + \tfrac{1}{2}\zeta^{-1}.$$

Therefore λ corresponds to the average velocity per particle. To compute the averages of $\frac{1}{2}p^2 + u$, and ℓ, recall the definition of $Z(\tau, \zeta)$ in (4.37), then the averages are given by

$$\langle u \rangle_\nu = -\partial_\zeta \log Z(\tau, \zeta)$$

$$\langle \ell \rangle_\nu = -\partial_\tau \log Z(\tau, \zeta).$$

It follows that the average energy is

$$\langle \tfrac{1}{2}p^2 + u \rangle_\nu = \tfrac{1}{2}\lambda^2 + \tfrac{1}{2}\zeta^{-1} - \partial_\zeta \log Z(\tau, \zeta).$$

Taking into account the thermal contribution $\frac{1}{2}\zeta^{-1}$ appearing above, we define the average internal energy $\bar{u}$ as the average energy $\langle E \rangle_\nu$ minus the contribution from the kinetic energy $\frac{1}{2}\lambda^2$ associated to the mean velocity λ,

$$\bar{e} = \tfrac{1}{2}\zeta^{-1} - \partial_\zeta \log Z(\tau, \zeta).$$

Similarly we define the average cell length $\bar{\ell}$ as

$$\bar{\ell} = -\partial_\tau \log Z(\tau, \zeta).$$

Given a prescribed mean length and internal energy $(\bar{\ell}, \bar{u})$, it is straight forward the the strict convexity of $\widehat{S}$ that one may solve the system of equations

$$\begin{aligned}\bar{u} &= \tfrac{1}{2}\zeta^{-1} - \partial_\zeta \log Z(\tau, \zeta)\\ \bar{\ell} &= -\partial_\tau \log Z(\tau, \zeta),\end{aligned}$$

for (ζ, τ), to obtain

$$\begin{aligned}\zeta &= \partial_{\bar{u}}\widehat{S}(\bar{\ell}, \bar{u})\\ \tau &= \partial_{\bar{\ell}}\widehat{S}(\bar{\ell}, \bar{u}).\end{aligned} \tag{4.38}$$

In calling $\widehat{S}$ the global entropy, we have implied that $\widehat{S}$ satisfies the first law of thermodynamics. This would suggest that ζ corresponds to the inverse temperature, while τ corresponds to the pressure multiplied by the inverse temperature. In fact this can be seen explicitly using properties of the one-particle measure g. Recall that the first law states that $\beta = \partial_u S$. Using this and integration by parts, we find

$$\begin{aligned}Z(\tau, \zeta) &= \int_{\mathbb{R}^2_+} \exp\big\{-\zeta u - \tau\ell\big\}\partial_u \exp\big\{S(\ell, u)\big\}\mathrm{d}\ell\mathrm{d}u\\ &= \zeta \int_{\mathbb{R}^2_+} \exp\big\{-\zeta u - \tau\ell + S(\ell, u)\big\}\mathrm{d}\ell\mathrm{d}u\\ &= \zeta Z(\tau, \zeta)\langle\, T\,\rangle_\nu.\end{aligned}$$

Dividing both sides by $Z(\tau, \zeta)$ and multiplying by ζ^{-1} we obtain

$$\langle\, T\,\rangle_\nu = \zeta^{-1}, \tag{4.39}$$

so that ζ^{-1} corresponds to the average temperature of the cells. Similarly, using the fact that $\beta P = \partial_\ell S$, we find that

$$\begin{aligned}\langle P \rangle_\nu &= \frac{1}{Z(\tau,\zeta)} \int_{\mathbb{R}_+^2} \exp\big\{-\zeta u - \tau\ell\big\}\partial_\ell \exp\big\{S(\ell,u)\big\} \mathrm{d}\ell \mathrm{d}u \\ &= \frac{\tau}{Z(\tau,\zeta)} \int_{\mathbb{R}_+^2} \exp\big\{-\zeta u - \tau\ell + S(\ell,u)\big\} \mathrm{d}\ell \mathrm{d}u \\ &= \tau\langle T \rangle_\nu,\end{aligned}$$

Using the relation (4.39) obtained for the average temperature, find that

$$\langle P \rangle_\nu = \tau\zeta^{-1}.$$

So that τ corresponds to the average pressure divided by the average temperature. This verifies the role played by $\widehat{S}$ as the global entropy with the relation (4.38) implying,

$$\begin{aligned}\partial_{\bar u}\widehat{S}(\bar\ell,\bar u) &= \langle T \rangle_\nu^{-1} \\ \partial_{\bar\ell}\widehat{S}(\bar\ell,\bar u) &= \langle T \rangle_\nu^{-1}\langle P \rangle_\nu.\end{aligned} \tag{4.40}$$

We will define global pressure $\widehat{P}(\bar\ell,\bar u)$ and global temperature $\widehat{T}(\bar\ell,\bar u)$, by

$$\widehat{P}(\bar\ell,\bar u) := \langle P \rangle_\nu, \quad \widehat{T}(\bar\ell,\bar u) := \langle T \rangle_\nu.$$

where the dependence on $(\bar\ell,\bar u)$, comes from the fact that the parameters (ζ,τ) in the measure $\nu = \nu_{\tau,0,\zeta}$ are given by relation (4.38). It follows from (4.40) that $\widehat{P}$ and $\widehat{T}$ satisfy the first law of thermodynamics

$$\partial_{\bar u}\widehat{S} = \widehat{\beta}, \quad \partial_{\bar\ell}\widehat{S} = \widehat{\beta}\widehat{P}.$$

Ideal Gas Fixed Point

As we saw in Section 4.5.2, the process of coarse-graining in equilibrium gives a procedure for obtaining the discrete Euler dynamics (4.30) from the microscopic Hamiltonian system (4.10). When the cell size K is large, we know from Theorem 4.2.16 that the volume entropy S_K governing the discrete Euler equations approaches the thermodynamic entropy S and we obtain the (infinite) discrete Euler system

$$\begin{aligned}
\tfrac{\mathrm{d}}{\mathrm{d}t}\ell_i &= p_i - p_{i-1}, \\
\tfrac{\mathrm{d}}{\mathrm{d}t}p_i &= P(\ell_i, e_i - \tfrac{1}{2}p_i^2) - P(\ell_{i+1}, e_{i+1} - \tfrac{1}{2}p_{i+1}^2) \\
\tfrac{\mathrm{d}}{\mathrm{d}t}e_i &= -p_{i-1}P(\ell_i, e_i - \tfrac{1}{2}p_i^2) + p_i P(\ell_{i+1}, u_{i+1} - \tfrac{1}{2}p_{i+1}^2)
\end{aligned} \tag{4.41}$$

corresponding to the thermodynamic entropy S. Following the renormalization group approach in statistical mechanics, we can view this as a mapping between models. Naturally, we are interested in applying this coarse-graining procedure again to the infinite discrete Euler system through the map

$$\{(\ell_i, p_i, e_i)\}_{i\in\mathbb{Z}} \mapsto \{(\hat{\ell}_i, \hat{p}_i, \hat{e}_i)\}_{i\in\mathbb{Z}}$$

where

$$(\hat{\ell}_i, \hat{p}_i, \hat{e}_i) = \frac{1}{K}\sum_{j\in\Lambda_i}(\ell_i, p_i, e_i),$$

and $\{\Lambda_i\}_{i\in\mathbb{Z}}$ is a partition of $\mathbb{Z}$ with cells of size K. We will denote such a map by $\hat{h}(\mathbf{y}) = \hat{\mathbf{y}}$, where $\mathbf{y} = \{(\ell_i, p_i, e_i)\}_{i\in\mathbb{Z}}$ and $\hat{\mathbf{y}} = \{(\hat{\ell}_i, \hat{p}_i, \hat{e}_i)\}_{i\in\mathbb{Z}}$. We consider the invariant probability measure

$$\mathrm{d}\nu_\alpha^\infty = \prod_{i\in\mathbb{Z}} \frac{1}{Z(\alpha)} \mathrm{e}^{-\alpha_1\ell_i - \alpha_2 p_i - \alpha_3\frac{1}{2}p_i^2 - \alpha_3 u_i + S(\ell_i,u_i)} \beta(\ell_i, u_i)\, \mathrm{d}\ell_i \mathrm{d}p_i \mathrm{d}u_i,$$

associated to (4.41) and denote $\langle \cdot \,|\, \hat{\mathbf{y}} \rangle_{\nu^\infty, K}$ the conditional probability measure obtained by conditioning ν_α^∞ on the event $\{\hat{\mathbf{h}}(\mathbf{y}) = \hat{\mathbf{y}}\}$ and denote the mapping in each cells Λ_i by $h_i(\mathbf{y})$. If one denotes

$$\mathcal{L} = \sum_{i \in \mathbb{Z}} -p_i(\partial_{\ell_i} - \partial_{\ell_{i-1}}) + P_i(\partial_{p_i} - \partial_{p_{i-1}}) + p_{i-1}P_i(\partial_{e_i} - \partial_{e_{i-1}})$$

the generator of the (4.41), then we aim to study the coarse-grained generator

$$\widehat{\mathcal{L}}_K \phi(\hat{\mathbf{y}}) = \langle \mathcal{L}(\phi \circ \hat{\mathbf{h}}) \,|\, \hat{\mathbf{y}} \rangle_{\nu^\infty, K}$$

where ϕ is a local function on $\Gamma^{\mathbb{Z}} = (\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}_+)^{\mathbb{Z}}$. Using the equivalence of ensembles Theorem A.3.7, and the properties of averages with respect to the generalized canonical ensemble obtained in Section 4.6 we obtain the following limit

$$\widehat{\mathcal{L}}\phi = \lim_{K \to \infty} \widehat{\mathcal{L}}_K \phi,$$

where

$$\widehat{\mathcal{L}} = \sum_{i \in \mathbb{Z}} -\hat{p}_i(\partial_{\hat{\ell}_i} - \partial_{\hat{\ell}_{i-1}}) + \widehat{P}_i(\partial_{\hat{p}_i} - \partial_{\widehat{p_{i-1}}}) + \hat{p}_{i-1}\widehat{P}_i(\partial_{\hat{e}_i} - \partial_{\widehat{e_{i-1}}})$$

and $\widehat{P}_i$ is the pressure associated with the global entropy function $\widehat{S}$. In particular, this shows that the discrete Euler system retains its form under consecutive coarse-graining procedures. In particular, if one considers the case where the entropy function is an *Ideal gas*

$$S(\ell, e) = (c_P - c_V)\log \ell + c_V \log e + C_2,$$

then the volume ℓ_i and the internal energy u_i follow a Gamma distribution under the measure ν_α^∞. Since such distributions are stable, it is a straightforward to see that entropy (which is just the large deviations rate function associated to the single particle measure) must invariant up to an additive constant, namely

$$\hat{S}(\hat{\ell}, \hat{e}) = (c_P - c_V) \log \hat{\ell} + c_V \log \hat{e} + C_2.$$

Therefore the discrete Euler system with Ideal gas equation of state constitute a fixed point of the renormalization group procedure.

In the case of an ideal gas, it is not hard to see that the discrete Euler equations (4.41) to the so-called *gamma-law* equations

$$\begin{aligned} \dot{\ell}_i &= p_i - p_{i-1} \\ \dot{p}_i &= (1-\gamma) e^{S_0/R} \left(\frac{1}{(\ell_{i+1})^\gamma} - \frac{1}{(\ell_i)^\gamma} \right), \end{aligned}$$

where S_0 is the initial entropy, $R = c_p - c_V$ is the gas constant and $\gamma = c_P/c_v > 1$ is the heat capacity ratio. What's interesting is that this system is again a one-dimensional particle chain with Hamiltonian

$$H_{\text{Ideal}} = \sum_i \frac{1}{2} p_i^2 + e^{S_0/R} (\ell_i)^{1-\gamma}.$$

Non-Equilibrium Coarse-graining and Corrections

We now want to coarse-grain our particle system in a fully non-equilibrium setting. If the particles are initially distributed on the phase space $\Omega^{\mathbb{Z}_N}$ according to a distribution $f_0^N(\mathbf{z}_N)$. Then the distribution $f_t^N(\mathbf{z}_N)$ at time $t > 0$ is governed by the Liouville equation

$$\partial_t f^N + \mathcal{A}_N f^N = 0, \quad f^N|_{t=0} = f_0^N. \tag{4.42}$$

Let $\Phi_t^N(\mathbf{z}_N)$ be flow the generated by the ODE (4.10). Then the solution to (4.42) can be represented by

$$f_t^N(\mathbf{z}_N) = f_0^N(\Phi_{-t}^N(\mathbf{z}_N)) \equiv e^{-\mathcal{A}_N t} f_0^N.$$

We will assume that the initial distribution is of the form $f_0^N = \hat{f}_t^M \circ \widehat{\mathbf{w}}$, for some $\hat{f}_0^M$ which is a *cyclically symmetric* distribution on Γ^N, meaning that $\hat{f}_0^M$ is invariant with respect to cyclic permutations in the indices. Of course, the fact that $\widehat{\mathbf{w}}$ is permutation symmetric inside the cells $\{\Lambda_i\}_{i\in\mathbb{Z}_N}$ implies that f_0^N is also cyclically symmetric on Ω^N. It is a simple consequence of the dynamics that cyclic permutations are preserved by the flow $\Phi_t^N(\mathbf{z}_N)$ and therefore that f_t^N is also cyclically symmetric.

Remark 4.8.1. The reason for symmetry with respect to cyclic permutations, as opposed any permutation, is a direct consequence of the nearest neighbor interactions of the particle system. Indeed, it is clear that nearest neighbor interactions would not be preserved under the flow if one swaps any two arbitrary indices. However, it is important to remark that the quantities L_N, P_N and H_N *are* symmetric with respect to *any* permutation. Therefore both the grand-canonical measure $\mu_\alpha^N(\mathrm{d}\mathbf{z}_N)$ and the micro-canonical measure $\mu_N(\mathrm{d}\mathbf{z}_N|\mathbf{y}_M)$ are symmetric with respect to any permutation. The intuition here being that once the particles are in equilibrium, they no longer feel the nearest neighbor interactions.

If $\mathbf{z}_N$ is distributed according to f_t^N, let $\hat{f}_t^M$ be the distribution of $\widehat{\mathbf{w}}(\mathbf{z})$, defined by pushforward of f_t^N as a measure

$$\hat{f}^M(t, \mathbf{y}_M)\mathrm{d}\mathbf{y}_M = \widehat{\mathbf{w}}_{\#}(f^N(t, \mathbf{z}_N)\mathrm{d}\mathbf{z}_N).$$

In addition, let $f_t^N(\mathrm{d}\mathbf{z}_N \,|\, \mathbf{y}_M)$ be the conditional measure obtained by conditioning $f_t^N(\mathbf{z}_N)\mathrm{d}\mathbf{z}_N$ on the event $\{\mathbf{y}_M = \widehat{\mathbf{w}}\}$. It is important to note that $f_t^N(\mathbf{z}_N \,|\, \mathbf{y}_M)$ is no-longer invariant with respect any type of permutation (this will in fact a have a profound effect on the behavior of the fluctuations).

Clearly, by the results of Section 4.5.2, if f_0 is distributed according to μ_α^N, then is is easy to see that $\hat{f}_t^M$ exactly solves the coarse-grained Liouville equation

$$\partial_t \hat{f}_t^M - \widehat{\mathcal{A}}_M^* \hat{f}_t^M = 0,$$

where $\widehat{\mathcal{A}}_M^*$ denotes the formal adjoint of $\widehat{\mathcal{A}}_M$ with respect to Lebesgue measure. Our goal will be to determine to what extent this is true when f_0^N is not in equilibrium. In this case $\hat{f}_t^M$ instead solves,

$$\partial_t \hat{f}_t^M - \widehat{\mathcal{A}}_M^* \hat{f}_t^M = \widetilde{R}_t[f_t^N], \tag{4.43}$$

where $\widetilde{R}_t[f_t^N]$ is a quantity describing the deviation from the equilibrium behavior, and depends on the microscopic distribution f_t^N. Our goal will be to understand the behavior of deviation $\widetilde{R}_t[f_t^N]$.

In what follows, we will denote the average with respect to $f_t^N(\mathrm{d}\mathbf{z}_N \,|\, \mathbf{y}_M)$ by $\langle \cdot \,|\, \mathbf{y}_M \rangle_{f_t^N}$ and the averages with respect to $f_t^N(\mathrm{d}\mathbf{z}_N)\mathrm{d}\mathbf{z}_N$ and $\hat{f}_t^M(\mathbf{y}_M)\mathrm{d}\mathbf{y}_M$ by $\langle \cdot \rangle_{f_t^N}$ and $\langle \cdot \rangle_{\hat{f}_t^M}$.

Our first step is to make precise the equation (4.43) $\hat{f}^M$ in weak form,

Lemma 4.8.2. *Let $\phi \in C_b^1(\Gamma^M)$, then we have the following evolution equation for $\hat{f}_t^M$*

$$\partial_t \langle \phi \rangle_{\hat{f}_t^M} - K^{-1} \big\langle \widehat{\mathcal{A}}_M \phi \big\rangle_{\hat{f}_t^M} = -K^{-1} \sum_{\delta=0}^{2} \sum_{i \in \mathbb{Z}_M} \Big\langle \langle \widetilde{\mathrm{J}}_i^\delta \,|\, \mathbf{y}_M \rangle_{f_t^N} D_i^\delta \phi \Big\rangle_{\hat{f}_t^M},$$

where

$$\widetilde{\mathrm{J}}_i^{\alpha}(\mathbf{z}_N) = \mathrm{J}^{\delta}_{l_i^-}(\mathbf{z}_N) - \widehat{\mathrm{J}}^{\delta}_i(\widehat{\mathbf{w}}(\mathbf{z}_N)), \quad \textit{and} \quad D_i^{\delta} = \partial_{y_i^{\delta}} - \partial_{y_{i-1}^{\delta}}.$$

Remark 4.8.3. Of course Lemma 4.8.2 implies that the deviation $\widetilde{R}_t[f_t^N]$ described in equation (4.43) can be written as

$$\widetilde{R}_t[f_t^N] = K^{-1} \sum_{\delta=0}^{2} \sum_{i\in\mathbb{Z}_M} D_i^{\delta} \left(\langle \widetilde{\mathrm{J}}_i^{\delta} \,|\, \mathbf{y}_M \rangle_{f_t^N} \hat{f}_t^M \right).$$

We would like to study the evolution of the term

$$\left\langle \langle \widetilde{\mathrm{J}}_i^{\delta} \,|\, \mathbf{y}_M \rangle_{f_t^N} D_i^{\delta} \phi \right\rangle_{\hat{f}_t^M}.$$

To do this, recall the decomposition

$$\mathcal{A}_N = \mathring{\mathcal{A}}_M + \overline{\mathcal{A}}_M,$$

where $\mathring{\mathcal{A}}_M$ and $\overline{\mathcal{A}}_M$ are the periodized Liouville and boundary interaction operators defined in equations (4.25) and (4.27) respectively. It is easy to see that $\mathring{\mathcal{A}}_M$ induces a well-defined unitary group $(e^{t\mathring{\mathcal{A}}_N})_{t\in\mathbb{R}}$ on $C_b(\Omega^N)$. Indeed, since each $\mathring{\mathcal{A}}_{\Lambda_i}$ only acts on indices in Λ_i, we can represent $(e^{t\mathring{\mathcal{A}}_N})_{t\in\mathbb{R}}$ by

$$e^{t\mathring{\mathcal{A}}_M} F(\mathbf{z}_N) = F\left(\Phi_t^K(\mathbf{z}_{\Lambda_1}), \Phi_t^K(\mathbf{z}_{\Lambda_2}), \ldots, \Phi_t^K(\mathbf{z}_{\Lambda_M})\right),$$

where Φ_t^K is the flow associated to the dynamics (4.10) with $N = K$. An immediate consequence of this representation is the following invariance property

$$e^{t\mathring{\mathcal{A}}_M} \phi \circ \widehat{\mathbf{w}} = \phi \circ \widehat{\mathbf{w}}, \quad \text{for each} \quad \phi \in C_b(\Gamma^M).$$

Moreover, since $\mu_K^M(\mathrm{d}\mathbf{z}_N \,|\, \mathbf{y}_M)$ is just a product of micro-canonical measures $\mu_K(\mathrm{d}\mathbf{z}_{\Lambda_i} \,|\, y_i)$ on each cell, Lemma 4.2.5 implies that $\mathring{\mathcal{A}}_M$ and is skew-symmetric with respect to

$\mu_K^M(\mathrm{d}\mathbf{z}_N \,|\, \mathbf{y}_M)$ and therefore

$$\left\langle G\, e^{t\mathring{\mathcal{A}}_M} F \,|\, \mathbf{y}_M \right\rangle_N = \left\langle e^{-t\mathring{\mathcal{A}}_M} G\, F \,|\, \mathbf{y}_M \right\rangle_N ,$$

for all $F, G \in C_b(\Omega^N)$. Of course, this implies that if ψ satisfies $\langle \psi \,|\, \mathbf{y}_M \rangle_N = 0$, then so does $e^{t\mathring{\mathcal{A}}_M}\psi$ for all $t \in \mathbb{R}$.

Using Duhammel's formula and the decomposition (4.28), we may write the Liouville equation (4.42) as

$$f_t^N = \hat{f}_0^M \circ \widehat{\mathbf{w}} - \int_0^t e^{-(t-s)\mathring{\mathcal{A}}_N} \overline{\mathcal{A}}_N f_s^N \mathrm{d}s. \tag{4.44}$$

This can easily be made rigorous by working in the weak form and choosing time dependent test functions. Setting, $\psi(\mathbf{z}_N) = \widetilde{\mathrm{J}}_i^\delta(\mathbf{z}_N) D_i^\delta \phi(\widehat{\mathbf{w}}(\mathbf{z}_N))$, and integrating both sides of (4.44) against ψ, we obtain the following formula for $\left\langle \langle \widetilde{\mathrm{J}}_i^\delta \,|\, \mathbf{y}_M \rangle_{f_t^N} D_i^\delta \phi \right\rangle_{\hat{f}_t^M}$.

Lemma 4.8.4. *For each $\phi \in C_b^2(\Gamma^M)$ we have*

$$\begin{aligned}\left\langle \langle \widetilde{\mathrm{J}}_i^\delta \,|\, \mathbf{y}_M \rangle_{f_t^N} D_i^\delta \phi \right\rangle_{\hat{f}_t^M} &= \int_0^t \left\langle \left\langle \overline{\mathcal{A}}_N e^{(t-s)\mathring{\mathcal{A}}_N} \widetilde{\mathrm{J}}_i^\delta \,\middle|\, \mathbf{y}_M \right\rangle_{f_s^N} D_i^\delta \phi \right\rangle_{\hat{f}_s^M} \mathrm{d}s \\ &\quad - K^{-1} \sum_{\gamma=0}^{2} \sum_{j \in \mathbb{Z}_M} \int_0^t \left\langle \left\langle (e^{(t-s)\mathring{\mathcal{A}}_N} \widetilde{\mathrm{J}}_i^\delta) \mathrm{J}_{l_j^-}^\gamma \,\middle|\, \mathbf{y}_M \right\rangle_{f_s^N} D_j^\gamma D_i^\delta \phi \right\rangle_{\hat{f}_s^M} \mathrm{d}s\end{aligned}$$

Combining Lemmas have obtained an exact evolution equation for $\hat{f}^N$ of the form

$$\begin{aligned}\partial_t \hat{f}_t^M - \widehat{\mathcal{A}}_M^* \hat{f}_t^M - K^{-1} \sum_{\delta=0}^{2} \sum_{i \in \mathbb{Z}_N} D_i^\delta \left(\int_0^t \mathcal{B}_j^\delta [(t-s), f_s^N] \hat{f}_s^M \mathrm{d}s \right) \\ - K^{-2} \sum_{\delta,\gamma=0}^{2} \sum_{i,j \in \mathbb{Z}_M} D_i^\delta D_j^\gamma \left(\int_0^t \mathcal{K}_{i,j}^{\gamma,\delta} [(t-s), f_s^N] \hat{f}_s^M \mathrm{d}s \right) = 0.\end{aligned} \tag{4.45}$$

where

$$\mathcal{B}_j^\delta [t, f_s^N](\mathbf{y}_M) := \left\langle \overline{\mathcal{A}}_N e^{t\mathring{\mathcal{A}}_N} \widetilde{\mathrm{J}}_i^\delta \,\middle|\, \mathbf{y}_M \right\rangle_{f_s^N} ,$$

and

$$\mathcal{K}^{\gamma,\delta}_{i,j}[t, f^N_s](\mathbf{y}_M) := \left\langle (e^{t\mathring{\mathcal{A}}_N}\widetilde{\mathrm{J}}^\delta_i)\mathrm{J}^\gamma_{l_j^-} \,\middle|\, \mathbf{y}_M \right\rangle_{f^N_s}.$$

This, non-local in time, equation is version of the Zwanzig/Nakajima master equation [99, 124] and is equivalent to the original Liouville equation for f^N_t.

Approximations to the Coarse-Grained Evolution Equation

In this section, we discuss various approximations to equation (4.45), the produce certain exact coarse-grained approximations to the evolution of $\hat{f}^M_t$.

To begin, we assume a hyperbolic scaling for the dynamics, that is, we consider the dynamics on times of order Kt. Under such a rescaling, equation (4.45) becomes

$$\begin{aligned} \partial_t \hat{f}^M_t - \widehat{\mathcal{A}}^*_M \hat{f}^M_t - K \sum_{\delta=0}^{2} \sum_{i\in\mathbb{Z}_N} D^\delta_i \bigg(\int_0^t \mathcal{B}^\delta_j[K(t-s), f^N_s] \hat{f}^M_s \mathrm{d}s \bigg) \\ - \sum_{\delta,\gamma=0}^{2} \sum_{i,j\in\mathbb{Z}_M} D^\delta_i D^\gamma_j \bigg(\int_0^t \mathcal{K}^{\gamma,\delta}_{i,j}[K(t-s), f^N_s] \hat{f}^M_s \mathrm{d}s \bigg) = 0. \end{aligned} \tag{4.46}$$

Relaxation approximation

Our first approximation will assume some level of scale separation on the dynamics, specifically, for large enough N we will assume that we can replace the conditional measures with respect to f^N_t conditional measures with respect to an equilibrium one,

$$\langle \cdot \,|\, \mathbf{y}_M \rangle_{f^N_t} \approx \langle \cdot \,|\, \mathbf{y}_M \rangle_N. \tag{4.47}$$

This approximation, which we refer to as the *relaxation approximation* is essentially a kind of ergodic hypothesis on f^N_t when N is large. Of course, for the deterministic evolution generated by $\mathring{\mathcal{A}}_N$, this is a difficult open problem. However, if one introduces a stochastic perturbation to the dynamics of the form discussed in Section 4.3, then the corresponding dynamics is ergodic (see [14] for a proof of this in context of anharmonic chains) and an approximation of the type (4.47) is more likely within reach.

At the level of equation (4.46) the relaxation approximation amounts to making the following approximations

$$\mathcal{B}_j^\delta[t, f_s^N](\mathbf{y}_M) \approx \mathcal{B}_j^\delta[t, \mu_\alpha^N](\mathbf{y}_M) \equiv \widehat{\mathcal{B}}_j^\delta(t, \mathbf{y}_M) = \left\langle \overline{\mathcal{A}}_N e^{t\mathring{\mathcal{A}}_N} \widetilde{\mathrm{J}}_i^\delta \,\middle|\, \mathbf{y}_M \right\rangle_N$$

and

$$\mathcal{K}_{i,j}^{\gamma,\delta}[t, f_s^N](\mathbf{y}_M) \approx \mathcal{K}_{i,j}^{\gamma,\delta}[t, \mu_\alpha^N](\mathbf{y}_M) \equiv \widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}(t, \mathbf{y}_M) = \left\langle (e^{t\mathring{\mathcal{A}}_N} \widetilde{\mathrm{J}}_i^\delta) \widetilde{\mathrm{J}}_j^\gamma \,\middle|\, \mathbf{y}_M \right\rangle_N.$$

Note that we have replaced $\mathrm{J}_{l_j^-}^\gamma$ with $\widetilde{\mathrm{J}}_j^\gamma$ in the definition of $\widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}$, since $e^{t\mathring{\mathcal{A}}_N}\widetilde{\mathrm{J}}_i^\delta$, is $\langle \cdot \,|\, \mathbf{y}_M \rangle_N$ mean 0 and therefore we may freely subtract $\widehat{\mathrm{J}}_j^\gamma = \left\langle \mathrm{J}_{l_j^-}^\gamma \,|\, \mathbf{y}_M \right\rangle_N$ from $\mathrm{J}_{l_j^-}^\gamma$ in the definition of $\widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}$.

Remark 4.8.5. It is important to remark that this approximation *does not* depend on the value of α in the grand-canonical measure μ_α^N, since the above quantities only depend on μ_α^N through its conditional measure $\mu_N(\,\cdot\,|\,\mathbf{y}_M)$.

While the matrix $\widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}(t)$ has explicit time dependence, it no longer depends on the microscopic solution f_t^N and can computed by solving a problem about current-current correlations of *periodized* evolution inside each cell under the micro-canonical measure. Such an approximation allows for computation of $\widehat{\mathcal{B}}_i^\delta$ and $\widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}$ in terms of small number coefficients. Firstly, $\widehat{\mathcal{B}}_i^\delta$ can be computed in terms of $\widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}$.

Lemma 4.8.6. *Let $g_K(y)$ be the density of states, and denote*

$$g_K^M(\mathbf{y}_M) = \prod_{i=1}^{M} g_K(y_i),$$

then the following formula holds

$$\widehat{\mathcal{B}}_i^{\delta} g_K^M = -K^{-1} \sum_{\delta=0}^{2} \sum_{j=1}^{M} D_j^{\gamma} \left(g_K^M \widehat{\mathcal{K}}_{i,j}^{\delta,\gamma} \right)$$

Proof. Let φ be a test function on Γ^M, then by definition of the condition measure $\mu_N(\,\cdot\,|\,\mathbf{y}_M)$, we can make the following computation,

$$\begin{aligned}
\int_{\Gamma^M} \varphi\, \widehat{\mathcal{B}}_i^{\delta}(t)\, g_K^M \mathrm{d}\mathbf{y}_M &= \int_{\Omega^N} \varphi \circ \widehat{\mathbf{w}}\, \overline{\mathcal{A}}_N \left(e^{t\mathring{\mathcal{A}}_N} \widetilde{\mathrm{J}}_i^{\delta} \right) \mathrm{d}\mathbf{z}_N \\
&= -\int_{\Omega^N} \overline{\mathcal{A}}_N(\varphi \circ \widehat{\mathbf{w}})\, e^{t\mathring{\mathcal{A}}_N} \widetilde{\mathrm{J}}_i^{\delta} \mathrm{d}\mathbf{z}_N \\
&= K^{-1} \sum_{\gamma=0}^{2} \sum_{j=1}^{M} \int_{\Omega^N} \mathrm{J}_{l_i^-}^{\gamma}\, D_i^{\gamma} \varphi \circ \widehat{\mathbf{w}} \left(e^{t\mathring{\mathcal{A}}_N} \widetilde{\mathrm{J}}_i^{\delta} \right) \mathrm{d}\mathbf{z}_N \\
&= -K^{-1} \sum_{\gamma=0}^{2} \sum_{j=1}^{M} \int_{\Gamma^M} D_i^{\gamma} \varphi\, \widehat{\mathcal{K}}_{i,j}^{\delta,\gamma}(t)\, g_K^M\, \mathrm{d}\mathbf{y}_M
\end{aligned}$$

This completes the proof. □

Lemma 4.8.6 implies that one only needs to compute $\widehat{\mathcal{K}}_{i,j}^{\delta,\gamma}$ since $\widehat{\mathcal{B}}_i^{\delta}$ can be computed explicitly in terms of $\widehat{\mathcal{K}}_{i,j}^{\delta,\gamma}$. In fact, using Lemma 4.8.6, we easily obtain

$$\begin{aligned}
&K \sum_{\delta=0}^{2} \sum_{i \in \mathbb{Z}_N} D_i^{\delta} \left(\int_0^t \widehat{\mathcal{B}}_j^{\delta}(K(t-s)) \hat{f}_s^M \mathrm{d}s \right) \\
&\qquad + \sum_{\delta,\gamma=0}^{2} \sum_{i,j \in \mathbb{Z}_M} D_i^{\delta} D_j^{\gamma} \left(\int_0^t \widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}(K(t-s)) \hat{f}_s^M \mathrm{d}s \right) \\
&= \sum_{\delta,\gamma=0}^{2} \sum_{i,j \in \mathbb{Z}_M} \int_0^t D_i^{\delta} \left(g_K^M\, \widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}(K(t-s)) D_j^{\gamma} \left(\frac{\hat{f}_s^M}{g_K^M} \right) \right) \mathrm{d}s.
\end{aligned}$$

It is not hard to see that the term on the right-hand side of the above identity vanishes when $\hat{f}_t^M$ is equal to the density of $\hat{\mu}_{K,\alpha}^M$. Therefore, as expected, the relaxation approximation is consistent with the equilibrium results of Section 4.5.2. After making the relaxation approximation, equation (4.46) becomes

$$\partial_t \hat{f}_t^M - \widehat{\mathcal{A}}_M^* \hat{f}_t^M = \sum_{\delta,\gamma=0}^{2} \sum_{i,j \in \mathbb{Z}_M} \int_0^t D_i^\delta \left(g_K^M \, \widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}(K(t-s)) D_j^\gamma \left(\frac{\hat{f}_s^M}{g_K^M} \right) \right) \mathrm{d}s. \quad (4.48)$$

As mentioned, the values $\mathcal{K}_{i,j}^{\gamma,\delta}$ can be computed explicitly in terms of much smaller number of terms. Indeed, using the skew symmetry of $\mathring{\mathcal{A}}_M$ with respect to $\langle \cdot \,|\, \mathbf{y}_M \rangle_N$, we have the following time-reversal relation

$$\widehat{\mathcal{K}}_{i,j}^{\gamma,\delta}(t) = \widehat{\mathcal{K}}_{j,i}^{\delta,\gamma}(-t). \quad (4.49)$$

Furthermore, since current $\widetilde{\mathrm{J}}_i^\delta$ lies on lower boundary values of the cell Λ_i, it can at most depend on values in Λ_{i-1} and Λ_i. Furthermore, since $e^{t\mathring{\mathcal{A}}_N}$ only evolves each particle within the cell it starts in and preserves the mean zero property of the fluxes, we necessarily have,

$$\widehat{\mathcal{K}}_{i+k,i}^{\gamma,\delta} = 0 \quad \text{if} \quad k \geq 2.$$

Moreover, by the symmetry relation (4.49), for each $i \in \mathbb{Z}_M$ it suffices to compute only $\mathcal{K}_{i,i}^{\gamma,\delta}$ and $\mathcal{K}_{i+1,i}^{\gamma,\delta}$, for each γ and δ, since $\mathcal{K}_{i,i+1}^{\gamma,\delta}$ can be computed from $\mathcal{K}_{i+1,i}^{\gamma,\delta}$. However, because of other symmetries of current-current correlation, the number of independent coefficients of $\mathcal{K}_{i,i}^{\gamma,\delta}$ and $\mathcal{K}_{i+1,i}^{\gamma,\delta}$ can be reduced further.

Specifically, for each $y \in \Gamma$ and $t > 0$ define the following correlation functions

$$\theta(t,y) = \frac{1}{T_K(y)} \left\langle \frac{1}{K} \sum_{j=1}^{K} \left(e^{t\mathcal{A}_K} v_j - p_i \right) \left(v_j - p_i \right) \Big| \, y \right\rangle_K$$

$$\zeta(t,y) = \frac{1}{T_K(y)} \left\langle \frac{1}{K} \sum_{j=1}^{K} \left(e^{t\mathcal{A}_K} v_j - p_i \right) \left(V'(r_j) + P_K(y_i) \right) \Big| \, y \right\rangle_K$$

$$\eta(t,y) = \frac{1}{T_K(y)} \left\langle \frac{1}{K} \sum_{j=1}^{K} \left(e^{t\mathcal{A}_K}(V'(r_j) + P_K(y_i)) \right) \left(V'(r_j) + P_K(y_i) \right) \middle| \, y \right\rangle_K ,$$

where $\langle \cdot \mid y \rangle_K$ denote the micro-canonical measure on Ω^K. The correlation functions θ and eta are auto correlation functions of for the volume and momentum fluxes inside a cell. We would also like to define the auto correlation function for the energy flux. However, since $\widetilde{\mathrm{J}}_i^2$ is evaluated on the boundary of a cell, it contains values in two different cells, therefore the auto-correlation function is naturally defined over two adjacent cells Hence we define the function

$$\kappa(t,y_1,y_2) = \frac{1}{T_K(y_1)T_K(y_2)} \left\langle \left(e^{t(\mathring{\mathcal{A}}_{\Lambda_1} + \mathring{\mathcal{A}}_{\Lambda_2})} \widetilde{\mathrm{J}}_1^2 \right) \widetilde{\mathrm{J}}_1^2 \, \middle| \, y_1, y_2 \right\rangle_K^{\otimes 2}$$

It is easy to see from the definition of $\widehat{\mathcal{K}}_{ij}^{\gamma,\delta}$ that

$$\widehat{\mathcal{K}}_{i,i}^{0,0}(t,\mathbf{y}_M) = T_K(y_{i-1})\theta(t,y_{i-1}),$$

$$\widehat{\mathcal{K}}_{i+1,i}^{0,1}(t,\mathbf{y}_M) = T_K(y_i)\zeta(t,y_i),$$

$$\widehat{\mathcal{K}}_{i,i}^{1,1}(t,\mathbf{y}_M) = T_K(y_i)\eta(t,y_i),$$

$$\widehat{\mathcal{K}}_{i,i}^{2,2}(t,\mathbf{y}_M) = T_K(y_{i-1})T_K(y_i)\kappa(t,y_{i-1},y_i).$$

As it turns out, because of the sharp division between cells, $\kappa(t,y_1,y_2)$ can be determined directly in terms of $\theta(t,y_1)$ and $\eta(t,y_2)$.

Lemma 4.8.7. *The following formula holds*

$$\kappa(t,y_1,y_2) = \theta(t,y_1)\eta(t,y_2) + \theta(t,y_1)\beta_K(y_2)P_K(y_2)^2.$$

Using time-reversal symmetry of the Hamiltonian evolution $e^{t\mathring{\mathcal{A}}_M}$ we find

Lemma 4.8.8. *The following time symmetry relations hold*

$$\theta(t,y) = \theta(-t,y), \quad \zeta(t,y) = -\zeta(-t,y), \quad \eta(t,y) = \eta(-t,y). \tag{4.50}$$

Proof. Consider the velocity inversion transformation

$$T(\mathbf{x}, \mathbf{v}) = (\mathbf{x}, -\mathbf{v}),$$

It is east to verify that the Hamiltonian flow $\Phi_t(\mathbf{x}, \mathbf{v})$ associated to $\mathring{\mathcal{A}}_M$ has the following time-symmetry property

$$T(\Phi_t(T(\mathbf{x}, \mathbf{v}))) = \Phi_{-t}(\mathbf{x}, \mathbf{v}),$$

and that $\langle \phi \circ T \,|\, \mathbf{y}_M \rangle = \langle \phi \,\big|\, \mathbf{y}_M \rangle$. This readily implies

$$\Big\langle (e^{-t\mathring{\mathcal{A}}_K} v_j) v_j \,|\, y \Big\rangle_K = \Big\langle (e^{t\mathring{\mathcal{A}}_K} v_j) v_j \,\big|\, y \Big\rangle_K,$$

$$\Big\langle (e^{-t\mathring{\mathcal{A}}_K} V'(r_i)) V'(r_i) \,|\, y \Big\rangle_K = \Big\langle (e^{t\mathring{\mathcal{A}}_K} V'(r_i)) V'(r_i) \,\big|\, y \Big\rangle_K,$$

$$\Big\langle (e^{-t\mathring{\mathcal{A}}_K} v_i) V'(r_i) \,|\, y \Big\rangle_K = -\Big\langle (e^{t\mathring{\mathcal{A}}_K} v_i) V'(r_i) \,|\, y \Big\rangle_K.$$

Using these identities along with the fact that

$$\Big\langle (e^{t\mathring{\mathcal{A}}_K} \widetilde{\mathrm{J}}_i^{\delta}) \widetilde{\mathrm{J}}_j^{\gamma} \,|\, y \Big\rangle_K = \Big\langle (e^{t\mathring{\mathcal{A}}_K} \mathrm{J}_i^{\delta}) \mathrm{J}_j^{\gamma} \,|\, y \Big\rangle_K - \big\langle \mathrm{J}_i^{\delta} \,|\, y \big\rangle_K \big\langle \mathrm{J}_j^{\gamma} \,|\, y \big\rangle_K$$

gives the time-symmetry relations (4.50). □

As it turns out, the entire correlation matrix $\mathcal{K}_{i,j}^{\gamma,\delta}$ can be computed in terms of θ, ζ and η.

Lemma 4.8.9. *The matrices* $(\widehat{\mathcal{K}}_{i,i})_{\gamma,\delta}$ *and* $(\widehat{\mathcal{K}}_{i+1,i})_{\gamma,\delta}$ *take the form*

$$\widehat{\mathcal{K}}_{i,i}(t) = \begin{pmatrix} T_{K,i-1}\theta_{i-1}(t) & 0 & 0 \\ 0 & T_{K,i}\eta_i(t) & p_{i-1}T_{K,i}\eta_i(t) \\ 0 & p_{i-1}T_{K,i}\eta_i(t) & T_{K,i}T_{K,i-1}\kappa_{i,i-1}(t) + T_{K,i}p_{i-1}^2\eta_i(t) \end{pmatrix},$$

and

$$\widehat{\mathcal{K}}_{i+1,i}(t) = \begin{pmatrix} 0 & T_{K,i-1}\zeta_i(t) & T_{K,i-1}\zeta_i(t)p_{i-1} \\ 0 & 0 & 0 \\ 0 & -T_{K,i-1}\zeta_i(t)P_{K,i+1} & 0 \end{pmatrix},$$

where $\theta_i, \zeta_i, \eta_i$ *and* $P_{K,i}$ *denote* $\theta(y_i)$, $\zeta(y_i)$, $\eta(y_i)$ *and* $P_K(y_i)$ *respectively.*

Note that, by the symmetry relations (4.49) and (4.50), we find

$$\widehat{\mathcal{K}}_{i,i+1}(t) = \widehat{\mathcal{K}}_{i+1,i}(-t)^{\top} = \begin{pmatrix} 0 & 0 & 0 \\ -T_{K,i-1}\zeta_i(t) & 0 & T_{K,i-1}\zeta_i(t)P_{K,i+1} \\ -T_{K,i-1}\zeta_i(t)p_{i-1} & 0 & 0 \end{pmatrix},$$

.

The coefficient $\eta_i(t)$ is the momentum-current-current correlation function within a cell. Such correlations are related to the emergence of bulk viscosity of the macroscopic dynamics.

The coefficient $\theta_i(t)$ is precisely the velocity auto-correlation function of a tagged particle in equilibrium evolving in the cell Λ_{i-1}. The emergence of the coefficients $\theta_i(t)$ and $\zeta_i(t)$ is related to the fact that our coarse-graining map has a sharp division between cells, and hence the current between cells is dictated by the value of the current on the boundary of the cells. Moreover since the conditional measure $\langle \cdot \,|\, \mathbf{y}_M \rangle_{f_t^N}$ has no permutation symmetry properties within cells, the boundary current cannot be replaced with a suitable summation of values in a cell. As a consequence, in the derivation of equation (4.45), one is restricted to looking at current-current correlations between currents on the cell boundaries. The result of this is the appearance of $\theta_i(t)$ and $\zeta_i(t)$, which, in some sense, encode non-trivial correlations in the currents that connect the cells.

The coefficient $\kappa_{i,i-1}(t)$ is the energy-current-current correlation functions across two cells and is related to the emergence of thermal conductivity between cells. It is rather remarkable in this case that it can be explicitly described in terms of θ_{i-1}, η_i and thermodynamic quantities $\beta_{K,i}$ and $P_{K,i}$.

Markov Approximation and Decay of Correlations

While the relaxation approximation is useful for studying the behavior of fluctuations around equilibrium, the utility of equation (4.48) is limited due to the non-local in time nature of the equation. Such an evolution equation gives rise to non-Markovian features of the dynamics with $\widehat{\mathcal{K}}K^{\gamma,\delta}_{ij}(t)$ playing the role of a memory kernel. In particular, this implies that it is precisely the persistence of two-time current-current correlations within cells that gives rise to memory effects. Indeed, if the size of the cell K (and consequently our choice of time-scale) is not too large, then non-Markovian effects in the coarse-grained dynamics is not entirely surprising, since the periodized dynamics inside the cells have not had enough time to forget their initial data. However, when K is large, one expects there to be a certain *decay of correlations* for large times, namely, when $t > 0$ we expect

$$\widehat{\mathcal{K}}^{\gamma,\delta}_{i,j}(Kt) = \left\langle (e^{Kt\mathring{\mathcal{A}}_M}\widetilde{\mathrm{J}}^{\delta}_i)\widetilde{\mathrm{J}}^{\gamma}_j \,\middle|\, \mathbf{y}_M \right\rangle_N \to 0, \tag{4.51}$$

as $K \to \infty$. Of course, for the deterministic evolution generated by $\mathring{\mathcal{A}}_M$, proving such decay is a very difficult mathematical problem and is likely out of the reach of current mathematical tools. Moreover, if there are other, apriori unknown, conserved quantities in the dynamics associated to $\mathring{\mathcal{A}}_M$, then, for instance the correlations might settle into a constant, non-zero, state (this is seen for instance in [98]). However, if one replaces the motion inside the cells with a stochastic component of the type introduced in Section 4.3, then one can likely obtain explicit (mixing) estimates on the decay of correlations (4.51), and ensure that (r_i, v_i, e_i) are the only locally conserved quantities.

Decay of correlations directly implies a loss of memory in the dynamics governed by (4.48). Indeed, if the decay is fast enough, we may localize the non-local nature of (4.48), such a localization in time is often referred to as a *Markov approximation* since the resulting evolution equation will be that of a Markov process, particularly a diffusion process. Specifically, treating $K\widehat{\mathcal{K}}^{\gamma,\delta}_{i,j}(Kt)$ as an *approximation of the identity* we write

$$\int_0^t \widehat{\mathcal{K}}^{\gamma,\delta}_{i,j}(K(t-s))\hat{f}^M_s \mathrm{d}s \approx K^{-1}\left(\int_0^K \widehat{\mathcal{K}}^{\gamma,\delta}_{i,j}(s)\mathrm{d}s\right)\hat{f}^M_t, \tag{4.52}$$

which is expected to hold when K is large. Note that we have truncated the time-integral above at $t = K$. The reason for this is that, in one dimension, the correlation matrix $\widehat{\mathcal{K}}^{\gamma,\delta}_{i,j}(t)$ typically decays to 0, but has "long-time tails" which are not integrable on $\mathbb{R}_+$. This effect was first noticed numerically by Alder and Wainwright [1, 2] for the velocity auto-correlation function $\theta_i(t)$, where it was observed that $\theta_i(t)$ decayed like $t^{-1/2}$. Indeed, this would imply that there is a divergence of the form

$$\int_0^K \theta_i(s)\mathrm{d}s \sim K^{1/2} \tag{4.53}$$

as $K \to \infty$. Note that, while $\int_0^K \theta_i(s)\mathrm{d}s$ diverges the quantity $K^{-1}\int_0^K \theta_i(s)\mathrm{d}s$ still vanishes for large K implying that the contribution to equation (4.48) is still small.

It is instructive to note that at time $t = 0$ we have

$$\theta_i(0, y_i) = \left\langle \frac{1}{K} \sum_{j \in \Lambda_{i-1}} (v_j - p_i)^2 \,\middle|\, \mathbf{y}_M \right\rangle_N ,$$

and by the equivalence of ensembles Theorem A.3.7, as $K \to \infty$, we have

$$\theta_i(0, \ell_i, p_i, e_i) \to T(\ell_i, e_i - \tfrac{1}{2}p_i^2),$$

where $T(\ell, e) = \beta(\ell, e)^{-1}$ is the thermodynamic temperature associated with the grand-canonical ensemble. Therefore $\theta_i(t))$ starts at a non-zero value and is expected to decay in time. However, for transient times, the decay will generally not be monotonic as might be suggested. Indeed, $\theta_i(t)$ may become negative and undergo oscillations on its approach to 0, further increasing the potential rate of divergence suggested in (4.53).

Also at time $t = 0$, we have

$$\zeta_i(0) = \left\langle \frac{1}{K} \sum_{j \in \Lambda_i} \Big(v_j - p_i\Big)\Big(V'(r_j) + P_K(y_i)\Big) \,\middle|\, \mathbf{y}_M \right\rangle_N = 0,$$

due to the fact that the measure $\langle \cdot \,|\, \mathbf{y}_M \rangle$ is symmetric with respect to permutation in the velocity and deformation indices separately, and therefore we may replace $v_j - p_i$ above with $\frac{1}{K}\sum_{j \in \Lambda_i} v_j - p_i$, which is equal to 0 on the micro-canonical surface. Therefore, contrary to θ_i we expect the grow from time zero, and oscillate with decreasing amplitude as $t \to \infty$. In light of this, we will assume that these oscillations average out over time and therefore

$$\int_0^K \zeta_i(s) \mathrm{d}s \sim 0,$$

as $K \to \infty$. As a result, we will typically neglect the contribution due to $\zeta_i(s)$ in the Markov approximation.

Consequently, we define the time integrated correlation functions

$$\bar{\theta}(y_i) := \int_0^K \theta(t, y_i)\mathrm{d}t, \quad \bar{\eta}(y_i) := \int_0^K \eta(t, y_i)\mathrm{d}t,$$

and, upon neglecting the time integral of $\zeta_i(t)$ for large enough K, we conclude

$$\int_0^K \mathcal{K}_{i,j}^{\gamma,\delta}(t, \mathbf{y}_M)\mathrm{d}t \approx \mathrm{d}^{\gamma,\delta}(y_{i-1}, y_i)\delta_{i,j}, \tag{4.54}$$

where $\mathrm{d}^{\gamma,\delta}(y_{i-1}, y_i) = (\mathrm{d}(y_{i-1}, y_i))_{\gamma,\delta}$ is the diffusion matrix defined by

$$\mathrm{d}(y_{i-1}, y_i) =$$

$$\begin{pmatrix} T_K(y_{i-1})\bar{\theta}(y_{i-1}) & 0 & 0 \\ 0 & T_K(y_i)\bar{\eta}(y_i) & p_{i-1}T_K(y_i)\bar{\eta}(y_i) \\ 0 & p_{i-1}T_K(y_i)\bar{\eta}(y_i) & T_K(y_{i-1})T_K(y_i)\bar{\kappa}(y_{i-1}, y_i) + T_K(y_i)\bar{\eta}(y_i)p_{i-1}^2 \end{pmatrix}$$

where

$$\bar{\kappa}(y_{i-1}, y_i) = \bar{\theta}(y_{i-1})\bar{\eta}(y_i) + \bar{\theta}(y_{i-1})\beta_K(y_i)P_K(y_i)^2.$$

The functions $\bar{\theta}(y)$, $\bar{\eta}(y)$ are given by time integrals of auto-correlation functions analogous to the *Green-Kubo* formula. It therefore follows from a standard application of the Wiener-Kinchin theorem that, for large enough K, we have

$$\bar{\theta}, \bar{\eta}, \bar{\kappa} \geq 0.$$

Applying the approximation (4.52) and (4.54) to equation (4.48), we obtain

$$\partial_t \hat{f}_t^M - \widehat{\mathcal{A}}_M^* \hat{f}_t^M = K^{-1} \sum_{\delta,\gamma=0}^{2} \sum_{i\in\mathbb{Z}_M} D_i^\delta \left(g_K^M \, \mathrm{d}_{i-1,i}^{\gamma,\delta} D_i^\gamma \left(\frac{\hat{f}_t^M}{g_K^M} \right) \right). \tag{4.55}$$

Equation (4.55) is the forward Kolmogorov equation for a diffusion process with generator

$$\begin{aligned}\widehat{\mathcal{L}}_M\phi &= \widehat{\mathcal{A}}_M\phi + K^{-1}\sum_{\delta,\gamma=0}^{2}\sum_{i\in\mathbb{Z}_M}\frac{1}{g_K^M}D_i^\gamma\left(g_K^M \mathrm{d}_i^{\gamma,\delta}D_i^\delta\phi\right)\\ &= \widehat{\mathcal{A}}_M\phi + K^{-1}\sum_{\delta=0}^{2}\sum_{i\in\mathbb{Z}_M}\mathcal{J}_{i,i-1}^\delta D_i^\delta\phi + K^{-1}\sum_{\gamma,\delta=0}^{2}\sum_{i\in\mathbb{Z}_M}\mathrm{d}_{i,i-1}^{\gamma,\delta}D_I^\gamma D_i^\delta\phi,\end{aligned} \tag{4.56}$$

where

$$\mathcal{J}_{i,i-1}^\delta = \sum_{\gamma=0}^{2}\mathrm{d}_i^{\gamma,\delta}D_i^\gamma\log g_K^M + D_i^\gamma(\mathrm{d}_i^{\gamma,\delta}),$$

is the dissipative flux.

The SDE system associated with the generator (4.56) is

$$\dot{y}_i = \widehat{\mathrm{J}}_{i+1} - \widehat{\mathrm{J}}_i + K^{-1}(\mathcal{J}_{i+1,i} - \mathcal{J}_{i,i-1}) + K^{-1/2}(\dot{\mathcal{M}}_{i+1,i} - \dot{\mathcal{M}}_{i,i-1}) \tag{4.57}$$

where $\widehat{\mathrm{J}}_i = (p_{i-1}, -P_i, -p_{i-1}P_i)$ are the discrete Euler currents, and $\mathcal{M}_{i-1,i} = (\mathcal{M}_{i-1,i}^\ell, \mathcal{M}_{i-1,i}^p, \mathcal{M}_{i-1,i}^e)$ is a vector of mean-zero martingales defined by stochastic integration against a collection of independent Wiener processes $\{W_i^\ell\}, \{W_i^p\}, \{W_i^e\}$

$$\begin{aligned}\dot{\mathcal{M}}_{i,i-1}^\ell &= \sqrt{2T_{i-1}\bar{\theta}_{i-1}}\dot{W}_i^\ell\\ \dot{\mathcal{M}}_{i,i-1}^p &= \sqrt{2T_i\bar{\eta}_i}\dot{W}_i^p\\ \dot{\mathcal{M}}_{i,i-1}^e &= u_{i-1}\dot{\mathcal{M}}_i^p + \sqrt{2\bar{\kappa}_{i,i-1}T_iT_{i-1}}\dot{W}_i^e.\end{aligned} \tag{4.58}$$

With a bit of work, the dissipative fluxes $\mathcal{J}_i = (\mathcal{J}_{i,i-1}^\ell, \mathcal{J}_{i,i-1}^p, \mathcal{J}_{i,i-1}^e)$ can be shown to be given by

$$\begin{aligned}\mathcal{J}_{i,i-1}^\ell &= T_{i-1}\bar{\theta}_{i-1}(\beta_iP_i - \beta_{i-1}P_{i-1}) + \beta_i\partial_\ell\bar{\theta}_{i-1} + \bar{\theta}_{i-1}\partial_\ell\bar{\eta}_i - P_i\partial_{e_i}\bar{\kappa}_{i-1,i}\\ \mathcal{J}_{i,i-1}^p &= (\bar{\eta}_i + T_i\partial_e\bar{\eta}_i)(p_i - p_{i-1})\\ \mathcal{J}_{i,i-1}^e &= p_{i-1}\mathcal{J}_{i,i-1}^p + T_i\bar{\eta}_i + \bar{\kappa}_{i,i-1}(T_i - T_{i-1}) - T_iT_{i-1}(\partial_{e_i}\bar{\kappa}_{i,i-1} - \partial_{e_{i-1}}\bar{\kappa}_{i,i-1}),\end{aligned}$$

In equations (4.7) and (4.8) we used subscripts to denote dependence on a certain coarse particle, for instance $P_i = P(y_i)$, and $\bar{\kappa}_{i-1,i} = \bar{\kappa}(y_{i-1}, y_i)$.

A Simplified Fluid-Particle Model

We now aim to introduce a simplified version of the model (4.57) in the case that the transport coefficients $\bar{\theta}_i, \bar{\eta}_i, \bar{\kappa}_{i-1,i}$ are constant and the dissipation in the volume term is gone. Such a model taken the form

$$\begin{aligned}
\dot{\ell}_i &= (p_i - p_{i-1}) \\
\dot{p}_i &= (P_i - P_{i+1}) + [\bar{\eta}(p_{i+1} - p_i) - \bar{\eta}(p_i - p_{i-1})] + \dot{\mathcal{M}}^p_{i+1} - \dot{\mathcal{M}}^p_i \\
\dot{e}_i &= (p_{i-1}P_i - p_i P_{i+1}) + \bar{\eta}\,[p_i(p_{i+1} - p_i) - p_{i-1}(p_i - p_{i-1})] \\
&\qquad + \bar{\kappa}\,[(T_{i+1} - T_i) - (T_i - T_{i-1})] + \bar{\eta}(T_i - T_{i-1}) \\
&\qquad + p_i\dot{\mathcal{M}}^p_{i+1} - p_{i-1}\dot{\mathcal{M}}^p_i + \dot{\mathcal{M}}^e_{i+1} - \dot{\mathcal{M}}^e_i.
\end{aligned}$$

where $(\mathcal{M}^p_{i-1,i}, \mathcal{M}^e_{i-1,i})$ are defined in 4.58. As in our discussion of the discrete Euler system in Section 4.6 we will find it useful to introduce the internal energy variables $u_i = e_i - \frac{1}{2}p_i^2$, which transform the equations to

$$\begin{aligned}
\dot{\ell}_i &= (p_i - p_{i-1}) \\
\dot{p}_i &= (P_i - P_{i+1}) + [\bar{\eta}(p_{i+1} - p_i) - \bar{\eta}(p_i - p_{i-1})] + \dot{\mathcal{M}}^p_{i+1} - \dot{\mathcal{M}}^p_i \\
\dot{u}_i &= -(p_i - p_{i-1})P_i + \bar{\eta}(p_i - p_{i-1})^2 + \bar{\kappa}\,[(T_{i+1} - T_i) - (T_i - T_{i-1})] - 2\bar{\eta}T_i \\
&\qquad + (p_i - p_{i-1})\dot{\mathcal{M}}^p_i + \dot{\mathcal{M}}^e_{i+1} - \dot{\mathcal{M}}^e_i.
\end{aligned} \tag{4.59}$$

The generator of equation (4.59) is given by

$$\mathcal{L} = \mathcal{L}_{\mathsf{H}} + \mathcal{L}_{\mathsf{S}},$$

where

$$\mathcal{L}_{\mathsf{H}} = \sum_{i \in \mathbb{Z}_N} -p_{i-1}(\partial_{\ell_i} - \partial_{\ell_{i-1}}) + P_i\mathcal{X}_i$$

and

$$\mathcal{L}_{\mathsf{S}} = \sum_{i\in T_N} -\eta(p_i - p_{i-1})\mathcal{X}_i - \kappa(T_i - T_{i-1})\mathcal{Y}_i + \eta T_i \mathcal{X}_i^2 + T_i T_{i-1}\mathcal{Y}_i^2.$$

where $\{\mathcal{X}_i : i \in T_N\}$ and $\{\mathcal{Y}_i : i \in T_N\}$ are two families of differential operators representing vector fields tangent to certain manifolds defining pairwise momentum and energy exchange. They are given by

$$\mathcal{X}_i = \partial_{p_i} - \partial_{p_{i-1}} - (p_i - p_{i-1})\partial_{u_i}, \quad \mathcal{Y}_i = \partial_{u_i} - \partial_{u_{i-1}}.$$

The constants $\eta, \kappa > 0$ play the role of bulk-viscosity and thermal-conductivity in the model. The functions $\mathsf{L} = \sum_{i\in T_N} \ell_i$, $\mathsf{P} = \sum_{i\in T_N} p_i$, $\mathsf{H} = \sum_{i\in T_N} \frac{1}{2}p_i^2 + u_i$, corresponding to total length, momentum, and energy, are in the null space of $\mathcal{L}$, and therefore conserved by the dynamics.

Contrary to the Euler discretization, the entropy S_N is *not* conserved by the stochastic dynamics, as is to be expected it is a discrete model of the Navier-Stokes-Fourier system. Instead the entropy satisfies a discrete version of the Gibbs-Duhem relation. Indeed when computing the evolution of the entropy $S_i = S(\ell_i, u_i)$, Itô's formula implies,

$$\begin{aligned} \mathrm{d}S_i &= \partial_\ell S(\ell_i, u_i)\mathrm{d}\ell_i + \partial_u S(\ell_i, u_i)\mathrm{d}u_i \\ &\quad + \partial_u^2 S(\ell_i, u_i)\left[(p_i - p_{i-1})^2 \eta T_i + \kappa T_{i+1} T_i + \kappa T_i T_{i-1}\right]\mathrm{d}t. \end{aligned}$$

Using the thermodynamic relations $\partial_\ell S = \beta P$, and $\partial_u S = \beta$, and the evolution equations, we obtain,

$$\begin{aligned} \mathrm{d}S_i &= T_i\left(\partial_u^2 S_i + \beta_i^2\right)\left(\eta(p_i - p_{i-1})^2 + [\kappa T_{i+1} + \kappa T_{i-1}]\right)\mathrm{d}t \\ &\quad - (\kappa + \kappa)\mathrm{d}t - 2\eta\mathrm{d}t + \beta_i\Big[(p_i - p_{i-1})\mathrm{d}\mathcal{M}_i^p + \mathrm{d}\mathcal{M}_{i+1}^e - \mathrm{d}\mathcal{M}_i^e\Big]. \end{aligned}$$

As we can see, the total entropy $\mathsf{S} = \sum_{i\in T_N} S_i$ is not strictly dissipated as, as one might expect being a discrete version of Navier-Stokes. This barrier to dissipation is due to the noise (the same behavior is observed, for instance, in stochastic gradient dynamics). In general this can lead to problems of well posedness for the fluid-particle model, i.e. finite time blow up in the form of parcel volumes or energies collapsing to 0. However, certain assumptions on the concavity on $S(\ell, e)$ allow one to obtain enough dissipation of S to show existence and uniqueness of a process which stays in the interior of Γ^N. The main result of this section is the following theorem

Theorem 4.9.1. *Suppose that the entropy function $S(\ell, u)$ approaches $-\infty$ when either u or ℓ approach 0, grows sub-linearly when either u or ℓ approach ∞ and satisfies the lower bound*

$$\partial_u^2 S(\ell, u) \geq= (1-\gamma)T(\ell, u)^{-2} \tag{4.60}$$

for some $\gamma \in (0,1)$. Let $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t\geq 0}, W)$, where $W = \{(W_i^p, W_i^e)\}_{i\in\mathbb{Z}_N}$ is a family of independent one-dimensional Brownian motions relative to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Then for any N and z_0 in the interior of Γ^N, the SDE associated to $\mathcal{L}$ has a unique $(\mathcal{F}_t)$ measurable solution $z(t) = \{(\ell_i, p_i, u_i)(t)\}_{i\in\mathbb{Z}_N}$ which remains in the interior of Γ^N for all $t \geq 0$ and has continuous sample paths.

Remark 4.9.2. Note that the case when S is the entropy of a *one-dimensional monatomic ideal gas* assumption (4.60) is *not* satisfied. In fact, in this case we have,

$$\partial_u^2 S_i + \beta_i^2 = -\beta_i^2.$$

It appears that this condition is a size condition on the specific heat (at constant volume) associated to the entropy S. There are the negative contributions due to κ_i and η_i, which can hurt the entropy dissipation as well. Indeed it will be necessary to have control on the size of κ_i and η_i

Well-posedness

In this section, we prove Theorem 4.9.1. Note that this requires showing that the process $z(t) = \{(\ell_i, p_i, u_i)(t)\}_{i\in\mathbb{Z}_N}$ remains in the interior $\Gamma_0^N := \operatorname{Int}\Gamma^N$ for all time. This implies that if for each $i \in \mathbb{Z}_N$, (ℓ_i, u_i) start positive, then $(\ell_i, u_i)(t)$ remain strictly positive for all later times with probability one. As a consequence, since ℓ_i denotes the difference between particles q_i and q_{i-1}, if the particles start ordered on $\mathbb{Z}_N$, they will remained ordered on $\mathbb{Z}_N$ with probability one, that is the particles *cannot* pass through each other. We will find it useful to simplify notation and write the SDE (4.59) in the following standard Itô form

$$\mathrm{d}z = b(z)\,\mathrm{d}t + \sigma(z)\,\mathrm{d}w, \quad z(0) = z_0, \tag{4.61}$$

where $z(t)$ denotes the process $\{(\ell_i, p_i, u_i)(t)\}_{i\in\mathbb{Z}_N}$, represented as a vector in Γ^N, with $z_i = (\ell_i, p_i, u_i)$, and $w(t)$ is an $(\mathbb{R}^3)^N$ valued Brownian motion. Let $\mathcal{L}_z$ denote the generator $\mathcal{L}$ with coefficients evaluated at $z \in \Gamma^N$. The drift $b(z)$, $b : \Gamma^N \to (\mathbb{R}^3)^N$ can be defined by

$$b(z) = \mathcal{L}_z\, z$$

and the matrix $\sigma(z)$, $\sigma : \Gamma^N \to (\mathbb{R}^3)^N \otimes (\mathbb{R}^2)^N$ satisfies $\sigma^2 = a$, where

$$a(z) = \mathcal{L}_z\,(z \otimes z) - z \otimes (\mathcal{L}_z\, z) - (\mathcal{L}_z\, z) \otimes z.$$

We are now ready to prove the Theorem

Proof of Theorem 4.9.1. . Note that the functions b and a are not globally Lipschitz on Γ^N. Indeed they have singularities in as $\ell_i \to 0$ and grow quadratically as $p_i \to \infty$. However they are *locally Lipschitz* in the sense that for any compact set K contained in $\operatorname{Int}\Gamma^N$ then b and a are Lipschitz on K.

To prove existence up to a possible explosion time $\tau = \inf\{t : z(t) \notin \operatorname{Int}\Gamma^N\}$, we will define a function $F : \Gamma^N \to \mathbb{R}_+$ by

$$F(z) = \mathsf{H}_N(z) + \mathsf{L}_N(z) - \mathsf{S}_N(z) + C,$$

where C is an undetermined constant. As a consequence of hypothesis 4.6.2, F is a C^2 convex function on Γ^N and approaches ∞ as $z \to \partial\Gamma^N$ and as $|z| \to \infty$. Therefore F has a minimum value on Γ^N and the constant C may be chosen so that $F \geq 0$ on Γ^N. For each $R \geq 0$, define the following family of compact sets, strictly contained in Γ^N,

$$K_R = \{z \in \Gamma^N \, : \, F(z) \leq R\},$$

and let φ_R be a smooth cutoff function equal to 1 on K_R and equal to zero outside of K_{R+1}. Let $b_R(z) = \varphi_R(z)b(z)$ and $\sigma_R(z) = \varphi_R(z)\sigma(z)$ be the corresponding cutoff coefficients. Indeed b_R and σ_R are globally Lipschitz on Γ^N. Therefore by a standard Banach fixed point argument on $L^2(\Omega; C([0,\infty), \Gamma^N))$, there exists a unique pathwise solution to the following SDE

$$\mathrm{d}z_R = b_R(z)\mathrm{d}t + \sigma_R(z)\mathrm{d}W, \quad z(0) = z_{0,R} \in K_R.$$

Since $b(z) = b_R(z)$ and $\sigma(z) = \sigma_R(z)$ on K_R, the process $(z_R(t))_{t=0}^{\infty}$ is a solution

$(z(t))_{t=0}^{\infty}$ to (4.61) up to the stopping time

$$\tau_R = \inf\{t : z(t) \notin K_R\}.$$

In fact, this solution $\{z_R(t)\}_{t=0}^{T}$ is the *unique* solution to (4.61) with initial data $z_{0,R}$ on the interval $[0, \tau_R)$. Since the sets $\{K_R\}_{R\geq 0}$ increase as $\mathbb{R} \to \infty$ and $\bigcup_{R\geq 0} K_R = \operatorname{Int}\Gamma^N$, the stopping times $\{\tau_R\}_{R\geq 0}$ are increasing. Therefore, by uniqueness, if $R_2 \geq R_1$, then $z_{R_1}(t) = z_{R_2}(t)$ on $[0, \tau_{R_1})$. Now, let

$$\tau = \sup_{R\geq 0} \tau_R,$$

and for any $z_0 \in \operatorname{Int}\Gamma^N$, choose $R_0 \geq 0$ such that $z_0 \in K_R$ if $R \geq R_0$. We then construct the unique solution $z(t)$ on $[0, \tau)$ to (4.61) with initial data z_0, by

$$z(t) = z_{R_0}(t)\mathbb{1}_{[0,\tau_{R_0})}(t) + \sum_{j=0}^{\infty} z_{R_0+j+1}(t)\mathbb{1}_{[\tau_{R_0+j},\tau_{R_0+j+1})}(t).$$

To show well-posedness, we simply need to show non-explosion,

$$\mathbf{P}\{\tau = \infty\} = 1.$$

To do this, we will use a Lyapunov function method with the function $F(z)$. Indeed, since H_N and L_N are conserved, we find,

$$\begin{aligned} \mathcal{L}\,F(z) &= -\mathcal{L}\,\mathsf{S}_N(z) \\ &= \sum_{i\in\mathbb{Z}_N} \Big(2\kappa + 2\eta - T_i(\partial_e^2 S_i + \beta_i^2)\big[\eta(p_i - p_{i-1})^2 + \kappa T_{i+1} + \kappa T_{i-1}\big]\Big). \end{aligned}$$

Under the assumption that

$$\partial_e^2 S(\ell, e) \geq -\beta(\ell, e)^2,$$

we can show that

$$\mathcal{L}F(z) \leq C_N$$

for a constant C_N depending on N and the transport coefficients η, κ. Define $V(t,z) = e^{C_N t}F(z)$, then by Itô's formula the process $(M_V(t))_{t=0}^{\infty}$ defined by

$$M_V(t) = V(t, z(t)) - V(0, z_0) - \int_0^t \big(\partial_s V(s, z(s)) + \mathcal{L}V(s, z(s))\big)\mathrm{d}s$$

is a martingale. Using the fact that

$$\partial_t V(t,z) + \mathcal{L}V(t,z) \leq 0,$$

and $V(t,z) \geq 0$ we may conclude that for each $R > 0$ the stopped process $z(t \wedge \tau_R)$ satisfies for each t,

$$\begin{aligned}
V(0, z_0) &\geq \mathbf{E}[V(t, z(t \wedge \tau_R)] \\
&= \mathbf{P}\{\tau_R \geq t\}\mathbf{E}[V(t, z(t))] + \mathbf{P}\{\tau \leq t\}\mathbf{E}[V(t, z(\tau_R))] \\
&\geq \mathbf{P}\{\tau_R \leq t\}R.
\end{aligned}$$

Therefore we conclude that for all $R > 0$ and $t \geq 0$,

$$\mathbf{P}\{\tau \leq t\} \leq \mathbf{P}\{\tau_R \leq t\} = V(0, z_0)R^{-1}.$$

Sending $R \to \infty$ concludes the non-explosion condition. Therefore the solution $z(t)$ constructed above is the unique solution to the SDE (4.61). $\square$

Part II

Stochastic Transport

Elegance should be left to shoemakers and tailors.

Ludwig Boltzmann

Introduction to Part II

Stochastic Transport Equations

The study of stochastic differential equations (SDEs)

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t \tag{5.1}$$

with rough drift b and diffusion σ have received a lot of attention in recent years. In many applications in fluid mechanics (and kinetic theory) one is interested in solving (5.1) when b and σ are not Lipschitz (rough). The problem of existence of probabilistically strong, pathwise unique solutions to (5.1) when b and σ are rough have been studied in a number of works, some of the earlier work is by Krylov and Veretennikov [117, 118], Krylov and Röckner [79] and more recently by Champagnat and Jabin [25] and Rezakhanlou [107].

One approach to this problem is to study existence and uniqueness of strong solutions to the associated stochastic transport equation

$$\begin{aligned} &\partial_t f + \operatorname{div}(bf) - \operatorname{div}\operatorname{div}(af) + \operatorname{div}(\sigma f)\cdot \dot{W} = 0, \\ &f|_{t=0} = f_0, \end{aligned} \tag{5.2}$$

where $a = \frac{1}{2}\sigma\sigma^\top$. When a is a multiple of the identity, and b is rough, this problem was studied by Flandoli, Gubinelli and Priola [50, 51], as well as by [22, 100].

The hope is to generalize the DiPerna/Lions theory for the deterministic transport equation [34] to one for the stochastic transport equation (5.2). When σ is rough and degenerate, a version of the DiPerna/Lions theory for the associated Kolmogorov equation has been developed by Figalli [49] and by Lions/Le Bris [84]. However, there appear to be few results in the literature concerning solutions to the stochastic transport (5.2).

In Chapter 6 the theory of renormalized solutions for (5.2) when σ is rough is developed. We employed the usual commutator estimates used in [34], along with a new *double commutator* that arises due to the stochastic term. Interestingly, using this method it only seems possible to obtain uniqueness for solutions in L^p for $p > 2$, when $\sigma \in W^{1,2p/p-2}$ and $\operatorname{div} a \in W^{1,p/p-2}$. The existence and uniqueness of (probabilistically) strong solutions in L^p for $p \in [1, 2)$ when σ is rough appears to be rather non-trivial. This is consistent with the work of Lions/ LeBris [84].

Stochastically Forced Boltzmann Equation

Many models of turbulence involve forcing the equations of fluid mechanics by noise. From a physical perspective, this can be viewed as some kind of environmental shaking inciting the onset of turbulence. A natural question to ask is whether this noise can be deduced from a more general form of noise at the kinetic level. Of course, conditions for the well-posedness of such stochastic kinetic equations are of interest, as well as whether such noise may provide insights into the behavior of a turbulent fluid at the kinetic level.

In a collaboration with another student Scott Smith we initiated a study of the Boltzmann equation with stochastic forcing,

$$\begin{aligned} &\partial_t f + v \cdot \nabla_x f + \mathrm{div}_v (f \circ \dot{\mathcal{X}}) = B(f,f), \\ &f|_{t=0} = f_0. \end{aligned} \tag{5.3}$$

The forcing $\dot{\mathcal{X}}$ is a Gaussian noise, white in time, and colored in $(x, v) \in \mathbb{R}^n \times \mathbb{R}^n$, of the general form

$$\dot{\mathcal{X}}(t,x,v) = \sum_{k \in \mathbb{N}} \sigma_k(x,v) \dot{\beta}_k(t)$$

where $\{\sigma_k : k \in \mathbb{N}\}$ are a family of deterministic $\mathbb{R}^n$ valued vector fields over $\mathbb{R}^n \times \mathbb{R}^n$, and $\{\beta_k : k \in \mathbb{N}\}$ are independent one-dimensional Brownian motions. The product between f and $\dot{\mathcal{X}}$ is interpreted in the Stratanovich sense.

Such an equation describes the evolution of the one-particle phase space density $f(t, x, v)$ of a rarefied gas subject to *elastic binary collisions* and *environmental noise*. The elastic binary collisions are modeled by Boltzmann collision operator $f \mapsto B(f, f)$, a quadratic operator that acts pointwise (t, x) and non-locally in v. The environmental noise acts on the gas *externally* in the sense that each particle is driven by the same realization of the noise $\dot{\mathcal{X}}$. This is in contrast to *intrinsic noise* where each particle in the gas is driven by an independent realization of the noise. The environmental noise is modeled by stochastic transport on the left side of equation (5.3). Indeed, in the absence of collisions one may think of the particle in that gas as following certain stochastic characteristics (X_t, V_t) that solve the

Stratanovich SDE

$$\mathrm{d}X_t = V_t\,\mathrm{d}t, \qquad\qquad X_0 = x \in \mathbb{R}^n$$

$$\mathrm{d}V_t = \sum_{k\in\mathbb{N}} \sigma_k(X_t, V_t) \circ \mathrm{d}\beta_t, \quad V_0 = v \in \mathbb{R}^n.$$

We are interested in the existence of solutions to (5.3). With regards to existence of solutions to the Navier-Stokes equations driven by white noise, one of the first rigorous studies was undertaken by Bensoussan and Temam [13] and has since received much attention in the mathematical literature (a relatively recent survey of the many results is given in [39]).

In [106], we study the existence of global in time solutions to (5.3) for a general class of 'large' initial data in $L^1(\mathbb{R}^n \times \mathbb{R}^n)$ with certain entropy and moment bounds. In the deterministic setting, such a result was proven by DiPerna/Lions [36] for the Boltzmann equation in the *renormalized sense*, and improved in subsequent works [35, 36, 88, 89]. Our main result is a proof of the existence of, global in time, *probabilistically weak* (in the sense of a solution to the martingale problems) solutions to (5.3) in the renormalized sense (the same notion of solution used in [36]). The main theorem is stated informally as follows:

Theorem 5.2.1. *Let f_0 have finite mass, energy and entropy,*

$$\|(1 + |x|^2 + |v|^2 + |\log f|)f\|_{L^1_{x,v}} < \infty$$

and suppose that the coefficients $\{\sigma_k : k \in \mathbb{N}\}$, $\operatorname{div}_v \sigma_k = 0$, satisfy certain regularity and summability conditions. Then for a certain class of collision operators $B(f, f)$, there exists a probabilistically weak (martingale) solution $\{f_t : t \geq 0\}$ to (5.3) satisfied in the renormalized sense.

The process $\{f_t : t \geq 0\}$ *takes values in the cone of non-negative* $L^1(\mathbb{R}^n \times \mathbb{R}^n)$ *functions and has bounded p-th moments of mass, energy, entropy, and entropy dissipation,*

$$\mathbf{E}\|(1 + |x|^2 + |v|^2 + |\log f|)f\|^p_{L^\infty_t(L^1_{x,v})} \leq \infty, \quad \mathbf{E}\|D(f)\|^p_{L^1_{t,x}} < \infty,$$

for each $p \in [1, \infty)$*, where the entropy dissipation is defined by*

$$D(f) := -\int_{\mathbb{R}^d} (\log f)B(f, f)\mathrm{d}v.$$

Moreover, $\{f_t : t \geq 0\}$ *has a continuous modification with paths in* $C([0, T]; L^1(\mathbb{R}^n \times \mathbb{R}^n))$*.*

The proof of Theorem 5.2.1 largely inspired by techniques layed out in [36], and more specifically on the later work by Lions [89] on the Vlasov-Maxwell-Boltzmann equation. In the deterministic case, one of the key elements of the proof is the strong compactness obtained velocity averages of solutions to the transport equation [64, 66, 67]. In our paper we prove a *stochastic velocity averaging* result in L^1 which shows, under certain conditions, that a family of solutions $\{f_n : n \in \mathbb{N}\}$ to a stochastic kinetic transport equation has the property that the laws of the velocity averages are tight on $L^1_{t,x}$. This result should be compared with other stochastic velocity averaging results in the literature [48, 87].

Outline of Part II

Part II of the dissertation will be organized as follows:

In Chapter 6, we discuss stochastic transport in L^p equations with rough diffusion coefficients. We introduce a theory of renormalized solutions to such equations and deduce regularity conditions on the noise coefficients which imply pathwise uniqueness.

Chapter 7 is a joint work by the author and his collaborator, Scott Smith, concerning the Boltzmann equation with stochastic transport, modeling the influence of a random environmental forcing. We study the properties of stochastic transport equations and prove a renormalization and stochastic velocity averaging result. We prove existence of renormalized martingale solutions for a general class of noise coefficients and bounded collision kernel, using a generalization of the Skorohod theorem for non-metric spaces. We also obtain local conservation of mass, average global balance of momentum, and average global dissipation of energy and entropy for these solutions.

Renormalized Solutions to Stochastic Transport

We are primarily interested in the transport equation associated to the Itô stochastic differential equation

$$\mathrm{d}X_t = u(t, X_t)\mathrm{d}t + \sum_k \sigma^k(t, X_t)\mathrm{d}W_t^k$$

$$X_0 = x \in \mathbb{R}^n,$$

where u is the drift $u : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$, $\{\sigma^k\}$ are noise coefficients $\sigma^k : [0,T]\times\mathbb{R}^n \to \mathbb{R}^n$ and $\{W_t^k\}$ are independent one-dimensional Wiener processes.

Specifically we are interested in the associated stochastic transport equation

$$\partial_t f + \operatorname{div}(uf) - \operatorname{div}\operatorname{div}(af) + \sum_k \operatorname{div}(\sigma^k f)\dot{W}^k = 0 \qquad (6.1)$$

$$f|_{t=0} = f_0$$

where $a = \frac{1}{2}\sum_k \sigma^k \otimes \sigma^k$ is the diffusion matrix.

Existence

We begin by studying the existence problem for the stochastic transport equation (6.1). Fix a cannonical stochastic basis $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbf{P}, \{W^k\})$ and let $\Phi_{s,t}$ be the stochastic flow associated with the SDE

$$\mathrm{d}\Phi_{s,t}(x) = u(t, \Phi_{s,t}(x))\mathrm{d}t + \sum_k \sigma^k(t, \Phi_{s,t}(x))\mathrm{d}W_t^k, \quad \Phi_{s,s}(x) = x. \qquad (6.2)$$

We assume u and $\{\sigma^k\}$ are smooth enough with sub-linear growth so that $\Phi_{s,t}$ is a diffeomorphism and adapted to $\mathscr{F}_{s,t} = \sigma(\{W_t^k - W_s^k\} : 0 \le s \le t \le \infty)$, and its spatial inverse $\Psi_{s,t}(x) = \Phi_{s,t}^{-1}(x)$ is also $(\mathscr{F}_{s,t})$-adapted (pontwise in x). We will also denote $\Phi_t = \Phi_{0,t}$ and $\Psi_t = \Psi_{0,t}$. Suppose that we start from some smooth f_0, and that $g = 0$. Then we know that the unique solutions to the transport equation is given by

$$f(t,x) = f_0(\Psi_t(x)) \det \partial\Psi_t(x)$$

where $(\partial\Psi_t)_{ij} = \partial_j(\Psi_t)_i$. We have the following proposition regarding a formula for $\det \partial\Psi_t(x)$.

Proposition 6.1.1. *The quantity* $\det \partial\Psi_t(x)$ *can be written as,*

$$\det \partial\Psi_t(x) = \exp\Bigg\{ - \int_0^t \Big[\operatorname{div} u(\Psi_{s,t}(x)) - \tfrac{1}{2}\sum_k \operatorname{tr}\big((\partial\sigma^k)^2\big)(\Psi_{s,t}(x))\Big] \mathrm{d}s \\ - \int_0^t \sum_k \operatorname{div} \sigma^k(\Psi_{s,t}(x)) \mathrm{d}W_r^k \Bigg\}.$$

Proof. To study $\det \partial\Psi_t(x)$ further, we remark that it suffices to study $\det \partial\Phi_t(x)$, since we have $\partial\Phi_t(\Psi(x))\partial\Psi_t(x) = I$ and therefore

$$\det \partial\Psi_t(x) = [\det \partial\Phi_t(\Psi_t(x))]^{-1}.$$

The taking the derivative of both sides of the SDE with respect to the initial data, it is well known that the matrix $\partial\Phi_t$ satisfies

$$\mathrm{d}\partial\Phi_t = \partial u(t,\Phi_t)\partial\Phi_t \mathrm{d}t + \sum_k \partial\sigma^k(t,\Phi_t)\partial\Phi_t \mathrm{d}W_t^k,$$

To study the determinant of $\partial\Phi_t$, we use the fact that for any invertible matrix A, the Gateaux derivative of $F(A) := \log(\det A)$ in the direction U is

$$DF(A)\big[U\big] = \operatorname{tr}\big(UA^{-1}\big)$$

while the second order Gateaux derivative in the directions U, V is

$$D^2F(A)\big[U,V\big] = -\operatorname{tr}\big(UA^{-1}VA^{-1}\big).$$

Applying Itô's formula to quantity $F(\partial\Phi_t) = \log(\det \partial\Phi_t(x))$, and using the above formulas, we find

$$\begin{aligned}
\mathrm{d}F(\partial\Phi_t) &= DF(\partial\Phi_t)\big[\partial u(t,\Phi_t)\partial\Phi_t\big]\mathrm{d}t + \sum_k DF(\partial\Phi_t)\big[\partial\sigma^k(t,\Phi_t)\partial\Phi_t\big]\mathrm{d}W_t^k \\
&\qquad + \tfrac{1}{2}\sum_k D^2F(\partial\Phi_t)\Big[\partial\sigma^k(t,\Phi_t)\partial\Phi_t, \partial\sigma^k(t,\Phi_t)\partial\Phi_t\Big]\mathrm{d}t \\
&= \operatorname{tr}\big(\partial u\big)(t,\Phi_t(x))\mathrm{d}t + \sum_k \operatorname{tr}\big(\partial\sigma^k\big)(t,\Phi_t)\mathrm{d}W_t^k - \sum_k \frac{1}{2}\operatorname{tr}\big((\partial\sigma^k)^2\big)(t,\Phi_t)\mathrm{d}t \\
&= \operatorname{div} u(t,\Phi_t)\mathrm{d}t + \sum_k \operatorname{div}\sigma^k(t,\Phi_t)\mathrm{d}W_t^k - \frac{1}{2}\sum_k \operatorname{tr}\big((\partial\sigma^k)^2\big)(t,\Phi_t)\mathrm{d}t
\end{aligned}$$

Using the fact that $F(\partial\Phi_0) = 0$ concluded the proof. □

We now try to get L^p estimates on the solution $f(t,x)$. We have the following

Proposition 6.1.2. *Assume that f_0, u an $\{\sigma^k\}$ are smooth and compactly supported and let $f(t,x)$ be the unique classical solution to the transport equation. For each $p \in [1,\infty)$, we have the following inequality*

$$\mathbf{E}\int_0^T \int_{\mathbb{R}^n} |f(t,x)|^p \mathrm{d}x\mathrm{d}t \le C_{p-1,u,\sigma}\|f_0\|_{L_x^p}^p \tag{6.3}$$

where for each $q \in (0,\infty)$ the constant $C_{p-1,u,\sigma}$ is defined by

$$C_{q,u,\sigma} := \exp\left\{q\|\operatorname{div} u\|_{L_t^1(L_x^\infty)} + \sum_k \frac{1}{2}q\|\operatorname{tr}\big((\partial\sigma^k)^2\big)\|_{L_t^1(L_x^\infty)} + \sum_k \frac{1}{2}q^2\|\operatorname{div}\sigma^k\|_{L_t^2(L_x^\infty)}^2\right\},$$

Proof. To esimtate this, we note that

$$\int_{\mathbb{R}^n} |f(t,x)|^p \mathrm{d}x = \int_{\mathbb{R}^n} |f_0(x)|^p |\det \partial\Psi_t(\Phi_t(x))|^{p-1}\mathrm{d}x$$

From the formula for $\det \partial\Psi_t(x)$ it readily follows that for any $q > 0$ and $x \in \mathbb{R}^n$

$$|\det \partial\Psi_t(\Phi_t(x))|^q \leq C_{q,u,\sigma}\, \mathcal{E}\left(-q\sum_k \int_0^t \operatorname{div}\sigma^k(\Phi_s(x))\mathrm{d}W_s^k\right),$$

where $C_{q,u,\sigma}$ is the contant defined in the statement of the proposition and $\mathcal{E}(X_t)$ denotes the Dooleans exponential of a martingale X_t, specifically in our case

$$\mathcal{E}\left(-q\sum_k \int_0^t \operatorname{div}\sigma^k(\Phi_s(x))\mathrm{d}W_s^k\right) = \exp\left\{-\int_0^t \sum_k q \operatorname{div}\sigma^k(\Phi_s(x))\mathrm{d}W_s^k - \frac{1}{2}q^2\sum_k \int_0^t |\operatorname{div}\sigma^k|^2(\Phi_s(x))\mathrm{d}s\right\}.$$

Using the fact that $\mathcal{E}(X_t)$ is again a martingale and therefore $\mathbf{E}\mathcal{E}(X_t) = \mathbf{E}\mathcal{E}(X_0)$ concludes the proof. □

Our definition of solution is as follows

Definition 6.1.3. Let $p \in [1,\infty]$, $q = p/(p-1)$ ($q = 1$ if $p = \infty$). Suppose for each compact $K \subseteq \mathbb{R}^n$, $u \in L^1([0,T], L^q(K))$, $\sigma = \{\sigma^k\} \in L^2([0,T], L^{2q}(K))$ and $(\Omega, \mathscr{F}, \mathbf{P}, (\mathscr{F}_t), \{W^k\})$ and stochastic basis. A weak L^p solution to the stochastic continuity equation is an $(\mathscr{F}_t)$ progessively measurable process $f : \Omega\times[0,T] \to L^p_{x,\mathrm{loc}}$ which almost surely solves the stochastic transport equation in weak, time-integrated form. That is, for every $\varphi \in C_c^\infty(\mathbb{R}^n)$ and $\mathbf{P}\otimes \mathrm{d}t$ almost every (t,ω) we have

$$\langle f_t, \varphi\rangle = \langle f_0, \varphi\rangle + \int_0^t \langle f_s, \mathcal{L}\varphi\rangle \mathrm{d}s + \sum_k \int_0^t \langle f_s, \sigma^k\cdot\nabla\varphi\rangle \mathrm{d}W_s^k,$$

where $\mathcal{L} = u\cdot\nabla + a : \nabla^2$ is the generator of the diffusion (6.2).

Our main existence theorem is the following

Theorem 6.1.4. *Let $p \in [2,\infty]$, $q = p/(p-1)$ ($q = 1$ if $p = \infty$) and assume that $f_0 \in L^p$ and $u \in L^1_t(L^q_{x,\mathrm{loc}})$ and $\sigma = \{\sigma^k\} \in L^2_t(L^{2q}_{x,\mathrm{loc}})$. If, in addition, $\operatorname{div}\sigma \in L^2_t(L^\infty_x)$ and $\operatorname{tr}((\partial\sigma)^2) \in L^1_t(L^\infty_x)$, then for any stochastic basis*

$(\Omega, \mathscr{F}, \mathbf{P}, (\mathscr{F}_t), \{W^k\})$, *there exists a weak* L^p *solution to the stochastic continuity equation and* $f \in L^p(\Omega \times [0,T] \times \mathbb{R}^n)$.

Proof. The proof is straight forward. We first approximate u, σ, f_0 by smooth functions $(u)_n$, $(\sigma)_n$, $(f_0)_n$ which satisfy the uniform bound

$$\sup_n C_{p-1,(u)_n,(\sigma)_n} < \infty,$$

such that

$$(u)_n \to u \text{ in } L^1_t(L^q), \quad (\sigma)_n \to \sigma \text{ in } \ell^2(\mathbb{N}; L^2_t(L^q)), \quad (f_0)_n \to f_0 \in L^p. \tag{6.4}$$

Let f_n be the unique classical solution to the stochastic transport equation associated to $(u)_n$, $(\sigma)_n$, $(f_0)_n$ (see [81]). We remark that the smoothness (in x) of f_n implies that f_n is progessively measurable as a process with values in L^p. Using the estimate in proposition 6.1.2 we may conclude that $\{f_n\}$ is uniformly bounded in $L^2(\Omega \times [0,T]; L^p)$. Therefore $\{f_n\}$ has a weakly converging subsequence in $L^2(\Omega\times[0,T]; L^p)$, which we still denote $\{f_n\}$. Moreover since the space of progressively measurable processes in $L^2(\Omega\times[0,T]; L^p)$ is closed, it follows that the limit f is also progressively measurable.

We now wish to pass the limit in the weak form. Let $Y \in L^2(\Omega)$ and $\phi \in C^\infty_c([0,T] \times \mathbb{R}^n)$, then for each $n \geq 0$ we have

$$\mathbf{E}\int_0^T \int_{\mathbb{R}^n} Y(\partial_t\phi + (\mathcal{L})_n\phi) f_n \mathrm{d}x\mathrm{d}t + \sum_k \mathbf{E}\int_0^T \int_{\mathbb{R}^n} Y((\sigma^k)_n \cdot \nabla\phi) f_n \mathrm{d}x\mathrm{d}W^k_t = 0, \tag{6.5}$$

where $(\mathcal{L})_n = (u)_n \cdot \nabla + (a)_n : \nabla^2$ and $(a)_n = \frac{1}{2}\sum_k (\sigma^k)^{\otimes 2}_n$. Clearly the weak convergence of $\{f_n\}$ in $L^2(\Omega \times [0,T]; L^p)$ and strong convergence of $Y(\partial_t\phi + (\mathcal{L})_n\phi)$

in $L^2(\Omega\times[0,T];L^q)$ is enough to pass the limit in the first integral in equation (6.5), which follow from the convergence properties (6.4). What remains is to pass the limit in the stochastic integral. Clearly we have

$$\int_{\mathbb{R}^n}(\sigma^k)_n\cdot\nabla\phi\, f_n \mathrm{d}x \to \int_{\mathbb{R}^n}\sigma^k\cdot\nabla\phi\, f\mathrm{d}x \quad \text{weakly in } L^2(\Omega\times[0,T]),$$

and since the stochastic integral is a weakly continuous linear mapping from $L^2(\Omega\times[0,T])$ to $L^2(\Omega)$, we may pass the limit term by term in the summation for the stochastic integral. If the summation is infinite then we use the fact that

$$\left|\mathbf{E}\int_0^T\int_{\mathbb{R}^n}Y((\sigma^k)_n\cdot\nabla\phi)f_n\mathrm{d}x\mathrm{d}W_t^k\right| \le \sup_n\{C_{p-1,(u)_n,(\sigma)_n}\}\|Y\|_{L^2(\Omega)}\|(\sigma^k)\|^2_{L^2_t(L^q)}\|f_0\|^2_{L^p}$$

and that $\sigma\in\ell^2(\mathbb{N};L^2_t(L^q))$ to pass the limit in the sum.

Finally, to obtain the almost sure, time integrated form, we remark that $\mathbf{P}$ almost surely,

$$\int_0^T\int_{\mathbb{R}^n}(\partial_t\phi+\mathcal{L}\phi)f\mathrm{d}x\mathrm{d}t+\sum_k\int_0^T\int_{\mathbb{R}^n}(\sigma^k\cdot\nabla\phi)f\mathrm{d}x\mathrm{d}W_t^k=0. \tag{6.6}$$

Now fix a $t\in[0,T]$ and choose a sequence of test functions $\phi^n(s,x)=\varphi(x)\psi^n(s)$, where $\psi^n(s)$ is a smooth approximation of the indicator $\mathbb{1}_{[0,t]}(s)$ so that $\partial_s\psi^n(s)$ is a symmetric approximation of a delta function centered at t. Using the integrability of f in time, Lebesgue's differentiation theorem implies that for almost every $t\in[0,T]$,

$$\int_0^T\int_{\mathbb{R}^n}\partial_s\phi^n(s,x)f(s,x)\mathrm{d}x\mathrm{d}s\to\langle f_t,\varphi\rangle.$$

Passing the limit in (6.6), for test function $\phi=\phi^n$, gives the time-integrated weak form, $\mathbf{P}$ almost surely. □

Remark 6.1.5. The statement of the extistence theorem can be somewhat improved. In fact, we will see that $f \in L^{\infty-}(\Omega; L^\infty_t(L^p_x))$, and f has a modification in $C_t([L^p_x]_{\mathrm{w}})$. While it might be possible to get this directly from the estimates on the flow and the Dooleans exponential, it will be more straight forward to work directly with the solution f_n to the approximating scheme presented above.

Renormalization

We now study the renormalization property for the stochastic continuity equation. For simplicity, we will study the following stochastic continuity equation with zero drift and one noise coefficient σ,

$$\partial_t f - \operatorname{div}\operatorname{div}(af) + \operatorname{div}(\sigma f)\dot{W} = 0, \tag{6.7}$$

where $a = \frac{1}{2}\sigma \otimes \sigma$. The extension to the more general case of non-zero (Sobolev regular) drift and countably many noise coefficients being straight forward, following the classical arguments of Diperna-Lions [34].

Let us asssume for the moment that σ is smooth and that f is a smooth classical solution to (6.7), that is, f is at least C^2 in x, is pointwise adapted to $(\mathscr{F}_t)$ and satisfies (6.7) in the time time integrated sense, pointwise in $\mathbb{R}^n$. Let $\Gamma : \mathbb{R} \to \mathbb{R}$ be a smooth function, then we will show that $\Gamma(f)$ satisfies a stochastic continuity equation of the form

$$\begin{aligned}\partial_t\Gamma(f) - \operatorname{div}\operatorname{div}(a\Gamma(f)) + \operatorname{div}(\sigma^k\Gamma(f))\dot{W} &= \operatorname{div}\big(\sigma(\operatorname{div}\sigma)G(f)\big) - G(f)\operatorname{div}\sigma\dot{W} \\ &\quad + \tfrac{1}{2}G(f)\operatorname{tr}\big((\partial\sigma)^2\big) + \tfrac{1}{2}H(f)(\operatorname{div}\sigma)^2,\end{aligned} \tag{6.8}$$

where $G(f) = f\Gamma'(f) - \Gamma(f)$ and $H(f) = fG'(f) - G(f)$. Such a procedure of solving the equation is called *renormalization* with the equation (6.8) being refered to as the renormalized equation. It's important to note that the above renormalized equation is in divergence form so that in distribution, this equation makes sense without any regularity requirements on f and no more regularity on u and σ than is required for the existence theorem (in fact it needs far less). Such a renormalization readily allows for bounds of the form

$$\begin{aligned}\mathbf{E}\int_{\mathbb{R}^n}\Gamma(f(t))\mathrm{d}x \leq \int_{\mathbb{R}^n}\Gamma(f_0)\mathrm{d}x + \frac{1}{2}\mathbf{E}\int_0^t\int G(f(s))\,\mathrm{tr}\left((\partial\sigma)^2\right)\mathrm{d}x\mathrm{d}s \\ + \frac{1}{2}\mathbf{E}\int_0^t\int_{\mathbb{R}^n} H(f(s))(\mathrm{div}\,\sigma)^2\mathrm{d}x\mathrm{d}s.\end{aligned}$$

In fact, one can do better. If $\Gamma(z) \geq 0$, and $\mathrm{div}\,\sigma \in L^2_t(L^\infty)$ and $\mathrm{tr}\,((\partial\sigma)^2) \in L^1_t(L^\infty)$ the Burkholder-Davis-Gundy inequality implies that for $r \in [1,\infty)$,

$$\begin{aligned}&\mathbf{E}\left(\sup_t\int_{\mathbb{R}^n}\Gamma(f(t))\mathrm{d}x\right)^r \\ &\leq C_{r,T}\left(\|\,\mathrm{div}\,\sigma\|^{2r}_{L^2_t(L^\infty_x)} + \|\,\mathrm{tr}\left((\partial\sigma)^2\right)\|^r_{L^1_t(L^\infty_x)}\right)\mathbf{E}\left(\sup_t\int_{\mathbb{R}^n}|G(f(t))| + |H(f(t))|\mathrm{d}x\right)^r.\end{aligned}$$

This bound, (by choosing bounded approximation of the function $\Gamma(z) = |z|^p$), implies, after an application of Grönwall's inequality, that for each $r \in [1,\infty)$,

$$\mathbf{E}\|f\|^r_{L^\infty_t(L^p_x)} \leq C\|f_0\|^r_{L^p},$$

where C depends continuously on r and T and on σ through the norms $\|\,\mathrm{div}\,\sigma\|_{L^2_t(L^\infty_x)}$ and $\|\,\mathrm{tr}\,((\partial\sigma)^2)\,\|_{L^1_t(L^\infty_x)}$. This bound is clearly and improvement over the one obtained in (6.3), and certainly not so obvious at the level of the stochastic flow and Dooleans exponential.

Derivation of the renormalized form

Here we detail, for convenience, the calculation for the renormalized form given in equation (6.8). Let f be a smooth solution (6.7), which we write in the following form

$$\partial_t f - 2\operatorname{div}(a)\cdot\nabla f - a:\nabla^2 f + \sigma\cdot\nabla f\dot{W} = (\operatorname{div}\operatorname{div} a)f - \operatorname{div}\sigma f\dot{W}$$

Let $\Gamma:\mathbb{R}\to\mathbb{R}$ be a smooth renormalizer, then, using Itôs formula, $\Gamma(f)$ satisfies

$$\begin{aligned}&\partial_t\Gamma(f) - 2\operatorname{div} a\cdot\nabla\Gamma(f) - a:\nabla^2\Gamma(f) + \sigma\cdot\nabla\Gamma(f)\dot{W}\\ &\qquad = f\Gamma'(f)(\operatorname{div}\operatorname{div} a) - \Gamma'(f)\operatorname{div}\sigma f\dot{W} + \tfrac{1}{2}\Gamma''(f)\left(\operatorname{div}(\sigma f)^2 - (\sigma\cdot\nabla f)^2\right).\end{aligned}$$

Writing the left-hand side above back in divergence form and utilitize some cancellation in the term that multiplies $\Gamma''(f)$ we have

$$\begin{aligned}\partial_t\Gamma(f) - \operatorname{div}\operatorname{div}(a\Gamma(f)) + \operatorname{div}(\sigma\Gamma(f))\dot{W} &= G(f)(\operatorname{div}\operatorname{div} a) - G(f)\operatorname{div}\sigma\dot{W}\\ &\quad + \tfrac{1}{2}G'(f)(\operatorname{div}\sigma)^2 f + G'(f)(\operatorname{div}\sigma)\sigma\cdot\nabla f.\end{aligned}$$

The terms on the right-hand side simplify nicely. Using the fact that

$$\operatorname{div}\operatorname{div} a = \tfrac{1}{2}\operatorname{tr}\left((\partial\sigma)^2\right) + \tfrac{1}{2}(\operatorname{div}\sigma)^2 + \sigma\cdot\nabla\operatorname{div}\sigma$$

and

$$G'(f)(\operatorname{div}\sigma)\sigma\cdot\nabla f = \operatorname{div}\left(\sigma(\operatorname{div}\sigma)G(f)\right) - (\operatorname{div}\sigma)^2 G(f) - \sigma\cdot\nabla\operatorname{div}\sigma G(f),$$

we can write the renormalized equation as

$$\begin{aligned}\partial_t\Gamma(f) - \operatorname{div}\operatorname{div}(a\Gamma(f)) + \operatorname{div}(\sigma^k\Gamma(f))\dot{W} &= \operatorname{div}\left(\sigma(\operatorname{div}\sigma)G(f)\right) - G(f)\operatorname{div}\sigma\dot{W}\\ &\quad + \tfrac{1}{2}G(f)\operatorname{tr}\left((\partial\sigma)^2\right) + \tfrac{1}{2}H(f)(\operatorname{div}\sigma)^2,\end{aligned}$$

Renormalization for rough σ

We now want study the renormalization property when f is not smooth. We will follow the strategy from the deterministic theory of DiPerna Lions. This involves regularizing a solution, renormalizing the regularized equation, and then show that the errors commited during this procedure can be written in terms of certain commutators between the differential action of a vector field and the smoothing operation. These commutators will vanish if one assumes the right Sobolev integrability on the vector field.

In what follows, we will find it useful to introduce the differential operators

$$\mathcal{L}_\sigma \phi := a : \nabla^2 \phi, \qquad \nabla_\sigma \phi := \sigma \cdot \nabla \phi$$
$$\mathcal{L}^*_\sigma \phi := \operatorname{div}\operatorname{div}(a\phi), \qquad \nabla^*_\sigma \phi := \operatorname{div}(\sigma\phi),$$

along with the quantities,

$$A_\sigma = \operatorname{tr}\left((\partial\sigma)^2\right), \quad D_\sigma = (\operatorname{div}\sigma)^2.$$

With this notation, the stochastic continuity (6.7) equation takes the form

$$\partial_t f - \mathcal{L}^*_\sigma f + \nabla^*_\sigma f \dot{W} = 0.$$

and the renormalized form (6.8) becomes,

$$\partial_t \Gamma(f) - \mathcal{L}^*_\sigma \Gamma(f) + \nabla^*_\sigma \Gamma(f)\dot{W} = \nabla^*_\sigma(\operatorname{div}\sigma\, G(f)) - \operatorname{div}\sigma\, G(f)\dot{W} + \tfrac{1}{2}G(f)A_\sigma + \tfrac{1}{2}H(f)D_\sigma.$$

We aim to prove the following theorem

Theorem 6.2.1. *Let $f \in L^\infty_t(L^p_{x,\mathrm{loc}})$, $p > 2$, be a weak L^p solution to (6.7), and suppose that $\sigma \in L^2_t(W^{1,2p/(p-2)}_{x,\mathrm{loc}})$. Then for any $\Gamma \in C^2_b(\mathbb{R}^n)$, such that $\sup_z z\Gamma'(z)$*

and $\sup_z z^2\Gamma''(z) < \infty$, $\Gamma(f)$ *solves the renormalized equation in time-integrated, weak form, namely for* $\mathbf{P} \otimes \mathrm{d}t$ *almost every* (t,ω) *we have*

$$\begin{aligned}\langle \Gamma(f(t)), \varphi\rangle = \langle \Gamma(f_0), \varphi\rangle &+ \int_0^t \langle \Gamma(f(s)), \mathcal{L}_\sigma \varphi\rangle \mathrm{d}s + \int_0^t \langle \Gamma(f(s)), \nabla_\sigma \varphi\rangle \mathrm{d}W(s)\\ &- \int_0^t \langle G(f(s)) \operatorname{div}\sigma, \nabla_\sigma\varphi\rangle \mathrm{d}s + \frac{1}{2}\int_0^t \langle G(f(s)) A_\sigma, \varphi\rangle \mathrm{d}s \\ &+ \frac{1}{2}\int_0^t \langle H(f(s)) D_\sigma, \varphi\rangle \mathrm{d}s - \int_0^t \langle G(f(s)) \operatorname{div}\sigma, \varphi\rangle \mathrm{d}W(s)\end{aligned} \tag{6.9}$$

Commutators

As in the deterministic theory, commutators of vector field operations with smoothing play an important role in the renomalization theory. Indeed indentifying the correct commutators is crucial for simplifying certain remainders in an efficient manner.

We start by considering $\eta : \mathbb{R}^n \to \mathbb{R}$ a smooth, symmetric function with support in the ball of radius 1 and with unit integral. For each $\epsilon > 0$ we denote by η_ϵ the rescaled function (mollifier) by

$$\eta_\epsilon(x) = \epsilon^{-n}\eta(\epsilon^{-1}x).$$

We define for any function $\phi : \mathbb{R}^n \to \mathbb{R}$, the mollified function $\phi_\epsilon = (\phi)_\epsilon = \eta_\epsilon \star \phi$ by it's convolution with η_ϵ. Define the following commutators

$$[\nabla_\sigma, \eta_\epsilon](f)(x) = \nabla_\sigma f_\epsilon(x) - (\nabla_\sigma^* f)_\epsilon(x) = \int_{\mathbb{R}^n} \nabla\eta_\epsilon(x-y)\cdot(\sigma(x)-\sigma(y))f(y)\,\mathrm{d}y.$$

and

$$\begin{aligned}\big[[\mathcal{L}_\sigma, \eta_\epsilon]\big](f)(x) &= \mathcal{L}_\sigma f_\epsilon(x) - \nabla_\sigma(\nabla_\sigma^* f)_\epsilon(x) + (\mathcal{L}_\sigma^* f)_\epsilon(x)\\ &= \frac{1}{2}\int_{\mathbb{R}^n} \nabla^2\eta_\epsilon(x-y) : (\sigma(x)-\sigma(y))^{\otimes 2} f(y)\,\mathrm{d}y.\end{aligned}$$

Note that these commutators differ from the typical commutators studied in DiPerna/Lions, since they do not contain any terms involving the divergence of a vector field. However, instead of these commutators vanishing, they will converge precisely to the divergence terms that they excluding. We have the following lemma

Lemma 6.2.2 (Commutator Lemma). *Let $f \in L^p_{x,\mathrm{loc}}$ and $\sigma \in W^{1,q}_{x,\mathrm{loc}}$, for $p, q \in [1,\infty]$. Then as $\epsilon \to 0$*

$$[\nabla_\sigma, \eta_\epsilon](f) \to \operatorname{div}\sigma \quad in \quad L^r_{x,\mathrm{loc}}, \quad for \quad \tfrac{1}{r} = \tfrac{1}{q} + \tfrac{1}{p}$$

and

$$\big[[\mathcal{L}_\sigma, \eta_\epsilon]\big](f) \to \frac{1}{2}(A_\sigma + D_\sigma) \quad in \quad L^r_{x,\mathrm{loc}}, \quad for \quad \tfrac{1}{r} = \tfrac{2}{q} + \tfrac{1}{p}.$$

Moreover for any compact $K \subseteq \mathbb{R}^n$ we have the following bounds

$$\|[\nabla_\sigma, \eta_\epsilon](f)\|_{L^r(K)} \le \|\nabla\sigma\|_{L^q(K)}\|f\|_{L^p(K)}, \quad for \quad \tfrac{1}{r} = \tfrac{1}{q} + \tfrac{1}{p}$$

$$\left\|\big[[\mathcal{L}_\sigma, \eta_\epsilon]\big](f)\right\|_{L^r(K)} \le \|\nabla\sigma\|^2_{L^q(K)}\|f\|_{L^p(K)}, \quad for \quad \tfrac{1}{r} = \tfrac{2}{q} + \tfrac{1}{p}.$$

Proof. We study $[\nabla_\sigma, \eta_\epsilon]$ first. Define for each $x, w \in \mathbb{R}^n$ the quantity

$$R_w(x) := \sigma(x) - \sigma(x-w) - \nabla\sigma(x)\cdot w = \int_0^t (\nabla\sigma(x + (\lambda-1)w) - \nabla\sigma(x))\cdot w \,\mathrm{d}\lambda,$$

so that we can write

$$[\nabla_\sigma, \eta_\epsilon](f)(x) = \int_{\mathbb{R}^n} \nabla\eta_\epsilon(x-y)\cdot(\nabla\sigma(x)\cdot(x-y))f(y)\mathrm{d}y + \int_{\mathbb{R}^n} \nabla\eta_\epsilon(y)R_y(x)f(x-y)\mathrm{d}y$$

Since $\sigma \in W^{1,q}_{x,\mathrm{loc}}$ we have that if $|w| < \epsilon$, then for any compact $K \subseteq \mathbb{R}^n$

$$\|R_w(x)\|_{L^q_x(K)} \le \epsilon \sup_{|y|<\epsilon} \|\delta_y \nabla\sigma\|_{L^q_x(K)},$$

where $\delta_y h(x) = h(x+y) - h(x)$ denotes the difference of for some function h and it's translation by y. Using the above bound, and the fact that $\epsilon \|\nabla \eta_\epsilon\|_{L^1_x}$ is uniformly bounded in ϵ, we find for

$$\begin{aligned}\left\| \int_{\mathbb{R}^n} \nabla \eta_\epsilon(y) R_y(x) f(x-y) \mathrm{d}y \right\|_{L^r_x(K)} &\leq \left(\int_{\mathbb{R}^n} \nabla \eta_\epsilon(y) \, \|R_y\|_{L^q_x(K)} \mathrm{d}y \right) \|f\|_{L^p_x(K)} \\ &\leq \epsilon \|\nabla \eta_\epsilon\|_{L^1_x} \sup_{|y|<\epsilon} \|\delta_y \nabla \sigma\|_{L^q_x(K)} \|f\|_{L^p_x(K)} \\ &\to 0 \quad \text{as} \quad \epsilon \to 0.\end{aligned}$$

for any compact set $K \subset \mathbb{R}^n$. Indeed this implies that for each $x \in \mathbb{R}^n$, and r satisfying $\frac{1}{r} = \frac{1}{q} + \frac{1}{p}$,

$$[\nabla_\sigma, \eta_\epsilon](f)(x) = \nabla \sigma(x) : (G_\epsilon \star f) + o(1)_{L^r_{x,\mathrm{loc}}}$$

where $G_\epsilon(x) = x \otimes \nabla \eta_\epsilon(x)$. This estimate directly implies the bound on $[\nabla_\sigma, \eta_\epsilon](f)$ stated in the lemma. Furthermore, using the fact that each component of $G_\epsilon(x) = \epsilon^{-d} G(\epsilon^{-1} x)$ is a symmetric approximation of a delta function, we can use the standard properties of mollifiers to find

$$G_\epsilon \star f \to \left(\int_{\mathbb{R}^n} x \otimes \nabla \eta(x) \mathrm{d}x \right) f \quad \text{in} \quad L^p_{x,\mathrm{loc}}.$$

Integration by parts, and the properties of η give to identity

$$\int_{\mathbb{R}^n} x \otimes \nabla \eta(x) \mathrm{d}x = I.$$

Therefore as $\epsilon \to 0$, the following convergence holds in $L^r_{x,\mathrm{loc}}$

$$[\nabla_\sigma, \eta_\epsilon](f) \to (\nabla \sigma : I) f = (\operatorname{div} \sigma) f.$$

Next we study the double commutator $\big[[\mathcal{L}_\sigma, \eta_\epsilon]\big](f)$. A similar arguement to the single commutator case implies that since $\sigma \in W^{1,q}_{x,\mathrm{loc}}$, we have that for each

$x \in \mathbb{R}^n$ and r such that $\frac{1}{r} = \frac{2}{q} + \frac{1}{p}$,

$$\big[[\mathcal{L}_\sigma, \eta_\epsilon]\big](f)(x) = \frac{1}{2}\int_{\mathbb{R}^n} \nabla^2\eta_\epsilon(x-y) : (\nabla\sigma(x)\cdot(x-y))^{\otimes 2} f(y)\mathrm{d}y + o(1)_{L^r_{x,\mathrm{loc}}}$$
$$= \frac{1}{2}\sum_{ijk\ell} \partial_i\sigma_k(x)\partial_j\sigma_\ell(x)\big(G^{kl}_{ij,\epsilon}\star f\big) + o(1)_{L^r_{x,\mathrm{loc}}},$$

where $G^{kl}_{ij,\epsilon}(x) = x_i x_j \partial_k\partial_\ell\eta_\epsilon(x)$. Again, this immediately implies the bound on $\big[[\mathcal{L}_\sigma, \eta_\epsilon]\big](f)$. Furthermore, since each $G^{kl}_{ij,\epsilon}(x)$ is a symmetric approximation of a delta function we have

$$G^{kl}_{ij,\epsilon}\star f \to \left(\int_{\mathbb{R}^n} x_i x_j \partial_k\partial_\ell\eta \mathrm{d}x\right) f \quad \text{in} \quad L^p_{x,\mathrm{loc}}.$$

Using the identity,

$$\int_{\mathbb{R}^n} x_i x_j \partial_k\partial_\ell\eta(x)\mathrm{d}x = \delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell},$$

and the fact that $\nabla\sigma \in L^q_{x,\mathrm{loc}}$, then the following convergence holds in $L^r_{x,\mathrm{loc}}$ as $\epsilon \to 0$

$$\big[[\mathcal{L}_\sigma, \eta_\epsilon]\big](f) \to \frac{1}{2}\sum_{ijk\ell}(\delta_{ij}\delta_{k\ell} + \delta_{ik}\delta_{j\ell})\partial_i\sigma_k\partial_j\sigma_\ell f = \frac{1}{2}(A_\sigma + D_\sigma).$$

□

Proof of renormalization result

Proof. As usual we begin we mollify the transport equation. For $\epsilon > 0$ we have,

$$f_\epsilon(t) = f_\epsilon(0) + \int_0^t (\mathcal{L}^*_\sigma f(s))_\epsilon \mathrm{d}s - \int_0^t (\nabla^*_\sigma f(s))_\epsilon \mathrm{d}W(s).$$

Then, using Itôs formula applied to $\Gamma(f_\epsilon)$, we have

$$\Gamma(f_\epsilon(t)) = \Gamma(f_\epsilon(0)) + \int_0^t \Gamma'(f_\epsilon(s))(\mathcal{L}^*_\sigma f(s))_\epsilon \mathrm{d}s - \int_0^t \Gamma'(f_\epsilon(s))(\nabla^*_\sigma f)_\epsilon \mathrm{d}W(s)$$
$$+ \frac{1}{2}\int_0^t \Gamma''(f_\epsilon(s))(\nabla^*_\sigma f)^2_\epsilon \mathrm{d}s$$

We can then write the equation above as a stochastic continuity equation for $\Gamma(f)$ plus some remainders. Specifically, we have

$$\begin{aligned}\Gamma(f_\epsilon(t)) = \Gamma(f_\epsilon(0)) + \int_0^t \mathcal{L}_\sigma^* \Gamma(f_\epsilon(s))\mathrm{d}s - \int_0^t \nabla_\sigma^* \Gamma(f_\epsilon(s))\mathrm{d}W(s) \\ + \int_0^t R_\epsilon^1(f(s))\mathrm{d}W(s) + \int_0^t R_\epsilon^2(f(s))\mathrm{d}s.\end{aligned} \tag{6.10}$$

where the remainders $R_\epsilon^1(f)$ and $R_\epsilon^2(f)$ are given by

$$R_\epsilon^1(f) = \nabla_\sigma^*\Gamma(f_\epsilon) - \Gamma'(f_\epsilon)(\nabla_\sigma^* f)_\epsilon$$

$$R_\epsilon^2(f) = \Gamma'(f_\epsilon)(\mathcal{L}_\sigma^* f)_\epsilon - \mathcal{L}_\sigma^*\Gamma(f_\epsilon) + \frac{1}{2}\Gamma''(\nabla_\sigma^* f)_\epsilon^2$$

In order to complete the proof we need to show that as $\epsilon \to 0$, the remainders $R_\epsilon^1(f)$ and $R_\epsilon^2(f)$ converge to the correct terms on the right-hand side of (6.8). To show this, we will make use of the following lemma which writes $R_\epsilon^1(f)$ and $R_\epsilon^2(f)$ in terms of the commutators $[\nabla_\sigma, \eta_\epsilon](f)$ and $\big[[\mathcal{L}_\sigma, \eta_\epsilon]\big](f)$.

Lemma 6.2.3. *We have the following identities*

$$R_\epsilon^1(f) = \Gamma'(f_\epsilon)[\nabla_\sigma, \eta_\epsilon](f) - \operatorname{div}\sigma\Gamma(f_\epsilon)$$

$$\begin{aligned}R_\epsilon^2(f) = \Gamma'(f_\epsilon)[[\mathcal{L}_\sigma, \eta_\epsilon]](f) + \frac{1}{2}\Gamma''(f_\epsilon)([\nabla_\sigma, \eta_\epsilon](f))^2 + \nabla_\sigma^*\big(\Gamma'(f_\epsilon)[\nabla_\sigma, \eta_\epsilon](f)\big) - \nabla_\sigma^*(\operatorname{div}\sigma\Gamma(f_\epsilon)) \\ - \Gamma'(f_\epsilon)\operatorname{div}\sigma[\nabla_\sigma, \eta_\epsilon](f) + \Gamma(f_\epsilon)D_\sigma\Gamma(f_\epsilon) - \frac{1}{2}\Gamma(f_\epsilon)A_\sigma - \frac{1}{2}\Gamma(f_\epsilon)D_\sigma.\end{aligned}$$

Before proving Lemma 6.2.3, let us see how to complete the proof assuming these identities. We need to pass the limit as $\epsilon \to 0$ in the weak, time integrated form of (6.10) which for each $\varphi \in C_c^\infty(\mathbb{R}^n)$ becomes

$$\begin{aligned}\langle\Gamma(f_\epsilon(t)), \varphi\rangle = \langle\Gamma(f_\epsilon(0)), \varphi\rangle + \int_0^t \langle\Gamma(f_\epsilon(s)), \mathcal{L}_\sigma\varphi\rangle\mathrm{d}s + \int_0^t \langle\Gamma(f_\epsilon(s)), \nabla_\sigma\varphi\rangle\mathrm{d}W(s) \\ + \int_0^t \langle R_\epsilon^1(f(s)), \varphi\rangle\mathrm{d}W(s) + \int_0^t \langle R_\epsilon^2(f(s)), \varphi\rangle\mathrm{d}s.\end{aligned} \tag{6.11}$$

The standard properties of mollifiers imply that for $\mathbf{P}\otimes \mathrm{d}t$ almost every $(\Omega, t) \in \Omega \times [0,T]$, $f_\epsilon(t) \to f(t)$ in $L^p_{x,\mathrm{loc}}$ and $f_\epsilon \to f$ pointwise $\mathbf{P} \otimes \mathrm{d}t \otimes \mathrm{d}x$ almost every where on $\Omega \times [0,T] \times \mathbb{R}^n$. It is a simple matter to show that this, the boundedness of $\Gamma(z)$ and the integrability conditions on σ imply that as $\epsilon \to 0$,

$$\begin{aligned}
&\langle \Gamma(f_\epsilon(0)), \varphi\rangle \to \langle \Gamma(f_0), \varphi\rangle,\\
&\langle \Gamma(f_\epsilon), \varphi\rangle \to \langle \Gamma(f), \varphi\rangle, \quad \text{in} \quad L^1(\Omega \times [0,T])\\
&\langle \Gamma(f_\epsilon), \mathcal{L}_\sigma\varphi\rangle \to \langle \Gamma(f), \mathcal{L}_\sigma\varphi\rangle, \quad \text{in} \quad L^1(\Omega \times [0,T])
\end{aligned}$$

while for the term in the stochastic intergral

$$\langle \Gamma(f_\epsilon), \nabla_\sigma\varphi\rangle \to \langle \Gamma(f), \nabla_\sigma\varphi\rangle \quad \text{in} \quad L^2(\Omega \times [0,T]).$$

Consequently we may pass the limit as $\epsilon \to 0$ in the first four terms of equation (6.11).

What remain are the terms involving $R^1_\epsilon(f)$ and $R^2_\epsilon(f)$. The commutator Lemma 6.2.2, the fact that $\sigma \in L^2_t(W^{1,2p/(p-2)}_{x,\mathrm{loc}})$, and the strong convergence properties of $f_\epsilon \to f$ are more than enough to conclude

$$\langle \Gamma'(f_\epsilon)[\nabla_\sigma, \eta_\epsilon](f), \varphi\rangle \to \langle f\Gamma'(f)\operatorname{div} u, \varphi\rangle, \quad \mathbf{P} \otimes \mathrm{d}t \text{ almost everywhere.}$$

Moreover, the bound provided in the commutator Lemma 6.2.2, and the fact that $\Gamma(z)$ and $\Gamma'(z)$ are bounded functions give the estimate

$$\mathbf{E}\int_0^T |\langle \Gamma'(f_\epsilon(s))[\nabla_\sigma, \eta_\epsilon](f(s)), \varphi\rangle|^2 \mathrm{d}s \lesssim \|\nabla\sigma\|^2_{L^2_t(L^{2p/(p-1)}(K))}\|f\|^2_{L^\infty_t(L^p(K))},$$

where K is a compact set containing the support of φ. Therefore the dominated convergence theorem and the fact that $\Gamma(f_\epsilon) \to \Gamma(f)$ in implies that

$$\langle R^1_\epsilon(f), \varphi\rangle \to \langle G(f)\operatorname{div}\sigma, \varphi\rangle \quad \text{in} \quad L^2(\Omega \times [0,T]),$$

whereby we may pass the limit in the stochastic integral for the fourth term on the right-hand side of (6.10). The last term $R^2_\epsilon(f)$, though complicated, is straightforward and can be treated in a similar manner as $R^1_\epsilon(f)$. Indeed similar arguments to those above show that

$$\langle(\Gamma'(f_\epsilon)[\nabla_\sigma,\eta_\epsilon](f)-\operatorname{div}\sigma\Gamma(f_\epsilon)),\nabla_\sigma\varphi\rangle \to \langle G(f)\operatorname{div}\sigma,\nabla_\sigma\varphi\rangle \quad \text{in} \quad L^1(\Omega\times[0,T]),$$

and

$$\langle\Gamma'(f_\epsilon)\operatorname{div}\sigma[\nabla_\sigma,\eta_\epsilon](f),\varphi\rangle \to \langle f\Gamma'(f)D_\sigma,\varphi\rangle \quad \text{in} \quad L^1(\Omega\times[0,T]).$$

Moreover using the commutator Lemma 6.2.2 applied to the double commutator $\big[[\mathcal{L}_\sigma,\eta_\epsilon]\big](f)$ we also obtain

$$\Big\langle\Gamma'(f_\epsilon)\big[[\mathcal{L}_\sigma,\eta_\epsilon]\big](f)-\tfrac{1}{2}\Gamma(f_\epsilon)(A_\sigma+D_\sigma),\varphi\Big\rangle \to \langle\tfrac{1}{2}G(f)(A_\sigma+D_\sigma),\varphi\rangle \quad \text{in} \quad L^1(\Omega\times[0,T]).$$

The only term left to study in $\mathbb{R}^2_\epsilon(f)$ is $\frac{1}{2}\Gamma''(f_\epsilon)([\nabla_\sigma,\eta_\epsilon](f))^2$. In fact, it is precisely this term that dictate the $L^2_t(W^{1,2p/(p-2)}_x)$ condition on σ (as opposed to $L^2_t(W^{1,2p/(p-1)}_x)$ which is sufficient to obtain all the limits above). Precisely, using the commutator Lemma, and requiring that $\mathbf{P}$ almost surely $[\nabla_\sigma,\eta_\epsilon](f(t))$ converges in $L^2_t(L^2_{x,\text{loc}})$, gives the condition on σ, and along with the bounded of $\Gamma''(z)$ implies that

$$\langle\tfrac{1}{2}\Gamma''(f_\epsilon)([\nabla_\sigma,\eta_\epsilon](f))^2,\varphi\rangle \to \langle\tfrac{1}{2}\Gamma''(f)f^2(\operatorname{div}\sigma)^2,\varphi\rangle \quad \text{in} \quad L^1(\Omega\times[0,T]).$$

The above limits can be collected to conclude that

$$\langle R^2_\epsilon(f),\varphi\rangle \to -\langle G(f)\operatorname{div}\sigma,\nabla_\sigma\varphi\rangle+\frac{1}{2}\langle G(f)A_\sigma,\varphi\rangle+\frac{1}{2}\langle H(f)D_\sigma,\varphi\rangle \quad \text{in} \quad L^1(\Omega\times[0,T]).$$

All of these convergence properties, allow us to pass the limit in each term of (6.11) in $L^1(\Omega \times [0,T])$ and therefore that equation (6.13) holds $\mathbf{P} \otimes \mathrm{d}t$ almost everywhere.

We now proceed to the proof of Lemma 6.2.3:

Proof of Lemma 6.2.3. We begin by remarking that this computation involves quantities, like $\operatorname{div}\operatorname{div} a$ and $\nabla_\sigma \operatorname{div} \sigma$ which are not well-defined functions given the regularity assumptions on σ. However as they are well-defined distributions and are only every multiplied by sooth functions the computations below make sense in the sense of distribution.

The proof of the identity for $R^1_\epsilon(f)$ is obvious given the definition of the commutator $[\nabla_\sigma, \eta_\epsilon](f)$. We focus on $R^2_\epsilon(f)$ and begin by expanding the term for $\mathcal{L}^*\Gamma(f)$,

$$\mathcal{L}^*\Gamma(f_\epsilon) = (\operatorname{div}\operatorname{div} a)\Gamma(f_\epsilon) + \Gamma'(f_\epsilon) 2 \operatorname{div} a \cdot \nabla f_\epsilon + \Gamma''(f_\epsilon)(\nabla_\sigma f_\epsilon)^2 + \Gamma'(f_\epsilon)\mathcal{L} f_\epsilon$$

So that R^2_ϵ becomes

$$\begin{aligned} R^2_\epsilon(f) = &-(\operatorname{div}\operatorname{div} a)\Gamma(f_\epsilon) - \Gamma'(f_\epsilon) 2 \operatorname{div} a \cdot \nabla f_\epsilon - \Gamma'(f_\epsilon)(\mathcal{L} f_\epsilon - (\mathcal{L}^* f)_\epsilon) \\ &+ \frac{1}{2}\Gamma''(f_\epsilon)((\nabla^*_\sigma f)^2_\epsilon - (\nabla_\sigma f_\epsilon)^2) \end{aligned}$$

We can write several expressions in terms of commutators

$$\begin{aligned} \mathcal{L} f_\epsilon - (\mathcal{L}^* f)_\epsilon &= -[[\mathcal{L}, \eta_\epsilon]](f) + 2\mathcal{L} f_\epsilon - \nabla_\sigma (\nabla^*_\sigma f)_\epsilon \\ &= -[[\mathcal{L}, \eta_\epsilon]](f) + 2\mathcal{L} f_\epsilon - \nabla_\sigma \nabla_\sigma f_\epsilon + \nabla_\sigma [\nabla_\sigma, \eta_\epsilon](f) \\ &= -[[\mathcal{L}, \eta_\epsilon]](f) - \nabla_\sigma \sigma \cdot \nabla f_\epsilon + \nabla_\sigma [\nabla_\sigma, \eta_\epsilon](f) \end{aligned}$$

and

$$\begin{aligned} (\nabla^*_\sigma f)^2_\epsilon - (\nabla_\sigma f_\epsilon)^2 &= (\nabla_\sigma f_\epsilon - [\nabla_\sigma, \eta_\epsilon](f))^2 - (\nabla_\sigma f_\epsilon)^2 \\ &= -2\nabla_\sigma f_\epsilon [\nabla_\sigma, \eta_\epsilon](f) + ([\nabla_\sigma, \eta_\epsilon](f))^2 \end{aligned}$$

Therefore we have

$$\frac{1}{2}\Gamma''(f_\epsilon)((\nabla^*_\sigma f)^2_\epsilon - (\nabla_\sigma f_\epsilon)^2) - \Gamma'(f_\epsilon)(\mathcal{L}f_\epsilon - (\mathcal{L}^* f)_\epsilon)$$

$$= \Gamma'(f_\epsilon)[[\mathcal{L},\eta_\epsilon]](f) + \frac{1}{2}([\nabla_\sigma,\eta_\epsilon](f))^2 - \Gamma''(f_\epsilon)\nabla_\sigma f_\epsilon[\nabla_\sigma,\eta_\epsilon](f) - \Gamma'(f_\epsilon)\nabla_\sigma[\nabla_\sigma,\eta_\epsilon](f)$$

$$+ \Gamma'(f_\epsilon)\nabla_\sigma\sigma\cdot\nabla f_\epsilon$$

$$= \Gamma'(f_\epsilon)[[\mathcal{L},\eta_\epsilon]](f) + \frac{1}{2}\Gamma''(f_\epsilon)([\nabla_\sigma,\eta_\epsilon](f))^2 + \nabla_\sigma\big(\Gamma'(f)[\nabla_\sigma,\eta_\epsilon](f)\big) + \Gamma'(f_\epsilon)\nabla_\sigma\sigma\cdot\nabla f_\epsilon$$

The remainder becomes

$$R^2_\epsilon(f) = \Gamma'(f_\epsilon)[[\mathcal{L},\eta_\epsilon]](f) + \frac{1}{2}\Gamma''(f_\epsilon)([\nabla_\sigma,\eta_\epsilon](f))^2 + \nabla_\sigma\big(\Gamma'(f_\epsilon)[\nabla_\sigma,\eta_\epsilon](f)\big) + \Gamma'(f_\epsilon)\nabla_\sigma\sigma\cdot\nabla f_\epsilon$$

$$- \operatorname{div}\operatorname{div} a\Gamma(f_\epsilon) - \Gamma'(f_\epsilon)2\operatorname{div} a\cdot\nabla f_\epsilon$$

Next we write

$$\Gamma'(f_\epsilon)\nabla_\sigma\sigma\cdot\nabla f_\epsilon - \Gamma'(f_\epsilon)2\operatorname{div} a\cdot\nabla f_\epsilon = -\operatorname{div}\sigma\nabla_\sigma\Gamma(f_\epsilon)$$

$$= -\nabla_\sigma(\operatorname{div}\sigma\Gamma(f_\epsilon)) + \nabla_\sigma\operatorname{div}\sigma\Gamma(f_\epsilon),$$

and use the fact that $(\operatorname{div}\operatorname{div} a) = \frac{1}{2}A_\sigma + \frac{1}{2}D_\sigma + \nabla_\sigma\operatorname{div}\sigma$ to simplify the remainder to

$$R^2_\epsilon(f) = \Gamma'(f_\epsilon)[[\mathcal{L},\eta_\epsilon]](f) + \frac{1}{2}\Gamma''(f_\epsilon)([\nabla_\sigma,\eta_\epsilon](f))^2 + \nabla_\sigma\big(\Gamma'(f_\epsilon)[\nabla_\sigma,\eta_\epsilon](f)\big) - \nabla_\sigma(\operatorname{div}\sigma\Gamma(f_\epsilon))$$

$$- \frac{1}{2}\Gamma(f_\epsilon)A_\sigma - \frac{1}{2}\Gamma(f_\epsilon)D_\sigma.$$

The lemma now follows by writing $\nabla_\sigma = \nabla^*_\sigma - \operatorname{div}\sigma$ in two of the terms above. □

This completes the proof of Theorem 6.2.1 □

Renormalization with drift and a family of noise coefficients

Theorem 6.2.1 can be easily generalized to equations with a drift u and a family of noise coefficients $\sigma = \{\sigma^k\}$,

$$\partial_t f + \operatorname{div}(uf) - \operatorname{div}\operatorname{div}(af) + \sum_k \operatorname{div}(\sigma^k f)\dot{W}^k = 0, \quad a = \frac{1}{2}\sum_k \sigma^k\otimes\sigma^k \tag{6.12}$$

as long as u satisfies the usual regularity requirements of the deterministic DiPerna-Lions theory and $\sigma = \{\sigma^k\}$ satisfy the appropriate summability conditions. In this case the renormalized form looks like

$$\partial_t \Gamma(f) - \mathcal{L}^*_{u,\sigma}\Gamma(f) + \nabla^*_\sigma \Gamma(f)\dot{W} = \nabla^*_\sigma(\operatorname{div}\sigma\, G(f)) - \operatorname{div} u\, G(f) - \operatorname{div}\sigma\, G(f)\dot{W} + \tfrac{1}{2}G(f)A_\sigma + \tfrac{1}{2}H(f)D_\sigma.$$

where $\mathcal{L}_{u,\sigma} = u \cdot \nabla + a : \nabla^2$. The corresponding renormalization result is given below:

Theorem 6.2.4. *Let $f \in L^\infty_t(L^p_{x,\mathrm{loc}})$, $p > 2$, be a weak L^p solution to (6.12). Suppose that $u \in L^1_t(W^{1,q}_{x,\mathrm{loc}})$ and $\sigma^k \in L^2_t(W^{1,2p/(p-2)}_{x,\mathrm{loc}})$ satisfing the summability condition*

$$\sum_k \|\sigma^k\|^2_{L^2_t(W^{1,2p/(p-2)}(K))} < \infty$$

for every compact $K \subseteq \mathbb{R}^n$. Then for any $\Gamma \in C^2_b(\mathbb{R}^n)$, such that $\sup_z z\Gamma'(z)$ and $\sup_z z^2\Gamma''(z) < \infty$, $\Gamma(f)$ solves the renormalized equation in time-integrated, weak form, namely for $\mathbf{P} \otimes \mathrm{dt}$ almost every (t, ω) and every $\varphi \in C^\infty_c(\mathbb{R}^n)$ the following equality holds

$$\begin{aligned} \langle \Gamma(f(t)), \varphi\rangle = \langle \Gamma(f_0), \varphi\rangle &+ \int_0^t \langle \Gamma(f(s)), \mathcal{L}_{u,\sigma}\varphi\rangle \mathrm{d}s + \sum_k \int_0^t \langle \Gamma(f(s)), \nabla_{\sigma^k}\varphi\rangle \mathrm{d}W^k(s) \\ &- \int_0^t \langle G(f(s)) \operatorname{div} u, \varphi\rangle \mathrm{d}s - \sum_k \int_0^t \langle G(f(s)) \operatorname{div}\sigma^k, \nabla_{\sigma^k}\varphi\rangle \mathrm{d}s \\ &+ \sum_k \frac{1}{2}\int_0^t \langle G(f(s))A_{\sigma^k}, \varphi\rangle \mathrm{d}s + \sum_k \frac{1}{2}\int_0^t \langle H(f(s))D_{\sigma^k}, \varphi\rangle \mathrm{d}s \\ &- \sum_k \int_0^t \langle G(f(s)) \operatorname{div}\sigma^k, \varphi\rangle \mathrm{d}W^k(s). \end{aligned} \tag{6.13}$$

Proof. The proof is an easy extension of the Theorem 6.2.1. Regulaizing and renormalizing just as in the proof of Theorem 6.2.1, we see that the drift u introduces another commutator $[\nabla_u, \eta_\epsilon](f)$ which satisfies

$$\langle \Gamma'(f_\epsilon)[\nabla_u, \eta_\epsilon](f) - \Gamma(f_\epsilon) \operatorname{div} u, \varphi \rangle \to \langle G(f) \operatorname{div} u, \varphi \rangle \quad \text{in} \quad L^1(\Omega \times [0,T])$$

as long as $u \in L^1([0,T]; W^{1,q}_{x,\mathrm{loc}})$. Furthermore the summability condition on $\sigma = \{\sigma^k\}$ allows one to pass the limit in each term of the sum just as in Theorem 6.2.1 and then, using the fact that each term in the sum which isn't a stochastic integral has a uniform (in ϵ) bound in $L^1(\Omega \times [0,T])$ by some contant times $\|\nabla \sigma\|^2_{L^2_t(L^{2p/(p-2)}(K))}$, the dominated convergence for series allows one to pass the limit in the summation. The same arguement work for the stochastic integrals where instead the bound is in $L^2(\Omega \times [0,T])$. $\square$

The Stochastic Boltzmann Equation (w/ Scott Smith)

Introduction

The Boltzmann equation

$$\begin{aligned} &\partial_t f + v \cdot \nabla_x f + \mathrm{div}_v(\mathcal{X} f) = \mathcal{B}(f,f), \\ &f|_{t=0} = f_0, \end{aligned} \tag{7.1}$$

on $[0,T] \times \mathbb{R}^{2d}$ is a nonlinear integro-differential equation describing the evolution of a rarefied gas, dominated by binary collisions, and in the presence of a external force field $\mathcal{X}$. The function $f(t,x,v) \in \mathbb{R}$ describes the density of particles at time $t \in [0,T]$, position $x \in \mathbb{R}^d$, with velocity $v \in \mathbb{R}^d$, starting at $t = 0$ from an initial density $f_0(x,v)$. The nonlinear functional $f \mapsto \mathcal{B}(f,f)$, known as the collision operator, acts on the velocity variable only, and accounts for the effect of collisions between pairs of particles; it will be described in more detail below.

Several studies have been conducted regarding the well-posedness of the Cauchy problem for the Boltzmann equation (7.1) with a fixed (deterministic) external force, for instance [7, 12, 37, 115]. In general, the external force field $\mathcal{X}$ may depend on $(t,x,v) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$. Such external forces may arise when considering the influence of gravity such as in the treatment of the Rayleigh-Benard problem in the kinetic regime [5, 47]. In fact, many external forces are not fixed, and are instead coupled with the density f in a self consistent way. This is the case, for example, with the Vlasov-Poisson-Boltzmann and Vlasov-Maxwell-Boltzmann equations (see [23, 86] and references therein for more details on these systems).

This article focuses instead on the Cauchy problem for the Boltzmann equation with *random* external forcing. In particular, we are interested in the following SPDE

$$\begin{aligned} &\partial_t f + v \cdot \nabla_x f + \mathrm{div}_v(f\sigma_k \circ \dot{\beta}_k) = \mathcal{B}(f,f), \\ &f|_{t=0} = f_0, \end{aligned} \tag{SB}$$

where $\{\beta_k\}_{k\in\mathbb{N}}$ are one-dimensional Brownian motions and $\{\sigma_k\}_{k\in\mathbb{N}}$ are a family of vector fields $\sigma_k : \mathbb{R}^{2d} \to \mathbb{R}^d$ with $\mathrm{div}_v\, \sigma_k = 0$. An implicit summation is taken over $k \in \mathbb{N}$, and the expression $\mathrm{div}_v(f\sigma_k \circ \dot{\beta}_k)$ denotes a transport type multiplicative noise, white in time and colored in (x, v), where the product $\circ$ is interpreted in the Stratonovich sense.

Physically, we view the quantity

$$(t, x, v) \mapsto \sum_{k\in\mathbb{N}} \sigma_k(x, v)\dot{\beta}_k(t)$$

as an environmental noise acting on the gas. In the absence of collisions, all particles evolve according to the stochastic differential equation

$$\mathrm{d}X_t = V_t \mathrm{d}t, \quad \mathrm{d}V_t = \sum_{k\in\mathbb{N}} \sigma_k(X_t, V_t) \circ \mathrm{d}\beta_k(t) \tag{7.2}$$

and are only distinguished from one another according to their initial location in the phase space. Let $\Phi_{s,t}(x, v)$ be the stochastic flow associated with the SDE (7.2), that is, $t \mapsto \Phi_{s,t}(x, v) = (X_t, V_t)$ solves (7.2) and satisfies $\Phi_{s,s}(x, v) = (x, v)$. The Stratonovich form of the noise and the fact that $\mathrm{div}_v\, \sigma_k = 0$ ensures that the flow

$\Phi_{s,t}$ is volume preserving (with probability one). The density of the collision-less gas is then given by $f_t(x,v) = f_0(\Phi_{0,t}^{-1}(x,v))$ and evolves according to the free stochastic kinetic transport equation

$$\partial_t f + v \cdot \nabla_x f + \mathrm{div}_v(f\sigma_k \circ \dot{\beta}_k) = 0,$$
$$f|_{t=0} = f_0.$$

The presence of collisions interrupts the stochastic transport process. In the low volume density regime, binary collisions are dominant and can be described by the Boltzmann collision operator $\mathcal{B}(f,f)$. The stochastic Boltzmann equation (SB) accounts for both stochastic transport and binary collisions. In fact, formally (SB) can be written in mild form,

$$f_t = f_0 \circ \Phi_{0,t}^{-1} + \int_0^t \mathcal{B}(f_s, f_s) \circ \Phi_{s,t}^{-1} \mathrm{d}s.$$

The stochastic Boltzmann equation (SB) can be interpreted as the so-called Boltzmann-Grad limiting description of interacting particles subject to the *same* environmental noise. In the deterministic setting, the Boltzmann-Grad problem has been studied extensively in the literature (see [60] for a recent review). In the stochastic setting, the Boltzmann-Grad problem has (to our knowledge) not yet been studied. However, a mean field limit to the Vlasov equation with stochastic kinetic transport has been shown recently by Coghi and Flandoli [27].

To our knowledge, this is the first study to obtain mathematically rigorous results on the Boltzmann equation with a random external force. However, a number of results on the *fluctuating Boltzmann equation* are available in the Math and Physics literature [16, 56, 94, 109–111, 114]. In particular, the articles of

Bixon/Zwanzig [16] and Fox/Uhlenbeck [56] outline a formal derivation of Landau and Lifshitz's equations of fluctuating hydrodynamics [83], from the fluctuating linear Boltzmann equation. The connection with macroscopic fluid equations arises from studying the correlation structure of the fluctuations at the level of the kinetic description. A more rigorous treatment of the fluctuation theory for the Boltzmann equation and its connection to the Boltzmann-Grad limit is given by Spohn [109–111].

Although our perspective differs from that of [56] and [16], we do expect to obtain various stochastic hydrodynamic equations (with colored noise) in different asymptotic regimes, using a Chapman-Enskog expansion and the moments method of Bardos/Golse/Levermore [9]. In fact, one of the original motivations for this article was to understand which of the common forms of noise in the stochastic fluids literature can be obtained by considering fluctuations of the stochastic kinetic description relative to an equilibrium state. This will be addressed in detail in future works.

The goal of this article is to investigate global solutions to (SB) starting from general 'large' initial data $f_0 \in L^1(\mathbb{R}^{2d})$. If the noise coefficients σ are identically zero, then this problem has already been addressed in the seminal work of DiPerna/Lions [36], where existence of renormalized solutions is proved. Our work is heavily inspired by [36], relying on a number of their insights together with various classical properties of the Boltzmann equation. Rather than give a detailed review, in the next subsection we will explain how these observations from the deterministic theory lead to the notion of renormalized martingale solution to (SB) in the present context. Finally, we should mention that our initial motivation for the choice of noise was heavily inspired by a number of interesting works on stochastic transport equations (see for instance [33, 48, 52, 53]). Finally, we should mention the work [20] on the 2-d stochastic Euler equations with a very similar noise to the one in this paper.

Statement of the main result

Let us begin by discussing the basics of the Boltzmann equation and introduce the analytical framework for the problem. We refer the reader to the books [23, 24] and the excellent set of notes [65] for a comprehensive introduction to the Boltzmann equation, as well as the review [119].

The collision operator $\mathcal{B}(f, f)$ describes the rate of change in particle density due to collisions. It contains all the information about collision rates between particles with different velocities. More precisely, it is defined through its action in v as

$$\mathcal{B}(f,f)(v) = \iint_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (f' f'_* - f f_*)(v, v_*, \theta) b(v - v_*, \theta) \mathrm{d}\theta \mathrm{d}v_*, \tag{7.3}$$

where f_*, f', and f'_* are shorthand for $f(v_*)$, $f(v')$, and $f(v'_*)$, while (v', v'_*) denote pre-collisional velocities

$$\begin{cases} v' = v - (v - v_*) \cdot \theta \, \theta \\ v'_* = v_* + (v - v_*) \cdot \theta \, \theta. \end{cases}$$

Note that (v', v'_*), parametrized by $\theta \in \mathbb{S}^{d-1}$, are solutions to the equations describing pairwise conservation of momentum and energy,

$$v' + v'_* = v + v_*$$

$$|v'|^2 + |v'_*|^2 = |v|^2 + |v_*|^2.$$

The collision kernel $b(v - v_*, \theta) \geq 0$ is determined by details of the inter-molecular forces between particles and describes the rate at which particles with relative velocity $v - v_*$ collide with deflection angle $\theta \cdot (v - v_*)/|v - v_*|$. In this article, for technical reasons and simplicity of exposition, we restrict our attention to bounded, integrable kernels, though we intend to investigate (in a future work) the possibility of treating more singular kernels as in Alexandre/Villani [3] and other works. Our assumption on the collision kernel is the following:

Hypothesis 7.1.1. *The collision kernel $b(z, \theta)$ depends solely on $|z|$ and $|z \cdot \theta|$ only, and satisfies,*

$$b \in L^1(\mathbb{R}^{2d} \times \mathbb{S}^{d-1}) \cap L^\infty(\mathbb{R}^{2d} \times \mathbb{S}^{d-1}).$$

Since the nonlinear term $\mathcal{B}(f, f)$ is quadratic in f, further properties of the operator must be exploited in order to obtain a priori bounds. A classical observation is that the symmetry assumptions on the collision kernel b imposed in Hypothesis 7.1.1 and the definition of (v', v'_*) imply that for each smooth $\xi : \mathbb{R}^d \to \mathbb{R}$,

$$\begin{aligned}&\int_{\mathbb{R}^d} \xi(v)\mathcal{B}(f, f)\mathrm{d}v \\ &\qquad = \frac{1}{4}\iiint_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} (f'f'_* - ff_*)(\xi_* + \xi - \xi'_* - \xi')\, b(v - v_*, \theta)\, \mathrm{d}\theta \mathrm{d}v_* \mathrm{d}v.\end{aligned} \tag{7.4}$$

Any quantity $\xi(v)$ such that $\xi_* + \xi = \xi'_* + \xi'$, is called a collision invariant. For any collision invariant $\xi(v)$, (7.4) implies that

$$\int_{\mathbb{R}^d} \xi(v)\mathcal{B}(f, f)(v)\mathrm{d}v = 0.$$

As a result of the definition of (v', v'_*), the quantities $\{1, \{v_i\}_{i=1}^d, |v|^2\}$ are collision invariants. Therefore, multiplying both sides of (SB) by a collision invariant and integrating in v, the collision operator vanishes

$$\partial_t\Big(\int_{\mathbb{R}^d} \xi(v) f \,\mathrm{d}v\Big) + \mathrm{div}_x\Big(\int_{\mathbb{R}^d} v\xi(v) f \,\mathrm{d}v\Big) = \Big(\int_{\mathbb{R}^d} \nabla\xi(v)\cdot\sigma_k\, f\,\mathrm{d}v\Big)\circ\dot{\beta}_k. \tag{7.5}$$

In the case that $\xi(v) = 1+|v|^2$ in (7.5), one can close on estimate on $\xi(v)f$, provided we have the following coloring hypothesis on σ:

Hypothesis 7.1.2. *For each $k \in \mathbb{N}$, the noise coefficient $\sigma_k : \mathbb{R}^{2d} \to \mathbb{R}^d$ satisfies* $\mathrm{div}_v\,\sigma_k = 0$*. In addition, the sequence $\sigma = \{\sigma_k\}_{k\in\mathbb{N}}$ obeys:*

$$\|\sigma\|_{\ell^2(\mathbb{N};L^\infty_{x,v})} = \Big(\sum_{k\in\mathbb{N}} \|\sigma_k\|^2_{L^\infty_{x,v}}\Big)^{1/2} < \infty \tag{H1}$$

$$\|\sigma\cdot\nabla_v\sigma\|_{\ell^1(\mathbb{N};L^\infty_{x,v})} = \sum_{k\in\mathbb{N}} \|\sigma_k\cdot\nabla_v\sigma_k\|_{L^\infty_{x,v}} < \infty. \tag{H2}$$

More generally, in Section 7.2 we show that Hypothesis 7.1.2 implies that a solution f to (SB) satisfies the following formal a priori bound

$$\mathbf{E}\|(1+|x|^2+|v|^2)f\|^p_{L^\infty_t(L^1_{x,v})} \le C_p, \tag{7.6}$$

for all $p \in [1,\infty)$ and some positive constant C_p (depending on p). In addition, a further $L\log L$ estimate on f is available due to the entropy structure of (SB). To obtain this, let $\Gamma : \mathbb{R} \to \mathbb{R}$ be a sufficiently smooth function, which we will refer to as a renormalization. Since we use Stratonovich noise and $\mathrm{div}_v\,\sigma_k = 0$, if f is a solution of (SB), then formally $\Gamma(f)$ should satisfy:

$$\begin{aligned} &\partial_t\Gamma(f) + v\cdot\nabla_x\Gamma(f) + \mathrm{div}_v(\Gamma(f)\sigma_k\circ\dot{\beta}_k) = \Gamma'(f)\mathcal{B}(f,f),\\ &\Gamma(f)|_{t=0} = \Gamma(f_0). \end{aligned} \tag{RSB}$$

In particular, taking $\Gamma(f) = f \log f$ in (RSB) and integrating in v yields

$$\partial_t \Big(\int_{\mathbb{R}^d} f \log f \, \mathrm{d}v \Big) + \mathrm{div}_x \Big(\int_{\mathbb{R}^d} v \, f \log f \, \mathrm{d}v \Big) = -\mathcal{D}(f), \tag{7.7}$$

where

$$\begin{aligned} \mathcal{D}(f) &\equiv \frac{1}{4} \iiint_{\mathbb{R}^{2d} \times \mathbb{S}^{d-1}} d(f)(t,x,v,v_*,\theta) \, b(v - v_*, \theta) \, \mathrm{d}\theta \mathrm{d}v_* \mathrm{d}v, \\ d(f) &\equiv (f' f'_* - f f_*) \log \left(\frac{f' f'_*}{f f_*} \right) \geq 0. \end{aligned} \tag{7.8}$$

Equation (7.7) describes the local dissipation of the entropy density $\int_{\mathbb{R}^d} f \log f \mathrm{d}v$. The quantity $\mathcal{D}(f)$ is referred to as the entropy dissipation, and inherits non-negativity from $d(f)$. Since $f \log f$ is unsigned, we cannot immediately use (7.7) to obtain an $L \log L$ bound. However, combining this with (7.6), in Section 7.2 we show that for all $p \in [1, \infty)$

$$\mathbf{E} \| f \log f \|^p_{L^\infty_t(L^1_{x,v})}, \quad \mathbf{E} \| \mathcal{D}(f) \|^p_{L^1_{t,x}} \leq C_p. \tag{7.9}$$

Although the a priori bounds (7.6) and (7.9) provide a useful starting point, they are unfortunately insufficient to give a meaning to $\mathcal{B}(f,f)$ in the sense of distributions. For bounded kernels, one can obtain an L^1_v estimate on $\mathcal{B}(f,f)$,

$$\| \mathcal{B}(f,f) \|_{L^1_v} \leq C \| f \|^2_{L^1_v}.$$

However, since $\mathcal{B}(f,f)$ acts pointwise in x, the operator $f \mapsto \mathcal{B}(f,f)$ sends $L^1_{x,v}$ to $L^0_x(L^1_v)$ (a measurable function in x). A key observation of DiPerna and Lions [36] is that the renormalized collision operator $f \to (1+f)^{-1} \mathcal{B}(f,f)$ is better behaved. More precisely, the following inequality holds:

$$\| (1+f)^{-1} \mathcal{B}(f,f) \|_{L^1_{t,x,v}} \lesssim \| \mathcal{D}(f) \|_{L^1_{t,x}} + \| f \|_{L^1_{t,x,v}}. \tag{7.10}$$

Thus, if f satisfies the a priori bounds (7.6) and (7.9), the quantity $(1+f)^{-1}\mathcal{B}(f,f)$ is well defined in $L^{\infty-}(\Omega; L^1_{t,x,v})$. Hence, it becomes feasible to search for solutions satisfying (RSB) in the sense of distributions for a suitable class of renormalizations. Towards this end, we make the following definition:

Definition 7.1.3. Define the set of renormalizations $\mathcal{R}$ to consist of $C^1(\mathbb{R}_+)$ functions $\Gamma : \mathbb{R}_+ \to \mathbb{R}$ such that the mapping $x \mapsto (1+x)\,|\Gamma'(x)|$ belongs to $L^\infty(\mathbb{R}_+)$.

It is important to keep in mind that this class of renormalizations excludes the possibility of choosing $\Gamma(f) = f$ or $\Gamma(f) = f \log f$ and therefore extra care must be taken to obtain the a priori estimates (7.6) and (7.9) above.

We note that for analytical purposes, relating to martingale techniques, it is often more convenient to work with (RSB) in Itô form. Thus, we introduce the matrix

$$a(x,v) = \frac{1}{2}\sum_{k\in\mathbb{N}} \sigma_k(x,v) \otimes \sigma_k(x,v),$$

and define the operator

$$\mathcal{L}_\sigma \varphi = \mathrm{div}_v(a\nabla_v \varphi).$$

Using the divergence free assumption for each σ_k, the random transport term in (RSB) can be converted to Itô form via the relation

$$\mathrm{div}_v(\Gamma(f)\sigma_k \circ \dot{\beta}_k) = -\mathcal{L}_\sigma \Gamma(f) + \mathrm{div}_v(\Gamma(f)\sigma_k \dot{\beta}_k).$$

We are now ready to define our notion of solution for (RSB).

Definition 7.1.4. A density f is defined to be a renormalized martingale solution to (SB) provided there exists a stochastic basis $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t=0}^T, \{\beta_k\}_{k\in\mathbb{N}})$ such that the following hold:

1. For all $(t,\omega) \in [0,T] \times \Omega$, the quantity $f(t,\omega)$ is a non-negative element of $L^1_{x,v}$.

2. The mapping $f : [0,T] \times \Omega \to L^1_{x,v}$ defines an $(\mathcal{F}_t)_{t=0}^T$ adapted process with continuous sample paths.

3. For all renormalizations $\Gamma \in \mathcal{R}$, test functions $\varphi \in C_c^\infty(\mathbb{R}^{2d})$, and times $t \in [0,T]$; the following equality holds $\mathbf{P}$ almost surely:

$$
\begin{aligned}
&\iint_{\mathbb{R}^{2d}} \Gamma(f)(t)\varphi \mathrm{d}x\mathrm{d}v = \iint_{\mathbb{R}^{2d}} \Gamma(f_0)\varphi \mathrm{d}x\mathrm{d}v \\
&\quad + \int_0^t \iint_{\mathbb{R}^{2d}} [\Gamma(f) v \cdot \nabla_x \varphi + \Gamma'(f)\mathcal{B}(f,f)\varphi] \mathrm{d}x\mathrm{d}v\mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t \iint_{\mathbb{R}^{2d}} \Gamma(f)\mathcal{L}_\sigma \varphi \, \mathrm{d}x\mathrm{d}v\mathrm{d}s \; + \sum_{k\in\mathbb{N}} \int_0^t \iint_{\mathbb{R}^{2d}} \Gamma(f)\sigma_k \cdot \nabla_v \varphi \, \mathrm{d}x\mathrm{d}v\mathrm{d}\beta_k(s).
\end{aligned}
\tag{7.11}
$$

4. For all $p \in [1,\infty)$ there exists a positive constant C_p such that:

$$
\mathbf{E}\|(1+|x|^2+|v|^2+|\log f|)f\|^p_{L^\infty_t(L^1_{x,v})}, \quad \mathbf{E}\|\mathcal{D}(f)\|^p_{L^1_{t,x}} \leq C_p. \tag{7.12}
$$

Remark 7.1.5. In light of the estimate (7.10), the estimates in condition 4 of Definition 7.1.4 ensure that the weak form (7.11) is well defined and the stochastic integral is a continuous-time martingale.

At present, we require a further technical hypothesis on σ and $\sigma \cdot \nabla_v \sigma$. This is related to the regularity needed on σ to renormalize a linear, stochastic kinetic transport equation, a crucial procedure in our analysis. This is discussed in more detail in Section 7.1.2 below.

Hypothesis 7.1.6. *There exists an $\epsilon > 0$ such that:*

$$\|\sigma\|_{\ell^2(\mathbb{N};W^{1,2+\epsilon}_{x,v})} = \Big(\sum_{k\in\mathbb{N}} \|\sigma_k\|^2_{W^{1,2+\epsilon}_{x,v}}\Big)^{1/2} < \infty \tag{H3}$$

$$\|\sigma\cdot\nabla_v\sigma\|_{\ell^1(\mathbb{N};W^{1,1+\epsilon}_{x,v})} = \sum_{k\in\mathbb{N}} \|\sigma_k\cdot\nabla_v\sigma_k\|_{W^{1,1+\epsilon}_{x,v}} < \infty. \tag{H4}$$

The main result of this article is the following global existence theorem:

Theorem 7.1.7. *Let $\{\sigma_k\}_{k\in\mathbb{N}}$ be a collection of noise coefficients satisfying Hypotheses 7.1.2 and 7.1.6 and assume that the collision kernel b satisfies Hypothesis 7.1.1. For any initial data $f_0 : \mathbb{R}^{2d} \to \mathbb{R}_+$ satisfying*

$$(1 + |x|^2 + |v|^2 + |\log f_0|)f_0 \in L^1_{x,v},$$

there exists a renormalized martingale solution to (SB), starting from f_0 with noise coefficients $\{\sigma_k\}_{k\in\mathbb{N}}$.

Moreover f satisfies

- *almost sure local conservation of mass*

$$\partial_t \int_{\mathbb{R}^d} f\mathrm{d}v + \mathrm{div}_x \int_{\mathbb{R}^d} vf\mathrm{d}x = 0, \tag{7.13}$$

- *average global balance of momentum*

$$\mathbf{E}\iint_{\mathbb{R}^{2d}} vf(t)\mathrm{d}v\mathrm{d}x = \frac{1}{2}\sum_k \mathbf{E}\int_0^t \iint_{\mathbb{R}^{2d}} \sigma_k\cdot\nabla_v\sigma_k f(s)\mathrm{d}v\mathrm{d}x\mathrm{d}s + \iint_{\mathbb{R}^{2d}} vf_0\mathrm{d}v\mathrm{d}x, \tag{7.14}$$

- *average global energy inequality*

$$\mathbf{E}\iint_{\mathbb{R}^{2d}} \frac{1}{2}|v|^2 f(t)\mathrm{d}v\mathrm{d}x \le \sum_k \mathbf{E}\int_0^t \iint_{\mathbb{R}^{2d}} (v\cdot(\sigma_k\cdot\nabla_v\sigma_k) + |\sigma_k|^2)f(s)\mathrm{d}v\mathrm{d}x\mathrm{d}s$$
$$+ \iint_{\mathbb{R}^{2d}} \frac{1}{2}|v|^2 f_0\mathrm{d}v\mathrm{d}x, \tag{7.15}$$

- *almost sure global entropy inequality*

$$\iint_{\mathbb{R}^{2d}} f(t)\log f(t)\mathrm{d}v\mathrm{d}x + \int_0^t \int_{\mathbb{R}^d} \mathcal{D}(f)(s)\mathrm{d}x\mathrm{d}s \leq \iint f_0 \log f_0 \mathrm{d}v\mathrm{d}x. \quad (7.16)$$

The almost sure local conservation of mass holds **P** *almost surely in distribution, the average global momentum and energy balances hold for every* $t \in [0,T]$*, and the global entropy inequality holds* **P** *almost surely for every* $t \in [0,T]$.

Overview of the article

Our analysis begins with formal a priori estimates which point to the natural functional framework for (SB). Namely, in Section 7.2 we show that under the coloring Hypotheses (H1) and (H2), solutions to (SB) formally satisfy

$$\mathbf{E}\|(1+|x|^2+|v|^2+|\log f|)f\|^p_{L^\infty_t(L^1_{x,v})} \leq C_p,$$
$$\mathbf{E}\|\mathcal{D}(f)\|^p_{L^1_{t,x}} \leq C_p.$$

With these formal a priori bounds at hand, the remainder of the paper splits roughly into two parts. In Sections 7.3 and 7.4, we analyze linear stochastic kinetic equations, while Sections $5-8$ are devoted to the proof of Theorem 7.1.7.

In Sections 7.3 and 7.4 we move to a detailed discussion of stochastic kinetic equations of the form

$$\begin{aligned} &\partial_t f + v\cdot\nabla_x f + \mathrm{div}_v(f\sigma_k \circ \dot{\beta}_k) = g, \\ &f|_{t=0} = f_0. \end{aligned} \quad (7.17)$$

Here $f_0 \in L^1_{x,v}$ is a deterministic initial density, while g is a certain random variable with values in $L^1_{t,x,v}$. We will focus on so-called weak martingale solutions to (7.17).

Roughly speaking (see Definition 7.3.1 of Section 7.3.1 for the precise meaning), these are $L^1_{x,v}$ valued stochastic processes satisfying (7.17) weakly in both the PDE and the probabilistic sense. In this context, probabilistically weak means that the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T)$ and the Brownian motions $\{\beta_k\}_{k\in\mathbb{N}}$ are not fixed in advance, but found as solutions to the problem, along with the process f solving (7.17) in the sense of distribution.

For convenience we introduce the following language to refer to solutions of (7.17), we say that: *f is a solution to the stochastic kinetic equation driven by g and starting from f_0, relative to the noise coefficients σ and the stochastic basis* $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t=0}^T, \{\beta_k\}_{k\in\mathbb{N}})$. In the case that the coefficients σ, the filtration $(\mathcal{F}_t)_{t=0}^T$, and the Brownian motions $\{\beta_k\}_{k\in\mathbb{N}}$ are implicitly known or irrelevant, we may omit them from the statement, saying instead: *f is a solution to the stochastic kinetic equation driven by g and starting from f_0.*

A key workhorse for our analysis is a stability result (Proposition 7.3.5) for weak martingale solutions to stochastic kinetic equations. In the deterministic setting, this simply corresponds to the observation that the space of solutions to linear, kinetic equations is closed with respect to convergence in distribution. More precisely, if

$$\partial_t f_n + v \cdot \nabla_x f_n = g_n \quad \text{in} \quad \mathcal{D}'_{t,x,v},$$

$$f|_{t=0} = f_0^n,$$

and $\{(f_n, g_n, f_0^n\}_{n\in\mathbb{N}}$ converges to (f, g, f^0) in $[\mathcal{D}'_{t,x,v}]^3$, then it easily follows from the linear structure of the equation that

$$\partial_t f + v \cdot \nabla_x f = g \quad \text{in} \quad \mathcal{D}'_{t,x,v},$$

$$f|_{t=0} = f_0.$$

In the stochastic framework, an additional subtlety arises. Namely, one should distinguish between stability of stochastically strong solutions, where a stochastic basis has been fixed, and stability of stochastically weak solutions, where each solution comes equipped with its own stochastic basis. For a fixed stochastic basis $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t=0}^T, \{\beta_k\}_{k\in\mathbb{N}})$ and noise coefficients $\{\sigma_k\}_{k\in\mathbb{N}}$, one can use the linearity of $f \to \operatorname{div}_v(f\sigma_k \circ \dot{\beta}_k)$ together with a method of Pardoux [104] to make a direct passage to the limit on both sides of the equation. However, for stochastically weak solutions, the Brownian motions are not fixed, and the mapping $(f, \beta_k) \mapsto \operatorname{div}_v(f\sigma_k \circ \dot{\beta}_k)$ is nonlinear, prohibiting the passage of weak limits. In this situation, a martingale method is used to overcome this difficulty and produce another weak martingale solution with a new stochastic basis. This result is detailed in Proposition 7.3.5.

Section 7.3.3 is devoted to renormalizing weak martingale solutions to stochastic kinetic equations. The technique of renormalization of *deterministic* transport equations originates from the now classical results of Di'Perna and Lions [34], where they were able to show uniqueness to certain linear transport equations when the drift has lower regularity that the classical theory of characteristics would allow. Formally, the strategy is as follows: if f satisfies (7.17) and $\Gamma : \mathbb{R} \to \mathbb{R}$ is a smooth renormalization, then $\Gamma(f)$ satisfies

$$\begin{aligned} &\partial_t \Gamma(f) + v \cdot \nabla_x \Gamma(f) + \operatorname{div}_v(\Gamma(f)\sigma_k \circ \dot{\beta}_k) = \Gamma'(f)g, \\ &\Gamma(f)|_{t=0} = \Gamma(f_0). \end{aligned} \tag{7.18}$$

If one can justify such a computation, then upon integrating both sides of the equation (7.18) for certain non-negative choices of $\Gamma(z)$ that vanish only at $z = 0$, for instance $\Gamma(z) = z/(1+z)$, then one can get explicit bounds on $\Gamma(f)$ in terms of the initial data, which, by linearity, implies uniqueness. However, since we are working with analytically weak solutions to (7.17), this formal calculation may fail if the individual σ_k are too rough. In particular (to our knowledge), only requiring the L^∞ coloring hypotheses (H1), (H2) are insufficient. The ability to renormalize stochastic kinetic transport equations will turn out to be a crucial property in the final stages of main existence proof. However, as in the case of the deterministic Boltzmann equation, it does not imply uniqueness of the equation, due to the nonlinear nature of the equation.

Our strategy in Section 7.3.3 uses the method of DiPerna and Lions reduces the renormalizability of stochastic kinetic equations to the vanishing of certain commutators between smoothing operators and the differential action of the rough vector fields. Specifically, given a smooth renormalization $\Gamma(z)$ with bounded first and second derivatives, we begin by smoothing a solution f to (7.17) in the (x, v) variables with mollifier η_ϵ. The regularity improvement allows us to renormalize the equation by Γ at the expense of a remainder $R_\epsilon(f)$ comprised of commutators and double commutators of $\sigma_k \cdot \nabla_v$ and convolution by η_ϵ,

$$[\eta_\epsilon, \sigma_k \cdot \nabla_v](f), \quad \big[[\eta_\epsilon, \sigma_k \cdot \nabla_v], \sigma_k \cdot \nabla_v\big](f).$$

As is well known from the classical theory of renormalization by [34] that the single commutator

$$[\eta_\epsilon, \sigma_k \cdot \nabla_v](f) \xrightarrow[\epsilon \to 0]{} 0 \quad \text{in} \quad L^r_{x,v}$$

as long as $\sigma \in W^{1,q}_{x,v}$ and $f \in L^p$ with $1/r = 1/p + 1/q$. As it turns out, the double commutator also vanishes

$$\big[[\eta_\epsilon, \sigma_k \cdot \nabla_v], \sigma_k \cdot \nabla_v\big](f) \xrightarrow[\epsilon \to 0]{} 0 \quad \text{in} \quad L^1_{x,v}$$

provided that $\sigma_k \in W^{1,\frac{2p}{p-1}}_{x,v}$ and $\sigma_k \cdot \nabla_v \sigma_k \in W^{1,\frac{p}{p-1}}_{x,v}$. However one of the primary differences between the deterministic and stochastic theory is an interesting consequence of Itô's formula. Specifically the remainder $R_\epsilon(f)$ involves the square of the single commutator $[\eta_\epsilon, \sigma_k \cdot \nabla_v](f)$. Due to the limited integrability and regularity of f, this imposes that $p \geq 2$ and $\sigma_k \in W^{1,\frac{2p}{p-2}}_{x,v}$ for this contribution to vanish in L^1 (see Proposition 7.3.8 for more details on this). Based on this method of proof, we are presently unable to treat the case $p \in [1,2)$. The main result of this section (Proposition 7.3.8) shows show that a weak martingale solution $f \in L^p(\Omega \times [0,T] \times \mathbb{R}^{2d})$, $p \geq 2$ to (7.17) is renormalizable provided we have the following regularity conditions on σ,

$$\sigma \in \ell^2(\mathbb{N}; W^{1,\frac{2p}{p-2}}_{x,v}) \quad \text{and} \quad \sigma \cdot \nabla_v \sigma \in \ell^1(\mathbb{N}; W^{1,\frac{p}{p-1}}_{x,v}). \tag{7.19}$$

We believe these results are consistent with the work of Lions/Le-Bris [19] on deterministic parabolic equations with rough diffusion coefficients. There should also be a connection with the more recent work of Bailleul/Gubinelli [8]. In the case that $f \in L^{\infty-}(\Omega \times [0,T] \times \mathbb{R}^{2d})$, the conditions (7.19) become precisely the assumptions (H3) and (H4) on the noise coefficients.

Section 7.4 concerns the subtle regularizing effects for stochastic kinetic equations. These are captured by studying the velocity averages of the solution, and have a long history in the deterministic literature [18, 66, 67, 73] as well as several more recent results in the SPDE literature [32, 61, 87]. Since equation (7.17) is of transport type, without more information on g, one does not expect to obtain any further regularity on the solution f than is present in the initial data f_0. However, in view of the deterministic theory it is natural to expect a small gain in the regularity of velocity averages $\langle f, \phi \rangle = \int_{\mathbb{R}^d} f\phi \mathrm{d}v$, where $\phi \in C_c^\infty(\mathbb{R}_v^d)$ is a test function in velocity only. Using a Fourier method of Bouchut/Desvillette [18], we prove that if f is a weak martingale solution to (7.17) and $f, g \in L^2(\Omega \times [0,T] \times \mathbb{R}^{2d})$, then $\langle f, \phi \rangle$ enjoys the following regularity estimate,

$$\mathbf{E}\|\langle f, \phi \rangle\|^2_{L^2_t(H^{1/6}_x)} \leq C_{\phi,\sigma}\big(\|f_0\|^2_{L^2_{x,v}} + \mathbf{E}\|f\|^2_{L^2_{t,x,v}} + \mathbf{E}\|g\|^2_{L^2_{t,x,v}}\big).$$

Combining this with a standard control on oscillations in time, one expects to obtain a form of strong compactness on the velocity averages. To formulate this directly in terms of f rather than its velocity averages, we introduce a topological vector space $L^p_{t,x}(\mathcal{M}^*_v)$ consisting of the space of $L^p_{t,x}$ functions taking values in the space of Radon measures $\mathcal{M}^*_v$ on $\mathbb{R}^d_v$ endowed with it's weak-$\star$ topology. The topology is designed so that sequential convergence in $L^p_{t,x}(\mathcal{M}^*_v)$ corresponds exactly to strong $L^p_{t,x}$ convergence of each sequence of velocity averages. We prove a characterization of compact sets in $L^p_{t,x}(\mathcal{M}^*_v)$ in the appendix. Using the regularity gain in L^2, we exhibit a sufficient criterion for a sequence $\{f_n\}_{n\in\mathbb{N}}$ of weak martingale solutions to a stochastic kinetic equation driven by $\{g_n\}_{n\in\mathbb{N}}$ to induce tight laws on $[L^2_{t,x}(\mathcal{M}^*_v)]_{\text{loc}}$. However, for applications to Boltzmann, one is mostly interested in the case where $\{g_n\}_{n\in\mathbb{N}}$ is only known to be uniformly bounded in $L^1(\Omega\times[0,T]\times\mathbb{R}^{2d})$, due to the very limited control provided by the a priori bounds on the renormalized collision operator $f\to\Gamma'(f)\mathcal{B}(f,f)$. The criteria for tightness in $L^1_{t,x}(\mathcal{M}^*_v)$ is the main result of Section 7.4. As in the deterministic setting (see [66, 67]), there is no easily quantifiable regularity gain for $f,g\in L^1(\Omega\times[0,T]\times\mathbb{R}^{2d})$, making the analysis more involved. At present, we can only treat well-prepared sequences of approximations for which the solution f_n and the source g_n are somewhat better behaved for fixed $n\in\mathbb{N}$. This is captured by Hypothesis 7.4.1.

At this point in the article, we have completed our analysis of the linear problem and proceed to apply our results from Section $3-4$ in the context of (SB). This begins in Section 5 with a construction of a sequence of approximations $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ satisfying a stochastic transport equation driven by a truncated collision operator

$$\mathcal{B}_n(f,f)=\frac{\widehat{\mathcal{B}}_n(f,f)}{(1+n^{-1}\langle f,1\rangle)}.$$

This truncation was introduced in [36] to make $\mathcal{B}_n(f,f)$ Lipschitz in $L^1_{t,x,v}$ while preserving it's conservation properties. After smoothing the noise coefficients and activating only finitely many Brownian motions, we obtain existence by way of the stochastic flow representation of Kunita [80], in combination with a fixed point argument. The main subtleties in comparison to the deterministic theory are due to the fact that the flow map is not explicit. To obtain the a priori bounds

$$\sup_n \mathbf{E}\|(1+|x|^2+|v|^2+|\log \tilde{f}_n|)\tilde{f}_n\|^p_{L^\infty_t(L^1_{x,v})} \le C_p, \quad \sup_n \mathbf{E}\|\mathcal{D}_n(\tilde{f}_n)\|^p_{L^1_{t,x,v}} \le C_p,$$

we require asymptotic growth estimates for the stochastic flow and a stopping time argument. A similar difficulty arises in the work of Hofmanova [93]. An additional difference with the deterministic theory is that we do not prove that our approximations are of Schwartz class in position and velocity. Instead, we use our renormalization lemma to establish the moment and entropy identities used in Section 7.2.

Let us now discuss the main features of the existence proof for Theorem 7.1.7 and some of the main difficulties. The main goal in sections $6-8$ is to extract an appropriate limit point f on a well prepared stochastic basis $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t=0}^T, \{\beta_k\}_{k\in\mathbb{N}})$ and verify that f is indeed a renormalized martingale solution to (SB). This requires a somewhat involved combination of the renormalization and stochastic velocity averaging lemmas together with the general line of arguments introduced by DiPerna and Lions [36] and a later work of Lions [86]. The argument requires a careful interpretation in the stochastic framework. We study the laws of the sequence $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ and use the velocity averaging and renormalization lemmas to show they are tight on $L^1_{t,x}(\mathcal{M}^*_v)\cap C_t([L^1_x]_w)$. A generalization of the Skorohod theorem due to Jakubowski [74] and Vaart/Wellner [116] gives a candidate limit f, which we endeavor to show is a renormalized martingale solution to (SB). The Skorohod theorem allows one to gain compactness of the nonlinear drift terms at the expense of the noise terms. Indeed, additional oscillations are introduced in the noise terms after switching probability spaces as $\operatorname{div}_v(\sigma_k^n f_n \dot{\beta}_k)$ is replaced by $\operatorname{div}_v(\sigma_k^n \tilde{f}_n \dot{\tilde{\beta}}_k^n)$, at which point we are setup to apply our weak stability result. However, this is done in a somewhat indirect way.

The procedure of identifying f with a solution of (RSB) requires two conceptually different steps. First, in Section 6 we fix a bounded renormalization Γ_m which converges to the identity as $m \to \infty$. With m fixed, we check the criterion necessary to apply our weak stability result to the sequence $\{\Gamma_m(f_n)\}_{n\in\mathbb{N}}$. This sequence is also shown to induce tight laws on $L^1_{t,x}(\mathcal{M}^*_v) \cap C_t([L^1_x]_w)$. The stability result implies its limit point $\overline{\Gamma}_m$ is a solution to a stochastic kinetic equation with a driver $\mathcal{B}_m$. To show this requires analysis of the laws induced by the sequence of renormalized collision operators $\{\Gamma_m'(f_n)\mathcal{B}_n(f_n, f_n)\}_{n\in\mathbb{N}}$.

At this stage, we do not yet have any sort of closed evolution equation for $\overline{\Gamma_m(f)}$. Indeed, it is unclear the relation between $\overline{\Gamma}_m$ and $\mathcal{B}_m$. Hence, our next step is to pass $m \to \infty$ and hope to obtain a closed evolution equation in the limit. As a result of the initial renormalization procedure $\overline{\Gamma_m(f)}$ converges strongly to f in $L^\infty_t(L^1_{x,v})$, $\mathbf{P}$ almost surely. Unfortunately, as $m \to \infty$ one does not have any good control on $\{\mathcal{B}_m\}_{m\in\mathbb{N}}$ in any space of distributions (only in the topology of measurable functions, which does not play well with the weak form). On the other hand, we do have control of $\{(1+\overline{\Gamma_m(f)})^{-1}\mathcal{B}_m\}_{m\in\mathbb{N}}$. Hence, the strategy is to renormalize again, this time with $\log(1+z)$, and apply again our stability result in the limit $m \to \infty$.

Section 7.7 is dedicated to analysis of the renormalized collision operator $\mathcal{B}_m$. As in the deterministic setting, we are able to obtain a pointwise (in Ω) continuity result

$$\frac{\mathcal{B}_n(f_n,f_n)}{1+\langle f_n,1\rangle} \to \frac{\mathcal{B}(f,f)}{1+\langle f,1\rangle} \quad \text{in} \quad L^1_{t,x}(\mathcal{M}^*_v),$$

as a consequence of the velocity averaging lemmas. Following the strategy in [86] and [65], we are able to conclude that

$$\frac{\overline{\mathcal{B}_m}}{1+\overline{\Gamma_m(f)}} \to \frac{\mathcal{B}(f,f)}{1+f} \quad \text{in} \quad L^2(\Omega;[L^1_{t,x,v}]_w),$$

allowing us to apply again the stability result.

We are then able to deduce that $\log(1+f)$ is a solution to a stochastic kinetic equation driven by $(1+f)^{-1}\mathcal{B}(f,f)$. Roughly speaking, the final step is verify the renormalized form of (SB) with an arbitrary renormalization. Since $\log(1+f) \in L^{\infty-}(\Omega\times[0,T]\times\mathbb{R}^{2d})$, the conditions on the noise coefficients (H3) and (H4) are exactly such that the renormalization Lemma 7.3.8 applies. This completes the proof.

Preliminaries

Notation

To simplify the appearance of the function spaces used in this paper, we will use a number of abbreviations. The notation $L^q_t(L^p_{x,v})$ denotes the space $L^q([0,T];L^p(\mathbb{R}^{2d}))$, and $L^p_{t,x,v}$ is short for $L^p([0,T]\times\mathbb{R}^{2d})$, with similar notation for Sobolev spaces. A Banach space B endowed with its weak topology is denoted $[B]_w$, and the space of weakly continuous functions from $[0,T]$ to B will be written as $C_t([B]_w)$. Finally, $[L^p_{t,x,v}]_{\text{loc}}$ denotes the space of locally integrable functions endowed with the natural topology of locally convex seminorms.

For a given probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and a Banach space B, we will denote by $L^p(\Omega; B)$ the measurable maps (random variables) from $\mathcal{F}$ to the Borel sigma algebra on B with p^{th} integrable norm. The space $L^{\infty-}(\Omega; B)$ consists of random variables belonging to $L^p(\Omega; B)$ for all $p \in [1, \infty)$.

Basic properties of the collision operator

In this section, we recall some basic properties of the collision operator $f \to \mathcal{B}(f, f)$ (defined in (7.3)) which will be used throughout the article. A more in-depth discussion can be found in [36]. To begin, we note that the collision operator may be split into gain and loss terms

$$\mathcal{B}(f, f) = \mathcal{B}^+(f, f) - \mathcal{B}^-(f, f),$$

with

$$\mathcal{B}^+(f, f) = \iint_{\mathbb{R}^d \times \mathbb{S}^{d-1}} f' f'_* b(v - v_*, \theta) \mathrm{d}\theta \mathrm{d}v_*, \quad \mathcal{B}^-(f, f) = f(\bar{b} * f),$$

and $\bar{b}$ defined by

$$\bar{b}(z) = \int_{\mathbb{S}^{d-1}} b(z, \theta) \mathrm{d}\theta.$$

The following inequality due to Arkeryd [4] relates the positive and negative parts of the collision operator through the entropy dissipation. Namely, for $K > 1$ and $f \in L^1_v$, it holds

$$\mathcal{B}^+(f, f)(v) \leq K \mathcal{B}^-(f, f)(v) + \frac{1}{\log K} \mathcal{D}^0(f)(v), \tag{7.20}$$

where $\mathcal{D}^0(f)$ is defined by

$$\mathcal{D}^0(f) = \frac{1}{4}\iint_{\mathbb{R}^d\times\mathbb{S}^{d-1}} d(f)b(v-v_*,\theta)\mathrm{d}\theta\mathrm{d}v_*.$$

Note that the quantity $\mathcal{D}^0(f)$ is *not* the entropy dissipation $\mathcal{D}(f)$ as defined in (7.8), but is instead related to $\mathcal{D}(f)$ by an integration in v,

$$\mathcal{D}(f) = \int_{\mathbb{R}^d} \mathcal{D}^0(f)\mathrm{d}v.$$

Formal a priori estimates

In this section, we will derive formal a priori estimates on the stochastic Boltzmann equation (SB) with $\{\sigma_k\}_{k\in\mathbb{N}}$ satisfying (H1) and (H2) and initial data f_0 satisfying

$$\|(1+|x|^2+|v|^2+|\log f_0|)f_0\|^p_{L^1_{x,v}} < \infty.$$

Specifically we will see that under these assumptions, there exists a positive constant $C \equiv C_{p,\sigma,T,f_0}$, depending on p, $\{\sigma_k\}_{k\in\mathbb{N}}$, T, and f_0 such that

$$\mathbf{E}\|(1+|x|^2+|v|^2+|\log f|)f\|^p_{L^\infty_t(L^1_{x,v})} \le C.$$

In addition the entropy dissipation $\mathcal{D}(f)$ satisfies

$$\mathbf{E}\|\mathcal{D}(f)\|^p_{L^1_{t,x}} \le C. \tag{7.21}$$

These a priori estimates are completely natural in the context of the deterministic Boltzmann equation and correspond to the physical assumptions of finite mass, momentum, energy, entropy, and entropy production (see for instance [24] or [65]).

Throughout the argument C will denote a positive, finite constant that depends on p, $\{\sigma_k\}_{k\in\mathbb{N}}$, T and f_0. It may change from line to line, and even within a line.

Moment Bound

We begin by showing that

$$\mathbf{E}\|(1+|x|^2+|v|^2)f\|^p_{L^\infty_t(L^1_{x,v})} \le C, \tag{7.22}$$

for $p>2$. To this end, we multiply the Boltzmann equation by $(1+|x|^2+|v|^2)$ in Itô form and integrate over $[0,t]\times\mathbb{R}^d_x\times\mathbb{R}^d_v$ to obtain

$$\begin{aligned}\frac{1}{2}\iint_{\mathbb{R}^{2d}}(1+|x|^2+|v|^2)f_t\,\mathrm{d}v\mathrm{d}x &= \frac{1}{2}\iint_{\mathbb{R}^{2d}}(1+|x|^2+|v|^2)f_0\,\mathrm{d}v\mathrm{d}x \\ &+\int_0^t\iint_{\mathbb{R}^{2d}}\sum_{k\in\mathbb{N}}|\sigma_k|^2 f_s\,\mathrm{d}x\mathrm{d}v\mathrm{d}s \\ &+\int_0^t\iint_{\mathbb{R}^{2d}}\Big(\sum_{k\in\mathbb{N}}(\sigma_k\cdot\nabla_v\sigma_k)+x\Big)\cdot vf_s\,\mathrm{d}x\mathrm{d}v\mathrm{d}s \\ &+\sum_{k\in\mathbb{N}}\int_0^t\Big(\int_{\mathbb{R}^{2d}}v\cdot\sigma_k\,f_s\,\mathrm{d}x\mathrm{d}v\Big)\mathrm{d}\beta_k(s).\end{aligned} \tag{7.23}$$

Applying Cauchy-Schwartz to the time integral the following estimate readily follows,

$$\begin{aligned}&\Big|\int_0^t\iint_{\mathbb{R}^{2d}}\Big(\sum_{k\in\mathbb{N}}(\sigma_k\cdot\nabla_v\sigma_k)+x\Big)\cdot vf_s\,\mathrm{d}x\mathrm{d}v\mathrm{d}s\Big|^p \\ &\qquad\le C\|\sigma\cdot\nabla_v\sigma\|^p_{\ell^1(\mathbb{N};L^\infty_{x,v})}\int_0^t\|(1+|x|^2+|v|^2)f_s\|^p_{L^1_{x,v}}\mathrm{d}s,\end{aligned} \tag{7.24}$$

and similarly

$$\Big|\int_0^t\int_{\mathbb{R}^{2d}}\sum_{k\in\mathbb{N}}|\sigma_k|^2 f_s\,\mathrm{d}x\mathrm{d}v\mathrm{d}s\Big|^p \le C\int_0^t\|(1+|x|^2+|v|^2)f_s\|^p_{L^1_{x,v}}\,\mathrm{d}s. \tag{7.25}$$

For the stochastic integral term in (7.23), the BDG (Burkholder-Davis-Gundy) inequality yields

$$\mathbf{E}\left|\sup_{r\in[0,t]}\sum_{k\in\mathbb{N}}\int_0^r\Big(\int_{\mathbb{R}^{2d}}v\cdot\sigma_k\,f_s\,\mathrm{d}v\mathrm{d}x\Big)\mathrm{d}\beta_k(s)\right|^p\le\mathbf{E}\bigg(\int_0^t\sum_{k\in\mathbb{N}}\Big(\int_{\mathbb{R}^{2d}}\sigma_k\cdot v f_s\,\mathrm{d}v\mathrm{d}x\Big)^2\mathrm{d}s\bigg)^{p/2}.$$

Therefore, after another application of Cauchy-Schwartz to the time integral, we conclude

$$\begin{aligned}\mathbf{E}&\left|\sup_{r\in[0,\,t]}\sum_{k\in\mathbb{N}}\int_0^r\Big(\int_{\mathbb{R}^{2d}}v\cdot\sigma_k\,f_s\,\mathrm{d}v\mathrm{d}x\Big)\mathrm{d}\beta_k(s)\right|^p\\&\le C\|\sigma\|^p_{\ell^2(\mathbb{N};L^\infty_{x,v})}\int_0^t\mathbf{E}\|(1+|x|^2+|v|^2)f_s\|^p_{L^1_{x,v}}\mathrm{d}s\end{aligned}\tag{7.26}$$

We may now combine estimates (7.24), (7.25) and (7.26) with (7.23) to obtain

$$\mathbf{E}\bigg(\sup_{r\in[0,t]}\|(1+|x|^2+|v|^2)f_r\|_{L^1_{x,v}}\bigg)^p\le C+C\int_0^t\mathbf{E}\bigg(\sup_{r\in[0,\,s]}\|(1+|x|^2+|v|^2)f_r\|_{L^1_{x,v}}\bigg)^p\mathrm{d}s.$$

Whereby Grönwall's Lemma gives (7.22).

Entropy Bound

Next, we show that

$$\mathbf{E}\|f\log f\|^p_{L^\infty_t(L^1_{x,v})}\le C.$$

This estimate, as in the deterministic case, is comprised of two parts, control of the entropy $f\log f$ from above by the entropy dissipation (7.7) and control of $f\log f$ from below using estimates (7.22) and a Maxwellian. Specifically, integrating the entropy dissipation law (7.7) in $[0,t]\times\mathbb{R}^d_x$ gives the $\mathbf{P}$ almost sure identity, for each $t\in[0,T]$,

$$\int_{R^{2d}}f_t\log f_t\mathrm{d}v\mathrm{d}x=\int_{R^{2d}}f_0\log f_0\mathrm{d}v\mathrm{d}x-\int_0^t\int_{\mathbb{R}^d}\mathcal{D}(f_s)\mathrm{d}x\mathrm{d}s,\tag{7.27}$$

and since $\mathcal{D}(f) \geq 0$, this yields the classical entropy inequality,

$$\int_{\mathbb{R}^{2d}} f_t \log f_t \mathrm{d}v\mathrm{d}x \leq \int_{\mathbb{R}^{2d}} f_0 \log f_0 \, \mathrm{d}v\mathrm{d}x.$$

Using this and standard estimates from kinetic theory (see [24]), we obtain $\mathbf{P}$ almost surely

$$\begin{aligned}\int_{\mathbb{R}^{2d}} f_t |\log f_t| \mathrm{d}v\mathrm{d}x &\leq \int_{R^{2d}} f_t \log f_t \mathrm{d}x\mathrm{d}v + 2\int_{\mathbb{R}^{2d}} (|x|^2 + |v|^2) f_t \mathrm{d}v\mathrm{d}x \\ &\quad + 2\frac{\log e}{e} \int_{\mathbb{R}^{2d}} e^{-\frac{1}{2}(|x|^2+|v|^2)} \, \mathrm{d}v\mathrm{d}x \\ &\leq \|f_0 \log f_0\|_{L^1_{x,v}} + C\|(1 + |x|^2 + |v|^2) f_t\|_{L^1_{x,v}} + C.\end{aligned}$$

Applying the previous estimate on $(1 + |x|^2 + |v|^2)f$ to the above inequality gives the desired estimate of $f \log f$.

Dissipation Bound

Finally with regard to the entropy dissipation estimate (7.21), observe that equation (7.27) also implies the $\mathbf{P}$ almost sure bound

$$\|\mathcal{D}(f)\|_{L^1_{t,x}} \leq \|f \log f\|_{L^1_{t,x,v}} + \|f_0 \log f_0\|_{L^1_{x,v}},$$

from which the estimate (7.21) clearly follows.

Stochastic Kinetic Transport Equations

In this section, we assume that a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ is given, together with a deterministic initial condition $f_0 \in L^1_{x,v}$ and a random variable $g \in L^1(\Omega; L^1_{t,x,v})$. Moreover, we have a collection of noise coefficients $\{\sigma_k\}_{k\in\mathbb{N}}$ satisfying the coloring Hypothesis 7.1.2. We analyze properties of solutions to stochastic kinetic equations of the type

$$\begin{aligned} &\partial_t f + v \cdot \nabla_x f + \mathrm{div}_v(f\sigma_k \circ \dot{\beta}_k) = g \\ &f\mid_{t=0}= f_0, \end{aligned} \tag{7.28}$$

where solutions are understood in the weak martingale sense, given precisely in Definition 7.3.1 below.

Weak martingale solutions

Definition 7.3.1 (Weak Martingale Solution)**.** A process $f : [0,T] \times \Omega \to L^1_{x,v}$ is a weak martingale solution of the stochastic kinetic transport equation driven by g with initial data f_0, provided the following is true:

1. For all $\varphi \in C_c^\infty(\mathbb{R}^{2d})$, the process $\langle f, \varphi\rangle : \Omega \times [0,T] \to \mathbb{R}$ admits $\mathbf{P}$ a.s. continuous sample paths. Moreover, f belongs to $L^2(\Omega; L^\infty_t(L^1_{x,v}))$.

2. There exists a collection of Brownian motions $\{\beta_k\}_{k\in\mathbb{N}}$ and a filtration $(\mathcal{F}_t)_{t=0}^T$ where the filtration $(\mathcal{F}_t)_{t=0}^T$ is generated by the $[L^1_{x,v}]_w$ valued processes $(f_t)_{t=0}^T$, $(\int_0^t g_s ds)_{t=0}^T$ and each Brownian motion $(\beta_k(t))_{t=0}^T$.

3. For all test functions $\varphi \in C_c^\infty(\mathbb{R}^{2d})$, the process $(M_t(\varphi))_{t=0}^T$ defined by

$$M_t(\varphi) = \iint_{\mathbb{R}^{2d}} f_t\varphi \mathrm{d}x\mathrm{d}v - \iint_{\mathbb{R}^{2d}} f_0\varphi \mathrm{d}x\mathrm{d}v - \int_0^t \iint_{\mathbb{R}^{2d}} f(v\cdot\nabla_x\varphi + \mathcal{L}_\sigma\varphi) + g\varphi \, \mathrm{d}x\mathrm{d}v\mathrm{d}s \tag{7.29}$$

   is an $(\mathcal{F}_t)_{t=0}^T$ martingale. Moreover, its quadratic variation and cross variation with respect to each β_k are given by:

$$\langle\!\langle M(\varphi), M(\varphi) \rangle\!\rangle_t = \sum_{k\in\mathbb{N}} \int_0^t \Big(\iint_{\mathbb{R}^{2d}} f_s \sigma_k \cdot \nabla_v \varphi \mathrm{d}x\mathrm{d}v \Big)^2 \mathrm{d}s.$$

$$\langle\!\langle M(\varphi), \beta_k \rangle\!\rangle_t = \int_0^t \iint_{\mathbb{R}^{2d}} f_s \sigma_k \cdot \nabla_v \varphi \mathrm{d}x\mathrm{d}v\mathrm{d}s.$$

Remark 7.3.2. Note that if f is a martingale solution to a stochastic kinetic equation driven by g and starting from f_0 relative to the stochastic basis $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t=0}^T, \{\beta_k\}_{k\in\mathbb{N}})$, then for all $t \in [0,T]$ the following identity holds $\mathbf{P}$ almost surely

$$\begin{aligned} \iint_{\mathbb{R}^{2d}} f_t \varphi \mathrm{d}x\mathrm{d}v = \iint_{\mathbb{R}^{2d}} f_0 \varphi \mathrm{d}x\mathrm{d}v + \int_0^t \iint_{\mathbb{R}^{2d}} [f_s(v\cdot\nabla_x + \mathcal{L}_\sigma)\varphi + g_s\varphi] \mathrm{d}x\mathrm{d}v\mathrm{d}s \\ + \sum_{k\in\mathbb{N}} \int_0^t \iint_{\mathbb{R}^{2d}} f_s\sigma_k \cdot \nabla_v \varphi \mathrm{d}x\mathrm{d}v\mathrm{d}\beta_k(s). \end{aligned} \tag{7.30}$$

This is guaranteed by Lemma B.1.13 of the appendix.

The following existence result may be proved with a small modification to the arguments given in [53] (which use a strategy developed already in the Ph.D thesis of E. Pardoux [104]).

Theorem 7.3.3 (Existence)**.** *Let $\{\beta_k\}_{k\in\mathbb{N}}$ be a given collection of $(\mathcal{F}_t)_{t=0}^T$ Brownian motions. Assume that $g \in L^\infty(\Omega; L^1_{t,x,v} \cap L^\infty_{t,x,v})$, and $(\int_0^t g_s \mathrm{d}s)_{t=0}^T$ is an $(\mathcal{F}_t)_{t=0}^T$ adapted process. Then there exists a weak martingale solution f (relative to the given stochastic basis) to the stochastic kinetic equation driven by g with initial data f_0. Moreover, we have the following estimate for every $p \in [1,\infty)$,*

$$\mathbf{E}\|f\|^p_{L^p_{t,x,v}} \lesssim \|f_0\|^p_{L^p_{x,v}} + \mathbf{E}\|g\|^p_{L^p_{t,x,v}}.$$

The next result is a time regularity estimate.

Lemma 7.3.4. *Let $q \in (2,\infty]$ and assume $f \in L^{\infty-}(\Omega; L^q_t(L^1_{x,v}))$ is a weak martingale solution to the stochastic kinetic transport equation driven by $g \in L^{\infty-}(\Omega; L^q_t(L^1_{x,v}))$ with with initial data $f_0 \in L^1_{x,v}$. Then for any test function $\varphi \in C^\infty_c(\mathbb{R}^{2d})$ and $p \in (\frac{2q}{q-2}, \infty)$, we have the following estimate*

$$\mathbf{E}\|\langle f, \varphi\rangle\|^p_{W^{\gamma,p}_t} \le C_{\varphi,\sigma}\Big(\mathbf{E}\|f\|^p_{L^q_t(L^1_{x,v})} + \mathbf{E}\|g\|^p_{L^q_t(L^1_{x,v})}\Big),$$

where $\gamma = \frac{1}{2} - \frac{1}{p} - \frac{1}{q}$.

Proof. Consider two times $t, s \in \mathbb{R}_+$, $t \neq s$. Writing (7.11) in Itô form, we can conclude that the difference $\langle f_t - f_s, \varphi\rangle$ satisfies

$$\begin{aligned}\langle f_t - f_s, \varphi\rangle = \int_s^t \iint_{\mathbb{R}^{2d}} (v\cdot\nabla\varphi + \mathcal{L}_\sigma\varphi) f \, dx\mathrm{d}v\mathrm{d}r + \int_s^t \iint_{\mathbb{R}^{2d}} \varphi g \,\mathrm{d}x\mathrm{d}v\mathrm{d}s \\ + \sum_{k\in\mathbb{N}} \int_s^t \Big(\iint_{\mathbb{R}^{2d}} f\sigma_k \cdot \nabla_v\varphi \,\mathrm{d}x\mathrm{d}v\Big)\mathrm{d}\beta_k(r).\end{aligned}$$

We would like to estimate $\mathbf{E}|\langle f_t - f_s, \varphi\rangle|^p$. To this end, since $v\cdot\nabla\varphi + \mathcal{L}_\sigma\varphi \in L^\infty_{x,v}$, we have the estimate

$$\Big|\int_s^t \iint_{\mathbb{R}^{2d}} (v\cdot\nabla\varphi + \mathcal{L}_\sigma\varphi) f \, dx\mathrm{d}v\mathrm{d}r\Big|^p \le C_{\varphi,\sigma}|t-s|^{p(1-\frac{1}{q})}\|f\|^p_{L^q_t(L^1_{x,v})},$$

and similarly

$$\Big|\int_s^t \varphi g \,\mathrm{d}x\mathrm{d}v\mathrm{d}s\Big|^p \le C_\varphi |t-s|^{p(1-\frac{1}{q})}\|g\|^p_{L^q_t(L^1_{x,v})}.$$

By the BDG inequality we may estimate the martingale term by

$$\begin{aligned}\mathbf{E}\Big|\sum_{k\in\mathbb{N}} \int_s^t \Big(\iint_{\mathbb{R}^{2d}} f\sigma_k\cdot\nabla\varphi \mathrm{d}x\mathrm{d}v\Big)\mathrm{d}\beta_k(r)\Big|^p &\le \mathbf{E}\Big(\int_s^t \sum_{k\in\mathbb{N}} \Big(\iint_{\mathbb{R}^{2d}} f\sigma_k\cdot\nabla\varphi \mathrm{d}x\mathrm{d}v\Big)^2 \mathrm{d}r\Big)^{p/2} \\ &\le C_{\varphi,\sigma}|t-s|^{\frac{p}{2}(1-\frac{2}{q})}\mathbf{E}\|f\|^p_{L^q_t(L^1_{x,v})}.\end{aligned}$$

Combining these estimates gives

$$\mathbf{E}|\langle f_t - f_s, \varphi\rangle|^p \le C_{\varphi,\sigma}|t-s|^{p(\frac{1}{2}-\frac{1}{q})}\Big(|t-s|^{\frac{p}{2}}\big(\mathbf{E}\|f\|^p_{L^q_t(L^1_{x,v})} + \mathbf{E}\|g\|^p_{L^q_t(L^1_{x,v})}\big) + \mathbf{E}\|f\|^p_{L^q_t(L^1_{x,v})}\Big).$$

We now estimate the regularity of $\langle f,\varphi\rangle$ via the Sobolev-Slobodeckij semi-norm $[\cdot]_{W_t^{\gamma,p}}$. For

$\gamma p+1=p(\frac{1}{2}-\frac{1}{q})$ we find

$$\mathbf{E}[\langle f,\varphi\rangle]^p_{W_t^{\gamma,p}}=\int_0^T\int_0^T\frac{\mathbf{E}|\langle f_t-f_s,\varphi\rangle|^p}{|t-s|^{\gamma p+1}}\mathrm{d}s\mathrm{d}t\leq C_{\varphi,\sigma}\Big(\mathbf{E}\|f\|^p_{L_t^\infty(L^1_{x,v})}+\|g\|^p_{L_t^\infty(L^1_{x,v})}\Big).$$

□

Stability of weak martingale solutions

In this section, we establish our main stability result for sequences of weak martingale solutions to stochastic kinetic equations. The result below will be used repeatedly throughout the article.

Proposition 7.3.5. *Let $f:\Omega\times[0,T]\to L^1_{x,v}$ be a stochastic process and $\{\beta_k\}_{k\in\mathbb{N}}$ be a collection of Brownian motions. Assume there exists a sequence of processes $\{f_n\}_{n\in\mathbb{N}}$ with the following properties.*

1. *For each $n\in\mathbb{N}$ there exist g_n, f_n^0, and σ_n such that f_n is a weak martingale solution to a stochastic kinetic equation driven by g_n with initial data f_n^0, relative to the noise coefficients $\sigma^n=\{\sigma_k^n\}_{k\in\mathbb{N}}$ and the stochastic basis $(\Omega,\mathcal{F},\mathbf{P},(\mathcal{F}_t^n)_{t=0}^T,\{\beta_k^n\}_{k\in\mathbb{N}})$.*

2. *The sequences $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ are bounded in $L^2(\Omega;L_t^\infty(L^1_{x,v}))$ and $L^2(\Omega;L^1_{t,x,v})$ respectively. Moreover, for each $\varphi\in C_c^\infty(\mathbb{R}^{2d})$,*

$$\langle f_n,\varphi\rangle\to\langle f,\varphi\rangle\quad in\quad L^2(\Omega;C_t),\tag{7.31}$$

and for each $t \in [0,T]$,

$$\left\langle \int_0^t g_n(s)ds, \varphi \right\rangle \to \left\langle \int_0^t g(s)ds, \varphi \right\rangle \quad in \quad L^2(\Omega). \tag{7.32}$$

3. *As $n \to \infty$, the following convergences hold:*

$$\{\beta_k^n\}_{k\in\mathbb{N}} \to \{\beta_k\}_{k\in\mathbb{N}} \quad in \quad L^2\left(\Omega; [C_t]^\infty\right).$$

$$f_n^0 \to f^0 \quad in \quad L^1_{x,v}.$$

$$\sigma^n \to \sigma \quad in \quad \ell^2(\mathbb{N}; L^\infty_{x,v}).$$

$$\sigma^n \cdot \nabla_v \sigma^n \to \sigma \cdot \nabla_v \sigma \quad in \quad \ell^1(\mathbb{N}; L^\infty_{x,v}).$$

Under these hypotheses, we may deduce that f is a weak martingale solution driven by g and starting from f_0, relative to the noise coefficients σ and the Brownian motions $\{\beta_k\}_{k\in\mathbb{N}}$.

Moreover, if $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t^n)_{t=0}^T, \{\beta_k^n\}_{k\in\mathbb{N}})$ is independent of $n \in \mathbb{N}$, then f can be built with respect to the same stochastic basis.

Proof. Define a collection of topological spaces $(E_t)_{t=0}^T$ by $E_t = C\big([0,t]; [L^1_{x,v}]^2_w\big) \times C[0,t]^\infty$. Let $r_t : E_T \to E_t$ be the corresponding restriction operators. Next define the $L^1_{x,v}$ valued processes $(G_t)_{t=0}^T$ and $(G_t^n)_{t=0}^T$ to be the running time integrals (starting from 0) of g and g_n, respectively. Use these to define the E_T valued random variables $X = (f, G, \{\beta_k\}_{k\in\mathbb{N}})$ and $X_n = (f_n, G_n, \{\beta_k^n\}_{k\in\mathbb{N}})$.

We will verify that f is a weak martingale solution relative to the filtration $(\mathcal{F}_t)_{t=0}^T$ given by $\mathcal{F}_t = \sigma(r_t X)$. With this filtration, Part 1 of Definition 7.3.1 certainly holds. Part 2 is true by assumption. Hence, if suffices to verify Part 3. Let $\varphi \in C_c^\infty(\mathbb{R}^{2d})$ and define the continuous process $(M_t(\varphi))_{t=0}^T$ by (7.29). Let $s < t$ be two times and suppose that $\gamma \in C_b(E_s;\mathbb{R})$. It suffices to show

$$\mathbf{E}\Big(\gamma(r_sX)\big(M_t(\varphi)-M_s(\varphi)\big)\Big)=0, \tag{7.33}$$

$$\mathbf{E}\Big(\gamma(r_sX)\big(M_t(\varphi)^2-M_s(\varphi)^2\big)\Big)=\sum_{k\in\mathbb{N}}\mathbf{E}\Big(\gamma(r_sX)\int_s^t\Big(\iint_{\mathbb{R}^{2d}} f\sigma_k\cdot\nabla_v\varphi \mathrm{d}x\mathrm{d}v\Big)^2\mathrm{d}r\Big), \tag{7.34}$$

$$\mathbf{E}\Big(\gamma(r_sX)\big(M_t(\varphi)\beta_k(t)-M_s(\varphi)\beta_k(s)\big)\Big)=\mathbf{E}\Big(\gamma(r_sX)\int_s^t\iint_{\mathbb{R}^{2d}}\sigma_k\cdot\nabla_v\varphi f\mathrm{d}x\mathrm{d}v\mathrm{d}r\Big). \tag{7.35}$$

Begin by defining the filtration $(\mathcal{F}_t^n)_{t=0}^T$ by the relation $\mathcal{F}_t^n=\sigma(r_tX_n)$. Let the $(\mathcal{F}_t^n)_{t=0}^T$ continuous martingale $(M_t^n(\varphi))_{t=0}^T$ defined by (7.29), with f_n, f_n^0, and σ^n replacing f, f^0, and σ. By the first assumption of the Proposition and Definition 7.3.1, we find that

$$\mathbf{E}\Big(\gamma(r_tX_n)\big(M_t^n(\varphi)-M_s^n(\varphi)\big)\Big)=0$$

Passing to a subsequence if necessary, the second and third assumptions of the Proposition imply that for each $t\in[0,T]$, the random variables $\{M_t^n(\varphi)\}_{n\in\mathbb{N}}$ converge to $M_t(\varphi)$ in $L^2(\Omega)$. Indeed, this hinges on the following facts. First, the sequences $\{\langle f_n(t),\varphi\rangle\}_{n\in\mathbb{N}}$ and $\{\langle G_n(t),\varphi\rangle\}_{n\in\mathbb{N}}$ converge to $\langle f(t),\varphi\rangle$ and $\langle G(t),\varphi\rangle$ in $L^2(\Omega)$. Second, the sequence $\{\mathcal{L}_{\sigma^n}\varphi\}_{n\in\mathbb{N}}$ converges to $\mathcal{L}_\sigma\varphi$ in $L^\infty_{x,v}$. A similar argument shows that for each $t\in[0,T]$, the random variables $\{\gamma(r_tX_n)\}_{n\in\mathbb{N}}$ converge to $\gamma(r_tX)$ in $L^{\infty-}(\Omega)$. To treat the sequence $\{G^n\}_{n\in\mathbb{N}}$, we use the fact that if a sequence of continuous functions converges pointwise to a continuous limit, then the convergence is also uniform. With these remarks, we may pass $n\to\infty$ and deduce

(7.33). Next we observe that:

$$\mathbf{E}\Big(\gamma(r_sX_n)\big(M_t^n(\varphi)^2-M_s^n(\varphi)^2\big)\Big)=\sum_{k\in\mathbb{N}}\mathbf{E}\Big(\gamma(r_sX_n)\int_s^t\Big(\iint_{\mathbb{R}^{2d}}f_n\sigma_k^n\cdot\nabla_v\varphi\mathrm{d}x\mathrm{d}v\Big)^2\mathrm{d}r\Big). \tag{7.36}$$

Using the facts mentioned above, we deduce that for each $k\in\mathbb{N}$

$$\begin{aligned}\lim_{n\to\infty}\mathbf{E}\Big(\gamma(r_sX_n)\int_s^t\Big(\iint_{\mathbb{R}^{2d}}f_n\sigma_k^n\cdot\nabla_v\varphi\mathrm{d}x\mathrm{d}v\Big)^2\mathrm{d}r\Big)\\ =\mathbf{E}\Big(\gamma(r_sX)\int_s^t\Big(\iint_{\mathbb{R}^{2d}}f\sigma_k\cdot\nabla_v\varphi\mathrm{d}x\mathrm{d}v\Big)^2\mathrm{d}r\Big).\end{aligned}$$

Moreover, we have the inequality

$$\mathbf{E}\Big(\gamma(r_sX_n)\int_s^t\Big(\iint_{\mathbb{R}^{2d}}f_n\sigma_k^n\cdot\nabla_v\varphi\mathrm{d}x\mathrm{d}v\Big)^2\mathrm{d}r\Big)\le\|\gamma\|_{C_b(E_s;\mathbb{R})}\|\nabla_v\varphi\|_{L^\infty_{x,v}}\mathbf{E}\|f_n\|^2_{L^2_t(L^1_{x,v})}\|\sigma_k^n\|^2_{L^\infty_{x,v}}.$$

Since $\{\sigma^n\}_{n\in\mathbb{N}}$ is strongly compact in $\ell^2(\mathbb{N};L^\infty_{x,v})$, it follows that

$$\lim_{N\to\infty}\sup_{n\in\mathbb{N}}\sum_{k=N}^\infty\|\sigma_k^n\|^2_{L^\infty_{x,v}}=0.$$

Since $\{f_n\}_{n\in\mathbb{N}}$ is uniformly bounded in $L^2(\Omega;L^2_t(L^1_{x,v}))$, by splitting the series into finitely many terms plus a uniformly controlled remainder, we find that:

$$\begin{aligned}\lim_{n\to\infty}\sum_{k\in\mathbb{N}}\mathbf{E}\Big(\gamma(r_sX_n)\int_s^t\Big(\iint_{\mathbb{R}^{2d}}f_n\sigma_k^n\cdot\nabla_v\varphi\mathrm{d}x\mathrm{d}v\Big)^2\mathrm{d}r\Big)\\ =\sum_{k\in\mathbb{N}}\mathbf{E}\Big(\gamma(r_sX)\int_s^t\Big(\iint_{\mathbb{R}^{2d}}f\sigma_k\cdot\nabla_v\varphi\mathrm{d}x\mathrm{d}v\Big)^2\mathrm{d}r\Big).\end{aligned}$$

We may now pass $n\to\infty$ on both sides of (7.36) to obtain (7.34). An entirely similar argument yields (7.35). This completes the proof. □

Renormalization

Formally, given a regular solution f to (7.28) and a smooth $\Gamma : \mathbb{R} \to \mathbb{R}$, Ito's formula implies that $\Gamma(f)$ satisfies

$$\begin{aligned} &\partial_t \Gamma(f) + v \cdot \nabla_x \Gamma(f) + \operatorname{div}_v(\Gamma(f)\sigma_k \circ \dot{\beta}_k) = \Gamma'(f)g, \\ &\Gamma(f)|_{t=0} = \Gamma(f_0). \end{aligned} \tag{7.37}$$

However, if we only impose Hypothesis 7.1.2 on the noise coefficients, it is not clear whether (7.37) can be justified when f is only a weak martingale solution to (7.28). In this section, we show that if f has increased local integrability in x, v and σ has sufficient Sobolev regularity, then (7.37) holds relative to a large class of renormalizations Γ. Towards this end, we introduce the notion of renormalized martingale solution to (7.28).

Definition 7.3.6 (Renormalized Martingale Solution)**.** Suppose that $(f_t)_{t=0}^T$ is a weak martingale solution to the stochastic kinetic equation driven by g with initial data f_0 and with with respect to the stochastic basis $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t=0}^T, \{\beta_k\}_{k\in\mathbb{N}})$. We say that $(f_t)_{t=0}^T$ is a renormalized weak martingale solution provided that for all renormalizations $\Gamma \in C^2(\mathbb{R})$ with $\sup_{z\in\mathbb{R}}(|\Gamma'(z)| + |\Gamma''(z)|) < \infty$ and $\Gamma(0) = 0$, the process $(\Gamma(f)_t)_{t=0}^T$ is weak martingale solution to the stochastic kinetic equation driven by $\Gamma'(f)g$ with initial data $\Gamma(f_0)$, and with respect to the same stochastic basis $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}_t)_{t=0}^T, \{\beta_k\}_{k\in\mathbb{N}})$.

Remark 7.3.7. It is important to note the assumptions on Γ ensure that a renormalized martingale solution is consistent with the notion of weak martingale solution given in Definition 7.3.1. Specifically, the assumptions $\sup_{z\in\mathbb{R}}|\Gamma'(z)| < \infty$ and $\Gamma(0) = 0$ given in definition 7.3.6 imply that $\Gamma(z) \le C|z|$. This means that when $f \in L^2(\Omega; L^\infty_t(L^1_{x,v}))$, so is $\Gamma(f)$. Likewise we see that $\Gamma(f_0) \in L^1_{x,v}$ when f_0 is and $\Gamma'(f)g \in L^1(\Omega; L^1_{t,x,v})$ when g is.

Proposition 7.3.8. *Let f be a weak martingale solution to the stochastic kinetic equation driven by g with initial data f_0. Assume that $f \in L^p(\Omega \times [0,T] \times \mathbb{R}^{2d})$ for some $p \in [2,\infty)$. If the noise coefficients satisfy $\sigma \in \ell^2(\mathbb{N}; W^{1,\frac{2p}{p-2}}_{x,v})$ and $\sigma \cdot \nabla_v \sigma \in \ell^1(\mathbb{N}; W^{1,\frac{p}{p-1}}_{x,v})$, then f is also a renormalized weak martingale solution.*

Proof. Let Γ satisfy the assumptions of definition 7.3.6 , then our goal is to establish that $\Gamma(f)$ is a weak martingale solution driven by $\Gamma'(f)g$ starting from $\Gamma(f_0)$. Towards this end, let η be a standard symmetric mollifier with support contained in the unit ball on $\mathbb{R}^d_x \times \mathbb{R}^d_v$ with $\int_{\mathbb{R}^{2d}} \eta(x,v)\mathrm{d}x\mathrm{d}v = 1$. Set $\eta_\epsilon(x,v) = \epsilon^{-2d}\eta(\epsilon^{-1}x, \epsilon^{-1}v)$ and denote by $f_{t,\epsilon} = \eta_\epsilon * f_t = (f_t)_\epsilon$ the mollified process.

Let $\varphi \in C^\infty_c(\mathbb{R}^{2d})$. The main step in this proof will be to establish that for all $t \in [0,T]$, the following identity holds $\mathbf{P}$ almost surely:

$$\begin{aligned}\iint_{\mathbb{R}^{2d}} \Gamma(f_{t,\epsilon})\varphi \mathrm{d}x\mathrm{d}v &= \iint_{\mathbb{R}^{2d}} \Gamma(f_{0,\epsilon})\varphi \mathrm{d}x\mathrm{d}v \\ &\quad + \int_0^t \iint_{\mathbb{R}^{2d}} [\Gamma(f_{s,\epsilon})(v\cdot\nabla_x + \mathcal{L}_\sigma)\varphi + \varphi\Gamma'(f_{s,\epsilon})g_{s,\epsilon}]\mathrm{d}x\mathrm{d}v\mathrm{d}s \\ &\quad + \sum_{k=1}^\infty \int_0^t \iint_{\mathbb{R}^{2d}} \Gamma(f_{s,\epsilon})\sigma_k \cdot \nabla_v\varphi \mathrm{d}x\mathrm{d}v\mathrm{d}\beta_k(s) + R^\varphi_\epsilon(t),\end{aligned} \tag{7.38}$$

for a process $(R^\varphi_\epsilon(t))_{t=0}^T$ such that for each $t \in [0,T]$,

$$R^\varphi_\epsilon(t) \to 0 \quad \text{in probability as} \quad \epsilon \to 0. \tag{7.39}$$

Assuming we can verify (7.38) and (7.39), let us complete the proof. Using standard properties of mollifiers, for almost every $(\omega, t, x, v) \in \Omega \times [0,T] \times \mathbb{R}^{2d}$ one has

$$\Gamma(f_\epsilon) \to \Gamma(f)$$

$$\Gamma(f_{0,\epsilon}) \to \Gamma(f_0),$$

and furthermore, using the boundedness of $\Gamma(z)$ and $\Gamma'(z)$, for each compact set $K \subseteq \mathbb{R}^{2d}$ one has

$$\Gamma(f_\epsilon) \to \Gamma(f) \quad \text{in} \quad L^2(\Omega \times [0,T] \times K),$$

$$\Gamma'(f_\epsilon) g_\epsilon \to \Gamma'(f) g \quad \text{in} \quad L^1(\Omega \times [0,T] \times K).$$

Using the convergence properties above along with the Itô isometry and the convergence of R^φ_ϵ to 0, we may pass the $\epsilon \to 0$ limit in each term of (7.38), where the convergence holds in $L^1(\Omega \times [0,T])$. We conclude that $\Gamma(f)$ solves

$$\begin{aligned}\iint_{\mathbb{R}^{2d}} \Gamma(f_t)\varphi \mathrm{d}x\mathrm{d}v = &\iint_{\mathbb{R}^{2d}} \Gamma(f_0)\varphi \mathrm{d}x\mathrm{d}v \\ &+ \int_0^t \iint_{\mathbb{R}^{2d}} [\Gamma(f_s)(v \cdot \nabla_x + \mathcal{L}_\sigma)\varphi + \varphi \Gamma'(f_s) g_s] \mathrm{d}x\mathrm{d}v\mathrm{d}s \\ &+ \sum_{k=1}^\infty \int_0^t \iint_{\mathbb{R}^{2d}} \Gamma(f_s)\sigma_k \cdot \nabla_v \varphi \mathrm{d}x\mathrm{d}v\mathrm{d}\beta_k(s),\end{aligned}$$

thereby completing the proof.

It now remains to verify identity (7.38) along with the vanishing of the remainder (7.39). We begin by considering the equation (7.30). We fix $z = (x,v) \in \mathbb{R}^{2d}$ and choose $\varphi(w) = \eta_\epsilon(z-w)$. This is equivalent to mollifying both sides of equation, giving

$$\begin{aligned}f_{t,\epsilon}(z) = f_{0,\epsilon}(z) &+ \int_0^t [(-v \cdot \nabla_x f_s)_\epsilon(z) + (\mathcal{L}_\sigma f)_\epsilon(z) + g_{s,\epsilon}(z)] \mathrm{d}s \\ &- \sum_{k \in \mathbb{N}} \int_0^t (\sigma_k \cdot \nabla_v f_s)_\epsilon(z) \mathrm{d}\beta_k(s).\end{aligned}$$

For each $z \in \mathbb{R}^{2d}$, we may renormalize by Γ by applying Itô's formula,

$$\begin{aligned}\Gamma(f_{t,\epsilon}(z)) = \Gamma(f_{0,\epsilon}(z)) + \int_0^t \Gamma'(f_{s,\epsilon}(z))[(-v \cdot \nabla_x f_s)_\epsilon(z) + (\mathcal{L}_\sigma f)_\epsilon(z) + g_{s,\epsilon}(z)]\mathrm{d}s \\ + \frac{1}{2}\sum_{k\in\mathbb{N}} \int_0^t \Gamma''(f_{s,\epsilon}(z))(\sigma_k \cdot \nabla_v f_s)^2_\epsilon(z)\mathrm{d}s \\ - \sum_{k\in\mathbb{N}} \int_0^t \Gamma'(f_{s,\epsilon}(z))(\sigma_k \cdot \nabla_v f_s)_\epsilon(z)\mathrm{d}\beta_k(s).\end{aligned}$$

Naturally we can force the form of (7.38) into view by the use of the commutators,

$$[\eta_\epsilon, v \cdot \nabla_x](f) = (v \cdot \nabla_x f)_\epsilon - v \cdot \nabla_x f_\epsilon$$

$$[\eta_\epsilon, \mathcal{L}_\sigma](f) = (\mathcal{L}_\sigma f)_\epsilon - \mathcal{L}_\sigma f_\epsilon$$

$$[\eta_\epsilon, \sigma_k \cdot \nabla_v](f) = (\sigma_k \cdot \nabla_v f)_\epsilon - \sigma_k \cdot \nabla_v f_\epsilon.$$

Specifically, using the fact that $\mathcal{L}_\sigma \Gamma(f) = \Gamma'(f)\mathcal{L}_\sigma f + \frac{1}{2}(\sigma_k \cdot \nabla_v f)^2 \Gamma''(f)$, we find

$$\begin{aligned}\Gamma(f_{t,\epsilon}) = \Gamma(f_{0,\epsilon}) + \int_0^t [(-v \cdot \nabla + \mathcal{L}_\sigma)\Gamma(f_{s,\epsilon}) + \Gamma'(f_{s,\epsilon})g_{s,\epsilon}]\mathrm{d}s \\ - \sum_{k\in\mathbb{N}} \int_0^t \sigma_k \cdot \nabla_v \Gamma(f_{s,\epsilon})\mathrm{d}\beta_k(s) + R_{t,\epsilon}\end{aligned} \tag{7.40}$$

where $R_{t,\epsilon}$ is a given by

$$\begin{aligned}R_{t,\epsilon} = \int_0^t \Gamma'(f_{s,\epsilon})\big(- [\eta_\epsilon, v \cdot \nabla_x](f_s) + [\eta_\epsilon, \mathcal{L}_\sigma](f_s)\big)\mathrm{d}s \\ + \sum_{k\in\mathbb{N}} \frac{1}{2} \int_0^t \Gamma''(f_{s,\epsilon})[(\sigma_k \cdot \nabla_v f_s)^2_\epsilon - (\sigma_k \cdot \nabla_v f_{s,\epsilon})^2]\mathrm{d}s \\ - \sum_{k\in\mathbb{N}} \int_0^t \Gamma'(f_{s,\epsilon})[\eta_\epsilon, \sigma_k \cdot \nabla_v](f_s)\mathrm{d}\beta_k(s).\end{aligned}$$

Integrating both sides of (7.40) against φ, we obtain (7.38).

It remains to show that for each $t \in [0, T]$,

$$R^\varphi_\epsilon(t) := \iint_{\mathbb{R}^{2d}} \varphi R_{t,\epsilon}\mathrm{d}x\mathrm{d}v \to 0 \quad \text{in probability as} \quad \epsilon \to 0.$$

This will be proved with the aid of standard commutator lemmas taken from [34].

Specifically, we use that $f \in L^p(\Omega \times [0,T] \times \mathbb{R}^{2d}) \cap L^p(\Omega \times [0,T]; L^1_{x,v})$ and for each

$k \in \mathbb{N}$, we have $\sigma_k \in W^{1,\frac{2p}{p-2}}_{x,v}$. It follows that for almost every $(\omega, t) \in \Omega \times [0,T]$ we have

$$[\eta_\epsilon, v \cdot \nabla_x](f_t) \to 0 \quad \text{in} \quad [L^2_{x,v}]_{\text{loc}}, \tag{7.41}$$

$$[\eta_\epsilon, \sigma_k \cdot \nabla_v](f_t) \to 0 \quad \text{in} \quad [L^2_{x,v}]_{\text{loc}}, \tag{7.42}$$

as well as the bound,

$$\|[\eta_\epsilon, \sigma_k \cdot \nabla_v](f_t)\|_{L^2_{x,v}} \leq \|\sigma_k\|_{W^{1,\frac{2p}{p-2}}_{x,v}} \|f_t\|_{L^p_{x,v}}. \tag{7.43}$$

In order to use the commutator results (7.41) and (7.42) to our advantage, we will need to manipulate $R_{t,\epsilon}$. First we write the commutator $[\eta_\epsilon, \mathcal{L}_\sigma](f)$ in terms of $[\eta_\epsilon, \sigma_k \cdot \nabla_v]$ as follows:

$$\begin{aligned}
[\eta_\epsilon, \mathcal{L}_\sigma](f) &= \frac{1}{2} \sum_{k \in \mathbb{N}} \Big((\sigma_k \cdot \nabla_v (\sigma_k \cdot \nabla_v f))_\epsilon - \sigma_k \cdot \nabla_v (\sigma_k \cdot \nabla_v f_\epsilon) \Big) \\
&= \frac{1}{2} \sum_{k \in \mathbb{N}} \Big([\eta_\epsilon, \sigma_k \cdot \nabla_v](\sigma_k \cdot \nabla_v f) + \sigma_k \cdot \nabla_v [\eta_\epsilon, \sigma_k \cdot \nabla_v](f) \Big).
\end{aligned}$$

The second observation is the following equalities

$$\begin{aligned}
&\frac{1}{2}\Gamma''(f_\epsilon)[(\sigma_k \cdot \nabla_v f)^2_\epsilon - (\sigma_k \cdot \nabla_v f_\epsilon)^2] \\
&\qquad = \frac{1}{2}\Gamma''(f_\epsilon)[\eta_\epsilon, \sigma_k \cdot \nabla_v](f)\big((\sigma_k \cdot \nabla_v f)_\epsilon + \sigma_k \cdot \nabla_v f_\epsilon\big) \\
&\qquad = \frac{1}{2}\Gamma''(f_\epsilon)\big([\eta_\epsilon, \sigma_k \cdot \nabla_v](f)\big)^2 + \Gamma''(f_\epsilon)[\eta_\epsilon, \sigma_k \cdot \nabla_v](f)\sigma_k \cdot \nabla_v f_\epsilon \\
&\qquad = \frac{1}{2}\Gamma''(f_\epsilon)\big([\eta_\epsilon, \sigma_k \cdot \nabla_v](f)\big)^2 - \Gamma'(f_\epsilon)\sigma_k \cdot \nabla[\eta_\epsilon, \sigma_k \cdot \nabla_v](f) \\
&\qquad\qquad + \sigma_k \cdot \nabla_v \left(\Gamma'(f_\epsilon)[\eta_\epsilon, \sigma_k \cdot \nabla_v](f)\right).
\end{aligned}$$

Adding the two identities above and introducing the double commutator $\big[[\eta_\epsilon, \sigma_k \cdot \nabla_v], \sigma_k \cdot \nabla_v\big]$ defined by

$$\big[[\eta_\epsilon, \sigma_k \cdot \nabla_v], \sigma_k \cdot \nabla_v\big](f) = [\eta_\epsilon, \sigma_k \cdot \nabla_v](\sigma_k \cdot \nabla_v f) - \sigma_k \cdot \nabla_v [\eta_\epsilon, \sigma_k \cdot \nabla_v](f),$$

we conclude that

$$\Gamma'(f_\epsilon)[\eta_\epsilon, \mathcal{L}_\sigma](f) + \sum_{k\in\mathbb{N}} \frac{1}{2}\Gamma''(f_\epsilon)[(\sigma_k\cdot\nabla_v f)_\epsilon^2 - (\sigma_k\cdot\nabla_v f_\epsilon)^2]$$
$$= \sum_{k\in\mathbb{N}} \Big(\sigma_k\cdot\nabla_v\left(\Gamma'(f_\epsilon)[\eta_\epsilon,\sigma_k\cdot\nabla_v](f)\right) + \frac{1}{2}\Gamma''(f_\epsilon)\big([\eta_\epsilon,\sigma_k\cdot\nabla_v](f)\big)^2$$
$$+ \frac{1}{2}\Gamma'(f_\epsilon)\big[[\eta_\epsilon,\sigma_k\cdot\nabla_v],\sigma_k\cdot\nabla_v\big](f)\Big).$$

The process $R_{t,\epsilon}$ is therefore given by

$$R_{t,\epsilon} = -\int_0^t \Gamma'(f_{s,\epsilon})[\eta_\epsilon, v\cdot\nabla_x](f_s)\mathrm{d}s + \sum_{k\in\mathbb{N}}\int_0^t \sigma_k\cdot\nabla_v\left(\Gamma'(f_\epsilon)[\eta_\epsilon,\sigma_k\cdot\nabla_v](f_s)\right)\mathrm{d}s$$
$$+ \frac{1}{2}\sum_{k\in\mathbb{N}}\int_0^t \Gamma''(f_{s,\epsilon})\big([\eta_\epsilon,\sigma_k\cdot\nabla_v](f_s)\big)^2\mathrm{d}s$$
$$+ \frac{1}{2}\sum_{k\in\mathbb{N}}\int_0^t \Gamma'(f_{s,\epsilon})\big[[\eta_\epsilon,\sigma_k\cdot\nabla_v],\sigma_k\cdot\nabla_v\big](f_s)\mathrm{d}s$$
$$+ \sum_{k\in\mathbb{N}}\int_0^t \Gamma'(f_{s,\epsilon})[\eta_\epsilon,\sigma_k\cdot\nabla_v](f_s)\mathrm{d}\beta_k(s).$$

Integrating $R_{t,\epsilon}$ against φ to obtain $R_\epsilon^\varphi(t)$, it is now possible to use the convergences (7.41), (7.42), the uniform bound (7.43), and our assumptions on the noise coefficients to show that each term in $R_\epsilon^\varphi(t)$ involving the single commutators, $[\eta_\epsilon, v\cdot\nabla_x](f)$ and $[\eta_\epsilon,\sigma_k\cdot\nabla_v](f)$, converges to 0 in probability for each $t\in[0,T]$.

It remains to estimate the double commutator term

$$I_{t,\epsilon} = \frac{1}{2}\sum_{k\in\mathbb{N}}\int_0^t \iint_{\mathbb{R}^{2d}} \varphi\Gamma'(f_{s,\epsilon})\big[[\eta_\epsilon,\sigma_k\cdot\nabla_v],\sigma_k\cdot\nabla_v\big](f_s)\mathrm{d}x\mathrm{d}v\mathrm{d}s.$$

We will prove that for each $t\in[0,T]$, $I_{t,\epsilon}\to 0$ in probability.

In what follows, to simplify notation, we will denote both $z=(x,v)$ and $w=(y,u)$ the phase space (position-velocity) coordinates in $\mathbb{R}^{2d}$ wherever possible, and define the translation operator

$$\delta_w f(z) := f(z+w) - f(z).$$

We will need to evaluate the double-commutator explicitly. This will be done piece by piece. For the first piece, since $\operatorname{div}_v \sigma_k = 0$, integrating by parts gives

$$[\eta_\epsilon\,,\, \sigma_k \cdot \nabla_v]\,(\sigma_k \cdot \nabla_v f_t)(z) = \int_{\mathbb{R}^{2d}} \nabla_v^2 \eta_\epsilon(z-w) : \sigma_k(w) \otimes [\sigma_k(w) - \sigma_k(z)]\, f_t(w)\,\mathrm{d}w$$
$$- \int_{\mathbb{R}^{2d}} \nabla_v \eta_\epsilon(z-w) \cdot (\sigma_k(w) \cdot \nabla_v \sigma_k(w))\, f_t(w)\,\mathrm{d}w,$$

and similarly, for the second piece, we have

$$\sigma_k \cdot \nabla_v\,[\eta_\epsilon\,,\, \sigma_k \cdot \nabla_v]\,(f_t)(z) = \int_{\mathbb{R}^{2d}} \nabla_v^2 \eta_\epsilon(z-w) : [\sigma_k(w) - \sigma_k(z)] \otimes \sigma_k(z)\, f_t(w)\,\mathrm{d}w$$
$$- \int_{\mathbb{R}^{2d}} \nabla_v \eta_\epsilon(z-w) \cdot (\sigma_k(z) \cdot \nabla_v \sigma_k(z))\, f_t(w)\,\mathrm{d}w.$$

Note that the operation $f \to \big[[\eta_\epsilon, \sigma_k \cdot \nabla_v], \sigma_k \cdot \nabla_v\big](f)$ vanishes on constant functions. Hence, in both identities above we may freely replace $f(w)$ by $f(w) - f(z)$. Therefore, using the symmetry of $\nabla_v^2 \eta_\epsilon$, and changing variables $w \to w + z$, we conclude that the double commutator can be written in the following form

$$\big[[\eta_\epsilon, \sigma_k \cdot \nabla_v], \sigma_k \cdot \nabla_v\big](f_t)(z) = \int_{\mathbb{R}^{2d}} \nabla_v^2 \eta_\epsilon(w) : (\delta_w \sigma_k(z) \otimes \delta_w \sigma_k(z))\delta_w f_t(z)\,\mathrm{d}w$$
$$+ \int_{\mathbb{R}^{2d}} \nabla_v \eta_\epsilon(w) \cdot \delta_w(\sigma_k \cdot \nabla_v \sigma_k)(z) \delta_w f_t(z)\,\mathrm{d}w.$$

Next we use the fact that for any $g \in W^{1,r}_{x,v}$, the following inequality holds pointwise in $w \in \mathbb{R}^{2d}$

$$|\delta_w g|_{L^r_{x,v}} \leq |w| |\nabla g|_{L^r_{x,v}}. \tag{7.44}$$

Using Holder's inequality, the estimate (7.44), and the fact that $|\nabla_v^2 \eta_\epsilon(w)|\,|w|^2$ and $|\nabla_v \eta_\epsilon(w)|\,|w|$ are uniformly bounded in L^1_w, we may estimate $I_{t,\epsilon}$ for each $t \in [0,T]$ and $\omega \in \Omega$,

$$|I_{t,\epsilon}| \leq C_\varphi \Big(\|\sigma\|^2_{\ell^2(\mathbb{N}; W^{1,\frac{2p}{p-2}}_{x,v})} + \|\sigma \cdot \nabla \sigma\|_{\ell^1(\mathbb{N}; W^{1,\frac{p}{p-1}}_{x,v})} \Big) \|\Gamma'(f_\epsilon)\|_{L^\infty_{t,x,v}} \sup_{|w|<\epsilon} \|\delta_w f\|_{L^p_{t,x,v}}.$$

Since $f \in L^p([0,T] \times \mathbb{R}^{2d})$ with probability one,

$$\sup_{|w|<\epsilon} \|\delta_w f\|_{L^p_{t,x,v}} \to 0, \quad \mathbf{P} \text{ almost surely.}$$

The proof of the Proposition is now complete since this implies $I_{t,\epsilon} \to 0$ in probability for each $t \in [0,T]$.

□

This section is now completed by checking that renormalized, weak martingale solutions to (7.28) with additional integrability have strongly continuous sample paths. The following lemma will be crucial for ultimately deducing strong continuity properties of the solution to the stochastic Boltzmann equation.

Lemma 7.3.9 (Strong Continuity)**.** *Let f be a renormalized weak martingale solution to the stochastic kinetic equation driven by g with initial data f_0. If f belongs to $L^\infty_t(L^p_{x,v})$ with probability one for some $p \in (1,\infty)$, then $f \in C_t(L^q_{x,v})$ with probability one for any $q \in (1,p)$.*

Proof. We begin by remarking that $f \in C_t([L^p_{x,v}]_w)$ with probability one. Indeed, let $\varphi \in C^\infty_c(\mathbb{R}^{2d})$. It follows directly from inspection of the weak form and elementary properties of stochastic integrals that the process $t \to \langle f_t, \varphi \rangle$ has continuous sample paths. Moreover, since f belongs to $L^\infty_t(L^p_{x,v})$ with probability one, it follows that $f \in C_t([L^p_{x,v}]_w)$ with probability one.

The next step is to upgrade to continuity with values in $L^q_{x,v}$ with the strong topology. Towards this end, let $\psi : \mathbb{R} \to \mathbb{R}$ be defined by $\psi(x) = |x|^q$. We may choose a sequence of smooth, truncations of ψ, denoted $\{\psi_k\}_{k\in\mathbb{N}}$ that satisfy the conditions on the renormalizations in Definition 7.3.6 such that ψ_k converge pointwise in $\mathbb{R}$ to ψ as $k \to \infty$. Moreover, these truncations can be chosen so that when $|x| < k$, $\psi_k(x) = \psi(x)$, and when $|x| > k$, $0 \leq \psi_k(x) \leq \psi(x)$. Applying Proposition 7.3.8, and using the fact that, with probability one, $\psi_k(f)$ is in $L^\infty_t(L^1_{x,v})$ and $\psi_k'(f)g$ is in $L^1_{t,x,v}$, we find that, for all $t \in [0,T]$, we have the **P**- a.s. identity,

$$\|\psi_k(f_t)\|_{L^1_{x,v}} = \|\psi_k(f_0)\|_{L^1_{x,v}} + \int_0^t \iint_{\mathbb{R}^{2d}} \psi_k'(f_s)g_s \mathrm{d}x\mathrm{d}v\mathrm{d}s.$$

In particular, this implies that $t \mapsto \|\psi_k(f_t)\|_{L^1_{x,v}}$ has continuous sample paths with probability one. Since weak martingale solutions are in $C_t([L^1_{x,v}]_w)$ with probability one, then by interpolation, f is in $C_t([L^q_{x,v}]_w)$ with probability one, and therefore for each $t \in [0,T]$, $\|\psi(f_t)\|_{L^1_{x,v}}$ is defined **P**- a.s.

Next, we claim that, **P** almost surely,

$$\|\psi_k(f)\|_{L^1_{x,v}} \to \|\psi(f)\|_{L^1_{x,v}} \quad \text{in} \quad L^\infty([0,T]),$$

whereby we may conclude that $t \to \|f_t\|_{L^q_{x,v}}$ has continuous sample paths with probability one. Indeed, we find

$$\begin{aligned}\sup_{t\in[0,T]} \big|\|\psi(f)\|_{L^1_{x,v}} - \|\psi_k(f)\|_{L^1_{x,v}}\big| &\leq \|\psi(f) - \psi_k(f)\|_{L^\infty_t(L^1_{x,v})} \leq \|\psi(f)\mathbb{1}_{f\geq k}\|_{L^\infty_t(L^1_{x,v})} \\ &\leq \|f\|^q_{L^\infty_t(L^p_{x,v})}\Big(\sup_{t\in[0,T]} |\{|f_t| \geq k\}|\Big)^{1-p/q} \\ &\leq \frac{1}{k^{p-q}}\|f\|^p_{L^\infty_t(L^p_{x,v})} \to 0 \quad \text{as} \quad k \to \infty.\end{aligned}$$

Since $L^q_{x,v}$ is a uniformly convex space for $q > 1$, the fact that f is in $C_t([L^q_{x,v}]_w)$ with probability one, combined with the fact $t \mapsto \|f_t\|_{L^q_{x,v}}$ has **P**-a.s. continuous sample paths implies that $f \in C_t(L^q_{x,v})$ with probability one. □

Stochastic Velocity Averaging

In Section 7.5, we will construct a sequence $\{f_n\}_{n\in\mathbb{N}}$ of approximations to the Boltzmann equation (SB) with stochastic transport. These will satisfy the formal a priori bounds (7.6), uniformly in $n \in \mathbb{N}$ enabling us to extract a weak limit f, which will be a candidate renormalized solution to (SB). However, we need a form of strong compactness to handle the stability of the non-linear collision operator. In this section we investigate some subtle regularizing effects for stochastic kinetic equations, inspired by the classical work of Golse/ Lions/ Perthame/ Sentis [67]. These will be applied in Section 6 to obtain a form of strong compactness of $\{f_n\}_{n\in\mathbb{N}}$. In fact, we allow for a nontrivial probability of oscillations in the velocity variable, so the strong compactness is only in space and time.

It turns out that the criteria for renormalization obtained in Section 3 plays an important role in the proof of our stochastic velocity averaging results. As a consequence, we are only able to establish our compactness criterion for sequences of well-prepared approximations.

Indeed for each $n \in \mathbb{N}$, suppose that f_n is a weak martingale solution to the stochastic kinetic equation driven by g_n and starting from f_n^0, relative to the noise coefficients $\sigma^n = \{\sigma_k^n\}_{k\in\mathbb{N}}$ and the stochastic basis $(\Omega_n, \mathcal{F}_n, (\mathcal{F}_t^n)_{t\in[0,T]}, \{\beta_k^n\}_{k\in\mathbb{N}}, \mathbf{P}_n)$. Then we make the following assumptions on f_n, f_0^n, g_n, and σ_n,

Hypothesis 7.4.1.

1. *Both f_n and g_n belong to $L^{\infty-}(\Omega; L^1_{t,x,v} \cap L^\infty_{t,x,v})$.*

2. f_n^0 *is in* $L^1_{x,v} \cap L^\infty_{x,v}$*, and* $\{f_n^0\}_{n\in\mathbb{N}}$ *is uniformly integrable* $L^1_{x,v}$

3. σ^n *satisfies Hypothesis 7.1.6, and* $\{\sigma^n\}_{n\in\mathbb{N}}$ *satisfies Hypothesis 7.1.2 uniformly.*

Our main stochastic velocity averaging result can now be stated as follows:

Lemma 7.4.2. *Let* $\{f_n\}_{n\in\mathbb{N}}$ *be a sequence of weak martingale solutions to a stochastic kinetic equation satisfying Hypothesis 7.4.1 and suppose that* $\{g_n\}_{n\in\mathbb{N}}$ *is uniformly bounded in* $L^1(\Omega \times [0,T] \times \mathbb{R}^{2d})$ *and induces a tight family of laws on* $[L^1_{t,x,v}]_{\mathrm{w,loc}}$*.*

1. *Then for each* $\varphi \in C_c^\infty(\mathbb{R}^d)$*,* $\{\langle f_n, \varphi\rangle\}_{n\in\mathbb{N}}$ *induces a tight family of laws on* $[L^1_{t,x}]_{\mathrm{loc}}$*.*

2. *If in addition, for each* $\eta > 0$ *the velocity averages* $\{\langle f_n, \varphi\rangle\}_{n\in\mathbb{N}}$ *satisfy*

$$\lim_{R\to\infty} \sup_n \mathbf{P}\left(\|\langle f_n, \varphi\rangle \mathbb{1}_{|x|>R}\|_{L^1_{t,x}} > \eta \right) = 0,$$

then for each $\varphi \in C_c^\infty(\mathbb{R}^d)$*,* $\{\langle f_n, \varphi\rangle\}_{n\in\mathbb{N}}$ *induces a tight family of laws on* $L^1_{t,x}$*.*

L^2 Velocity Averaging

As is typical with velocity averaging lemmas in L^1 (see [67]), we will find it useful first to prove an L^2 result. Roughly speaking, the L^1 case is then reduced to showing that the part of the solution sequence violating the hypotheses of the L^2 lemma has a high probability of being small in L^1.

Lemma 7.4.3. *Let* $\{f_n\}_{n\in\mathbb{N}}$ *be a sequence of martingale solutions to the stochastic kinetic equation satisfying Hypothesis 7.4.1. If* $\{f_n^0\}_{n\in\mathbb{N}}$ *is bounded in* $L^2_{x,v}$ *and* $\{g_n\}_{n\in\mathbb{N}}$*,* $\{f_n\}_{n\in\mathbb{N}}$ *are bounded in* $L^p(\Omega \times [0,T] \times \mathbb{R}^{2d})$ *for each* $p \geq 1$*, then for each* $\varphi \in C_c^\infty(\mathbb{R}^d_v)$*, the velocity averages* $\{\langle f_n, \varphi\rangle\}_{n\in\mathbb{N}}$ *induce tight laws on* $[L^2_{t,x}]_{\mathrm{loc}}$*.*

In the L^2 setting, Fourier methods yielding explicit regularity estimates on the velocity averages can be obtained. More explicitly, given a $\phi \in C_c^\infty(\mathbb{R}_v^d)$, we define the velocity averaged process by

$$\langle f, \phi \rangle(t,x) = \int_{\mathbb{R}^d} f(t,x,v)\phi(v)\,\mathrm{d}v.$$

Using an extension of the method outlined in [18], the following spatial regularity estimate on $\langle f, \phi \rangle$ can be established.

Lemma 7.4.4. *Let f be a weak martingale solution to the stochastic kinetic equation driven by g, with initial data f_0 relative to noise coefficients σ satisfying Hypothesis 7.1.2. If $f, g \in L^2(\Omega \times [0,T] \times \mathbb{R}^{2d})$ and $f_0 \in L^2_{x,v}$, then for any $\phi \in C_c^\infty(\mathbb{R}_v^d)$,*

$$\mathbf{E}\|\langle f, \phi \rangle\|^2_{L_t^2(H_x^{1/6})} \leq C_{\phi,\sigma}\big(\|f_0\|^2_{L^2_{x,v}} + \mathbf{E}\|f\|^2_{L^2_{t,x,v}} + \mathbf{E}\|g\|^2_{L^2_{t,x,v}}\big).$$

The proof is technical and left to Appendix B.2. We are now equipped to prove Lemma 7.4.3:

Proof of Lemma 7.4.3. Let $\phi \in C_c(\mathbb{R}^d)$ be arbitrary. We proceed by explicitly constructing sets $(K_\ell)_{\ell>0}$ which are compact in $[L^2_{t,x}]_{\mathrm{loc}}$ such that

$$\lim_{\ell\to\infty} \sup_n \mathbf{P}\{\langle f_n, \phi \rangle \notin K_\ell\} = 0.$$

Let $\{\varphi_j\}_{j=1}^\infty$ be a dense subset of L^2_x and $\{N_j\}_{j\in\mathbb{N}}$ be a positive, real-valued sequence to be selected later. Define the sets

$$E_\ell = \big\{\rho \in L^2_{t,x} \,:\, \|\rho\|_{L_t^2(H_x^{1/6})} \leq \ell\big\},$$
$$F_\ell = \bigcap_{j=1}^\infty \big\{\rho \in L^2_{t,x} \,:\, \|\langle \rho, \varphi_j \rangle\|_{W_t^{\gamma,p}} \leq (\ell\, N_j)^{\frac{1}{p}}\big\},$$

where $p > 4$ and $\gamma = \frac{1}{4} - \frac{1}{p}$. Let $K_\ell = E_\ell \cap F_\ell$ and observe this is a compact set in $[L^2_{t,x}]_{\text{loc}}$.

Applying the Chebyshev inequality followed by Lemma 7.4.4,

$$\mathbf{P}\big\{\langle f_n, \phi\rangle \notin E_\ell\big\} \leq \frac{1}{\ell}\mathbf{E}\|\langle f_n, \phi\rangle\|_{L^2_t(H^{1/6}_x)} \leq \frac{C_\phi}{\ell},$$

where C_ϕ depends on the uniform bounds for $\{f_n\}_{n=1}^\infty$, $\{g_n\}_{n=1}^\infty$, $\{f_n^0\}_{n\in\mathbb{N}}$, and $\{\sigma^n\}_{n\in\mathbb{N}}$. Similarly, for each $j \in \mathbb{N}$ we may appeal to Lemma 7.3.4 to find a constant C_{φ_j} (depending on the same uniform bounds) such that

$$\mathbf{P}\{\langle f_n, \phi\rangle \notin F_\ell\} \leq \sum_{j=1}^\infty \mathbf{P}\Big\{\big\|\big\langle\langle f_n, \phi\rangle, \varphi_j\big\rangle\big\|_{W_t^{\gamma,p}} < \ell N_j\Big\} \leq \sum_{j=1}^\infty \frac{C_{\varphi_j}}{\ell N_j}.$$

Choosing $N_j = 2^j C_{\varphi_j}$, we conclude that

$$\sup_n \mathbf{P}\{\langle f_n, \phi\rangle \notin K_\ell\} \leq \frac{1}{\ell}\sum_{j=1}^\infty 2^{-j} = \frac{1}{\ell}.$$

Taking $\ell \to \infty$ gives the result. □

Proof of Main lemma

In this section, we give the proof of the main result of the section, Lemma 7.4.2.

Proof of Lemma 7.4.2. Let $\{(\Omega_n, \mathcal{F}_n, (\mathcal{F}_t^n)_{t\in[0,T]}, \{\beta_k^n\}_{k\in\mathbb{N}}, \mathbf{P}_n)\}_{n\in\mathbb{N}}$ be the sequence of stochastic bases corresponding to $\{f_n\}_{n\in\mathbb{N}}$. Fix $\epsilon > 0$ and for each $n \in \mathbb{N}$, we begin by decomposing f_n as

$$f_n = f_n^{\leq L} + f_n^{>L},$$

such that $f_n^{\leq L}$ solves

$$\partial_t f_n^{\leq L} + v\cdot\nabla_x f_n^{\leq} + \sigma_k^n \cdot \nabla_v f_n^{\leq L} \circ \dot{\beta}_k^n = g_n \mathbb{1}_{|g_n|\leq L}, \quad f_n^{\leq L}|_{t=0} = f_n^0 \mathbb{1}_{|f_n^0|\leq L}.$$

and $f_n^{>L}$ solves

$$\partial_t f_n^{>L} + v \cdot \nabla_x f_n^{>L} + \sigma_k^n \cdot \nabla_v f_n^{>L} \circ \dot{\beta}_k^n = g_n \mathbb{1}_{|g_n|>L}, \quad f_n^{>L}|_{t=0} = f_n^0 \mathbb{1}_{|f_n^0|>L}$$

on the filtered probability space $(\Omega_n, \mathcal{F}_n, (\mathcal{F}_t^n)_{t\in[0,T]}, \mathbf{P}_n)$. Since $g_n \mathbb{1}_{|g_n|\leq L}$ belongs to the space $L^{\infty-}(\Omega; L^1_{t,x,v} \cap L^\infty_{t,x,v})$ by Hypothesis 7.4.1, we can build the above decomposition in the following way. First apply the existence result, Theorem 7.3.3 to obtain $f_n^{\leq L}$ as a solution to the equation above. Then, by linearity, the process $f_n^{>L} := f_n - f_n^{\leq L}$ must solve it's corresponding equation above. Moreover, since f_n and $f_n^{\leq L}$ are both in $L^{\infty-}(\Omega \times [0,T] \times \mathbb{R}^{2d})$, so is $f_n^{>L}$. In view of our assumptions on the noise coefficients made in Hypothesis 7.4.1 we may apply Proposition 7.3.8 to deduce that $f_n^{>L}$ is in fact a renormalized solution.

The strategy of the proof will be to show that the process $\langle f_n^{\leq L}, \varphi\rangle$ is tight in n using the L^2 velocity averaging Lemma 7.4.3 and that the remaining processes, $f_n^{\geq L}$, can be made uniformly small in n by taking L sufficiently large and therefore appealing to Lemma B.1.4.

First we apply our L^2 velocity averaging lemma to $\{f_n^{\leq L}\}_{n\in\mathbb{N}}$. Note that $\{f_n^0 \mathbb{1}_{|f_n^0|\leq L}\}_{n\in\mathbb{N}}$ is bounded in $L^2_{x,v}$ (by interpolation) and $\{g_n \mathbb{1}_{|g_n|\leq L}\}_{n\in\mathbb{N}}$ is bounded in $L^{\infty-}(\Omega \times [0,T] \times \mathbb{R}^{2d})$. Therefore, by the estimate given in Theorem 7.3.3, $\{f_n^{\leq L}\}_{n\in\mathbb{N}}$ is also bounded in $L^{\infty-}(\Omega \times [0,T] \times \mathbb{R}^{2d})$. Hence we have enough to apply Lemma 7.4.3 and conclude that $\langle f_n^{\leq L}, \varphi\rangle$ induced tight laws on $[L^2_{t,x}]_{\text{loc}}$.

Our next step is prove tightness of $\{\langle f_n, \varphi\rangle\}_{n\in\mathbb{N}}$ on $[L^1_{t,x}]_{\text{loc}}$ by estimating the sequence $\{\langle f_n^{>L}, \varphi\rangle\}_{n\in\mathbb{N}}$. Indeed, since

$$\|\langle f_n^{>L}, \varphi\rangle\|_{L^1_{t,x}} \leq \|f_n^{>L}\|_{L^1_{t,x,v}} \|\varphi\|_{L^\infty_v},$$

we only need to estimate $\{f_n^{>L}\}$ in $L^1_{t,x,v}$. Therefore, by Lemma B.1.4, it suffices to show that for any $\eta > 0$,

$$\lim_{L\to\infty} \sup_n \mathbf{P}\left(\|f_n^{>L}\|_{L^1_{t,x}} > \eta\right) = 0.$$

Since $f_n^{>L}$ is renormalized, the following inequality holds $\mathbf{P}$ almost surely:

$$\|f_n^{>L}\|_{L^1_{t,x,v}} \leq \|f_n^0 \mathbb{1}_{|f_0^n|>L}\|_{L^1_{x,v}} + \|g_n \mathbb{1}_{|g_n|>L}\|_{L^1_{t,x,v}} \tag{7.45}$$

Since Hypothesis 7.4.1 gives uniform integrability of $\{f_n^0\}_{n\in\mathbb{N}}$, we may choose an $L_0 > 0$ such that for $L > L_0$,

$$\sup_{n\in\mathbb{N}} \|f_n^0 \mathbb{1}_{|f_0^n|>L}\|_{L^1_{x,v}} \leq \eta/2.$$

Therefore by the inequality (7.45),

$$\mathbf{P}\left(\|f_n^{>L}\|_{L^1_{t,x,v}} > \eta\right) \leq \mathbf{P}\left(\|g_n \mathbb{1}_{|g_n|>L}\|_{L^1_{t,x,v}} > \eta/2\right). \tag{7.46}$$

Since $\{g_n\}_{n\in\mathbb{N}}$ induces a tight family of laws on $[L^1_{t,x,v}]_{w,\text{loc}}$, it follows from the tightness criterion on $[L^1_{t,x,v}]_{w,\text{loc}}$ given in Lemma B.1.6 the right-hand side of inequality (7.46) vanishes as $L \to \infty$, thereby proving tightness of the laws of $\{\langle f_n, \varphi\rangle\}_{n\in\mathbb{N}}$ on $[L^1_{t,x}]_{\text{loc}}$.

Next we show that if in addition, for every $\eta > 0$ and $\varphi \in C_c^\infty(\mathbb{R}^d_v)$ we have

$$\lim_{R\to\infty} \sup_n \mathbf{P}\left(\|\langle f_n, \varphi\rangle \mathbb{1}_{|x|>R}\|_{L^1_{t,x}} > \eta\right) = 0,$$

then $\{\langle f_n, \varphi\rangle\}_{n\in\mathbb{N}}$ has tight laws on $L^1_{t,x}$. To this end fix $\epsilon > 0$ and $\varphi \in C_c^\infty(\mathbb{R}^d_v)$ and use what we have just proved to produce a compact set $K \subseteq [L^1_{t,x}]_{\text{loc}}$ such that

$$\mathbf{P}(\langle f_n, \varphi\rangle \notin K) < \epsilon.$$

Next for each $k \in \mathbb{N}$, $k \geq 1$ choose R_k such that

$$\sup_n \mathbf{P}\left(\|\langle f_n, \varphi \rangle \mathbb{1}_{|x|>R_k}\|_{L^1_{t,x}} > 1/k \right) < \epsilon 2^{-k},$$

and define the closed set A_k

$$A_k = \left\{ f \in L^1_{t,x} : \|\langle f, \varphi \rangle \mathbb{1}_{|x|>R_k}\|_{L^1_{t,x}} \leq 1/k \right\}.$$

It it straight forward to conclude that

$$\hat{K} = \bigcap_{k=1}^{\infty} K \cap A_k$$

is tight (in the sense of functions in $[L^1_{t,x}]_{\text{loc}}$) and therefore $\hat{K}$ is compact in $L^1_{t,x}$. It follows that

$$\mathbf{P}\left(\langle f_n, \varphi \rangle \notin \hat{K} \right) \leq \mathbf{P}\left(\langle f_n, \varphi \rangle \notin K \right) + \sum_{k=1}^{\infty} \mathbf{P}\left(\langle f_n, \varphi \rangle \notin A_k \right) < 2\epsilon$$

□

Approximating Scheme

There are two main goals in this section. First, for each $n \in \mathbb{N}$ fixed we will construct a renormalized weak martingale solution to the SPDE

$$\begin{cases} \partial_t f_n + v \cdot \nabla_x f_n + \operatorname{div}_v (f_n \sigma_k^n \circ \dot{\beta}_k) = \mathcal{B}_n(f_n, f_n) \\ f_n \mid_{t=0} = f_n^0, \end{cases} \tag{7.47}$$

where the initial datum f_n^0 and the noise coefficients σ^n are sufficiently regular, and $\mathcal{B}_n$ is an approximation to $\mathcal{B}$ involving a truncation and a regularized collision kernel b_n. The second goal is to rigorously establish the uniform bounds on $\{f_n\}_{n \in \mathbb{N}}$

obtained formally in Section 2. Towards this end, our regularizations are chosen to satisfy the following hypotheses.

Hypothesis 7.5.1 (Initial Data)**.**

1. *For each $n \in \mathbb{N}$, f_n^0 is smooth, non-negative and bounded from above.*

2. *There exists a constant C_n such that for all $(x,v) \in \mathbb{R}^{2d}$, f_n^0 has the lower bound*

$$f_n^0(x,v) \geq C_n e^{-|x|^2-|v|^2}.$$

3. *For all $j \in \mathbb{N}$, $(1+|x|^2+|v|^2)^j f_n^0 \in L^1_{x,v}$,*

4. *The sequence $\{(1+|x|^2+|v|^2+|\log f_n^0|)f_n^0)\}_{n\in\mathbb{N}}$ is uniformly bounded in $L^1_{x,v}$ and $\{f_n^0\}_{n\in\mathbb{N}}$ converges to f_0 strongly in $L^1_{x,v}$.*

Hypothesis 7.5.2 (Noise Coefficients)**.**

1. *For each $k,n \in \mathbb{N}$, the noise coefficient $\sigma_k^n \in C^\infty(\mathbb{R}^{2d};\mathbb{R}^d)$ and $\operatorname{div}_v \sigma_k^n = 0$.*

2. *For $k > n$, the noise coefficient σ_k^n vanishes identically.*

3. *The sequences $\{\sigma^n\}_{n\in\mathbb{N}}$ and $\{\sigma^n\cdot\nabla_v\sigma^n\}_{n\in\mathbb{N}}$ converge pointwise to σ and $\sigma\cdot\nabla_v\sigma$, are uniformly bounded in the spaces $\ell^2(\mathbb{N};L^\infty_{x,v})$ and $\ell^1(\mathbb{N};L^\infty_{x,v})$. Furthermore we have*

$$\lim_{M\to\infty}\sum_{k=M}^\infty \|\sigma_k^n\|_{L^\infty_{x,v}} = 0, \qquad \lim_{M\to\infty}\sum_{k=M}^\infty \|\sigma_k^n\cdot\nabla_v\sigma_k^n\|_{L^\infty_{x,v}} = 0.$$

Hypothesis 7.5.3 (Collision Kernel)**.**

1. *For each $n \in \mathbb{N}$, b_n is smooth and compactly supported in $\mathbb{R}^d \times \mathbb{S}^{d-1}$.*

2. *The sequence* $\{b_n\}_{n\in\mathbb{N}}$ *is bounded in* $L^\infty(\mathbb{R}^d \times \mathbb{S}^{d-1})$ *and converges strongly to* b *in* $L^1(\mathbb{R}^d \times \mathbb{S}^{d-1})$.

Following DiPerna/Lions [36], the truncated collision operator $\mathcal{B}_n$ is defined for $f \in L^1(\mathbb{R}^d_v)$ by

$$\mathcal{B}_n(f,f) = \frac{1}{1 + n^{-1}\int_{\mathbb{R}^d} f \mathrm{d}v} \iint_{\mathbb{R}^d \times \mathbb{S}^{d-1}} (f' f'_* - f f_*) b_n(v - v_*, \theta) \mathrm{d}v_* \mathrm{d}\theta.$$

The following lemma provides the necessary boundedness and continuity properties of the operator $\mathcal{B}_n$. The method of proof is classical, see [36] or [24] for most of the ideas.

Lemma 7.5.4. *For each* $n \in \mathbb{N}$, *there exists a constant* C_n *such that*

1. *For all* $f, g \in L^1_{x,v}$ *it holds:*

$$\|\mathcal{B}_n(f,f) - \mathcal{B}_n(g,g)\|_{L^1_{x,v}} \leq C_n \|f - g\|_{L^1_{x,v}}.$$

2. *For all* f *such that* $(1 + |x|^2 + |v|^2)^k f \in L^1_{x,v}$ *and* $k \in \mathbb{N}$, *it holds*

$$\|(1 + |x|^2 + |v|^2)^k \mathcal{B}_n(f,f)\|_{L^1_{x,v}} \leq C_n \|(1 + |x|^2 + |v|^2)^k f\|_{L^1_{x,v}}.$$

3. *For all* $f \in L^\infty_{x,v}$ *it holds:*

$$\|\mathcal{B}_n(f,f)\|_{L^\infty_{x,v}} \leq C_n \|f\|_{L^\infty_{x,v}}.$$

The strategy for solving the SPDE (7.47) involves a sequence of successive approximations based on mild formulation of (7.47) in terms of stochastic flows. Namely, we fix a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and a collection of independent, one dimensional Brownian motions $\{\beta_k\}_{k\in\mathbb{N}}$. The filtration generated by the Brownian motions is denoted $(\mathcal{F}_t)_{t=0}^T$. For each $n \in \mathbb{N}$, the smoothing regularizations present in Hypothesis 7.5.2, in particular the L^∞ bounds on σ^n and $\sigma^n \cdot \nabla_v \sigma^n$ allow us to apply the results of Kunita [80] to obtain a collection of stochastic flows of volume preserving homeomorphisms $\{\Phi^n_{s,t}\}_{n\in\mathbb{N}}$, $0 \le s \le t \le T$, $\Phi^n_{s,s}(x,v) = (x,v)$, associated to the Stratonovich SDE

$$\mathrm{d}X^n_t = V^n_t \mathrm{d}t, \quad \mathrm{d}V^n_t = \sum_{j=1}^n \sigma^n_j(X^n_t, V^n_t) \circ \mathrm{d}\beta_j.$$

The corresponding inverse (in (x,v)) stochastic flows will be denoted $\{\Psi^n_{s,t}\}_{n\in\mathbb{N}}$. These objects have been studied at length by Kunita [80], so we will mostly defer to this reference for proofs of their properties. The main fact needed for our purposes concerns the following $\mathbf{P}$ almost sure growth estimates for the flow, which can be found as exercises (Exercises 4.5.9 and 4.5.10) in Kunita [80], Chapter 4, Section 5.

Lemma 7.5.5. *Let $\epsilon \in (0,1)$. For each $n \in \mathbb{N}$, the following limits holds $\mathbf{P}$ almost surely:*

$$\lim_{(x,v)\to\infty} \sup_{\{s,t\in[0,T], s\le t\}} \frac{|\Phi^n_{s,t}(x,v)|}{(1+|x|+|v|)^{1+\epsilon}} = 0,$$
$$\lim_{(x,v)\to\infty} \sup_{\{s,t\in[0,T], s\le t\}} \frac{(1+|x|+|v|)^{\epsilon}}{(1+|\Phi^n_{s,t}(x,v)|)} = 0.$$

Our next step is to apply Lemmas 7.5.4 and 7.5.5 to establish the following existence result.

Proposition 7.5.6. *Fix a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \{\beta_k\}_{k\in\mathbb{N}}, \mathbf{P})$. For each $n \in \mathbb{N}$ there exists an analytically weak, stochastically strong solution to the truncated*

Boltzmann equation

$$\partial_t f_n + v \cdot \nabla_x f_n + \sigma_k^n \cdot \nabla_v f_n \circ \dot{\beta}_k = \mathcal{B}_n(f_n, f_n)$$
$$f_n|_{t=0} = f_n^0.$$

such that f_n has the following properties:

1. *$f_n : \Omega \times [0,T] \to L^1_{x,v}$ is a $\mathcal{F}_t$ progressively measurable process.*

2. *f_n belongs to $L^2(\Omega; C_t(L^1_{x,v})) \cap L^\infty(\Omega \times [0,T] \times \mathbb{R}^{2d})$.*

3. *There exists a constant $\overline{C}_n$ such that for each $t \in [0,T]$, **P** almost surely*

$$f_n(t) \geq e^{-\overline{C}_n t} f_n^0 \circ \Psi^n_{0,t}. \tag{7.48}$$

4. *For all $j \in \mathbb{N}$, $(1 + |x|^2 + |v|^2)^j f_n$ is in $L^{\infty-}(\Omega; L^\infty_t(L^1_{x,v}))$.*

5. *The sequence $\{(1+|x|^2+|v|^2) f_n\}_{n\in\mathbb{N}}$ is uniformly bounded in $L^p\big(\Omega; L^\infty_t(L^1_{x,v})\big)$ for each $p \in [1,\infty)$.*

Proof. Begin by constructing a sequence of successive approximations $\{f_{n,k}\}_{k\in\mathbb{N}}$. For each $k \in \mathbb{N}$, define $\{f_{n,k}\}_{k\in\mathbb{N}}$ over $[0,T]$ by the relation

$$f_{n,k}(t) = f_n^0 \circ \Psi^n_{0,t} + \int_0^t \mathcal{B}_n\big(f_{n,k-1}(s), f_{n,k-1}(s)\big) \circ \Psi^n_{s,t}\, \mathrm{d}s, \quad f_{n,0} = 0. \tag{7.49}$$

Applying classical results of Kunita [80], it follows that $f_{n,k}$ is a stochastically strong, classical solution to

$$\begin{aligned} &\partial_t f_{n,k} + v \cdot \nabla_x f_{n,k} + \sigma_j^n \cdot \nabla_v f_{n,k} \circ \dot{\beta}_j = \mathcal{B}_n(f_{n,k}, f_{n,k}), \\ &f_{n,k}|_{t=0} = f_n^0. \end{aligned} \tag{7.50}$$

Let X_T be the Banach space of $(\mathcal{F}_t)_{t=0}^T$ progressively measurable processes $f : [0,T] \times \Omega \to L^1_{x,v}$ endowed with the $L^2(\Omega; C_t(L^1_{x,v}))$ norm. Let C_n be the constant corresponding to the continuity estimates for $\mathcal{B}_n$ from Lemma 7.5.4. In addition, observe that the Hypothesis $\operatorname{div}_v \sigma_k^n = 0$ implies that for every $s, t \in [0,T]$, $s < t$, the flow map $\Phi_{s,t}$ is almost surely volume preserving (see Kunita [80] Theorem 4.3.2 for more details). Taking $L^1_{x,v}$ norms on both sides of (7.49), maximizing over $[0,T]$, and using the Lipschitz continuity of $\mathcal{B}_n$ in $L^1_{x,v}$ obtained in Lemma 7.5.4, we find

$$\|f_{n,k+1} - f_{n,k}\|_{X_T} \leq (C_n T)^k \|f_n^0 \circ \Psi_{0,t}^n\|_{X_T} = (C_n T)^k \|f_n^0\|_{L^1_{x,v}},$$

for each $k \in \mathbb{N}$. Choosing T small enough, the sequence $\{f_{n,k}\}_{k\in\mathbb{N}}$ is Cauchy in X_T. Applying this argument a finite number of times, we may remove the constraint on T. Therefore, for each $n \in \mathbb{N}$, there exists an $f_n \in X_T$ such that $\{f_{n,k}\}_{k\in\mathbb{N}}$ converges to f_n in $L^2(\Omega; C_t(L^1_{x,v}))$. In view of Lemma 7.5.4, $\mathcal{B}_n$ is continuous on $L^2(\Omega; C_t(L^1_{x,v}))$. Therefore we have more then enough to pass the limit weakly in each term of equation (7.50)

Our next step is to verify the lower bound (7.48). Let $\overline{C}_n$ be a deterministic constant to be selected. In view of (7.49) and the fact that $\Psi_{s,t}^n \circ \Phi_{0,t}^n = \Phi_{0,s}^n$ for $s < t$, the following inequalities hold $\mathbf{P}$ almost surely:

$$\begin{aligned}
e^{\overline{C}_n t} f_{n,k+1}(t) \circ \Phi_{0,t}^n &= f_0^n + \int_0^t e^{\overline{C}_n s} \mathcal{B}_n\big(f_{n,k}(s), f_{n,k}(s)\big) \circ \Phi_{0,s}^n \mathrm{d}s + \overline{C}_n \int_0^t e^{\overline{C}_n s} f_{n,k}(s) \circ \Phi_{0,s}^n \mathrm{d}s \\
&\geq f_0^n - \int_0^t e^{\overline{C}_n s} \mathcal{B}_n^-\big(f_{n,k}(s), f_{n,k}(s)\big) \circ \Phi_{0,s}^n \mathrm{d}s + \overline{C}_n \int_0^t e^{\overline{C}_n s} f_{n,k}(s) \circ \Phi_{0,s}^n \mathrm{d}s \\
&\geq f_0^n + [\overline{C}_n - n|\bar{b}_n|_{L^\infty_v}] \int_0^t e^{\overline{C}_n s} f_{n,k}(s) \circ \Phi_{0,s}^n \mathrm{d}s.
\end{aligned}$$

In the last line, we used the explicit definition of the operator $\mathcal{B}_n^-$ together with

Young's inequality and the fact that the flow map is volume preserving. Choose $\overline{C}_n > n|\bar{b}_n|_{L^\infty_v}$ and apply the inequality above inductively to obtain the non-negativity of $f_{n,k}(t) \circ \Phi^n_{0,t}$, which consequently yields the more precise bound $e^{\overline{C}_n t} f_{n,k+1}(t) \circ \Phi^n_{0,t} \geq f^n_0$. Passing $k \to \infty$ and using the $L^2(\Omega; C_t(L^1_{x,v}))$ convergence of $\{f_{n,k}\}_{k\in\mathbb{N}}$ towards f_n, we find that $e^{\overline{C}_n t} f_n(t) \circ \Phi^n_{0,t} \geq f^n_0$ for all $t \in [0,T]$ with probability one. Composing with $\Psi^n_{0,t}$ on both sides gives the desired lower bound (7.48).

Our next step is prove that f_n is in $L^\infty(\Omega \times [0,T] \times \mathbb{R}^{2d})$. We will do this be first checking that the sequence $\{f_{n,k}\}_{k\in\mathbb{N}}$ is uniformly (in k only) bounded in $L^\infty(\Omega \times [0,T] \times \mathbb{R}^{2d})$. By Hypothesis 7.5.1, f^0_n is bounded. Taking $L^\infty_{x,v}$ norms on both sides of (7.49), then maximizing over $t \in [0,T]$ yields $\mathbf{P}$ almost surely

$$\|f_{n,k+1}\|_{L^\infty_{t,x,v}} \leq \|f^0_n\|_{L^\infty_{x,v}} + C_n T \|f_{n,k}\|_{L^\infty_{t,x,v}},$$

where C_n is the constant from Lemma 7.5.4. Iterating, and summing the geometric series, we find that if $T < C_n^{-1}$,

$$\|f_{n,k}\|_{L^\infty_{t,x,v}} \leq (1 - C_n T)^{-1} \|f^0_n\|_{L^\infty_{x,v}}.$$

Of course we may repeat this argument a finite number of times to remove the restriction on T. Taking $L^\infty(\Omega)$ norms on both sides of the above inequality yields the uniform bound. By weak-* L^∞ sequential compactness of $L^\infty(\Omega \times [0,T] \times \mathbb{R}^{2d})$, f_n belongs to $L^\infty(\Omega \times [0,T] \times \mathbb{R}^{2d})$.

Our next goal is to establish the following uniform estimate: for all $p \in (1,\infty)$

$$\sup_{k,n\in\mathbb{N}} \mathbf{E}\|(1 + |x|^2 + |v|^2) f_{n,k}\|^p_{L^\infty_t(L^1_{x,v})} \leq C_p,$$

where C_p depends only on f_0 and σ. If the process $(1 + |x|^2 + |v|^2) f_{n,k}$ was known a priori to belong to $L^{\infty-}(\Omega; C_t(L^1_{x,v}))$, we could argue exactly as in the formal estimates Section 7.2.3.1. Since this is a priori unknown, we proceed by a stopping time argument based on the characteristics. Define for each $R \geq 0$, the stopping time

$$\tau^n_R = \inf\left\{ t \in [0,T] \mid \sup_{s\in[0,t],(x,v)\in\mathbb{R}^{2n}} \frac{|\Phi^n_{s,t}(x,v)|}{(1+|x|+|v|)^2} \geq R \right\}.$$

To see that this stopping time is well defined it suffices to show that the process

$$t \mapsto \sup_{s\in[0,t],(x,v)\in\mathbb{R}^{2n}} \frac{|\Phi^n_{s,t}(x,v)|}{(1+|x|+|v|)^2} \tag{7.51}$$

is adapted to $(\mathcal{F}_t)_{t\geq 0}$ and has continuous sample paths. Indeed, Lemma 4.5.6 of [80] implies that $\Phi^n_{s,t}(x,v)$ is jointly continuous in (s,t,x,v) and therefore the suprema in (7.51) can be taken over a countable dense subset of $[0,t]\times\mathbb{R}^{2d}$, implying adaptedness. Furthermore, the decay estimate presented in Lemma 7.5.5 allows the supremum in (x,v) to be taken over a compact set in $\mathbb{R}^{2d}$. Continuity of the process in (7.51) follows from the fact that for any jointly continuous function $f(x,y)$, $f: X\times Y \to \mathbb{R}$, where X and Y are two compact metric spaces, the function $g(x) = \sup_{y\in Y} g(x,y)$ is continuous.

For each $t\in[0,T]$ we now define the stopped process $f^R_{n,k}(t) = f_{n,k}(t\wedge\tau^n_R)$. We will verify that for each $k,n\in\mathbb{N}$ and $R>0$, the process $(1+|x|^2+|v|^2)^j f^R_{n,k}$ belongs to the space $L^{\infty-}(\Omega; L^\infty_t(L^1_{x,v}))$ for all $j\geq 1$. The claim will be established by induction on $k\in\mathbb{N}$. Suppose the claim is true for step $k-1$. To check k, note that

$$\begin{aligned}\|(1+|x|^2+|v|^2)^j f^R_{n,k}(t)\|_{L^1_{x,v}} &\leq \|(1+|x|^2+|v|^2)^j f^0_n \circ \Psi^n_{0,t\wedge\tau^n_R}\|_{L^1_{x,v}} \\ &+ \int_0^T \left\| \mathbb{1}_{s\in[0,t\wedge\tau^n_R]}(1+|x|^2+|v|^2)^j \mathcal{B}_n\big(f_{n,k-1}(s), f_{n,k-1}(s)\big)\circ\Psi^n_{s,t\wedge\tau^n_R}\right\|_{L^1_{x,v}} \mathrm{d}s\end{aligned}$$

Using the volume preserving property of the stochastic flow, the right-hand side above is equal to

$$\|(1+|\Phi^n_{0,t\wedge\tau^n_R}(x,v)|^2)^j f^0_n\|_{L^1_{x,v}} \\ + \int_0^T \left\|\mathbb{1}_{s\in[0,t\wedge\tau^n_R]}(1+|\Phi^n_{s,t\wedge\tau^n_R}(x,v)|^2)^j \mathcal{B}_n\big(f_{n,k-1}(s),f_{n,k-1}(s)\big)\right\|_{L^1_{x,v}} \mathrm{d}s$$

Using the definition of the stopping time to bound the flow and the L^1 bound on $\mathcal{B}_n$ in Lemma 7.5.4, we obtain

$$\begin{aligned}\|(1+|x|^2+|v|^2)^j f^R_{n,k}(t)\|_{L^1_{x,v}} &\lesssim R^{2j}\|(1+|x|^2+|v|^2)^{2j} f^0_n\|_{L^1_{x,v}} \\ &\quad + R^{2j}\int_0^T \left\|\mathbb{1}_{s\in[0,t\wedge\tau^n_R]}(1+|x|^2+|v|^2)^{2j}\mathcal{B}_n\big(f_{n,k-1}(s),f_{n,k-1}(s)\big)\right\|_{L^1_{x,v}}\mathrm{d}s \\ &\lesssim (1+T)R^{2j}\|(1+|x|^2+|v|^2)^{2j} f^R_{n,k-1}\|_{L^\infty_t(L^1_{x,v})}.\end{aligned}$$

Taking the supremum in time, and the $L^p(\Omega)$ norm on both sides, we may use the inductive hypothesis to complete the inductive step. The base case is established in the same way. Therefore $(1+|x|^2+|v|^2)^j f^R_{n,k}$ belongs to the space $L^{\infty-}(\Omega; L^\infty_t(L^1_{x,v}))$ for all $j \geq 1$.

Now, if one follows the argument in the a priori moment bounds section 7.2.3.1, specifically multiplying the truncated Boltzmann equation for $f^R_{n,k}$ by $(1+|x|^2+|v|^2)$ and integrating in (x,v) so as to kill the collision operator, one may close the estimates on $(1+|x|^2+|v|^2)f^R_{n,k}$ uniformly in k using the BDG inequality, Grönwall's lemma and the uniform hypothesis 7.5.1 and 7.5.2 on the initial data and noise coefficients to find for all $R>0$

$$\mathbf{E}\|(1+|x|^2+|v|^2)f_{n,k}\mathbb{1}_{t\in[0,T\wedge\tau^n_R]}\|^p_{L^\infty_t(L^1_{x,v})} \leq C_{p,T}$$

It is important to note that the constant C_p above does not depend on R, n or k.

The independence of $C_{p,T}$ from R can be readily seen from the fact that the constant obtained in Section 7.2.3.1 depends only in an increasing way on the final time T.

Now we wish to send $R \to \infty$ on both sides of this inequality. To achieve this, note that Lemma 7.5.5 implies that $\mathbf{P}$ almost surely,

$$\left\| \sup_{s,t\in[0,T],s<t} \frac{|\Phi^n_{s,t}(x,v)|}{(1+|x|+|v|)^2} \right\|_{L^\infty_{x,v}} < \infty.$$

Hence,

$$\lim_{R\to\infty} \mathbf{P}(\tau^n_R \le T) = \lim_{R\to\infty} \mathbf{P}\left(\left\| \sup_{s,t\in[0,T],s<t} \frac{|\Phi^n_{s,t}(x,v)|}{(1+|x|+|v|)^2} \right\|_{L^\infty_{x,v}} \ge R \right) = 0.$$

Therefore, it follows that $\tau^n_R \wedge T$ converges in probability to T, and by the monotone convergence theorem we deduce that for any $p \in [1,\infty)$,

$$\mathbf{E}\|(1+|x|^2+|v|^2)f_{n,k}\|^p_{L^\infty_t(L^1_{x,v})} \le C_p.$$

Next we claim that the sequence $\{(1+|x|^2+|v|^2)^j f_{n,k}\}_{k\in\mathbb{N}}$ is uniformly bounded (in k) in $L^{\infty-}(\Omega; L^\infty_t(L^1_{x,v}))$. We can estimate $(1+|x|^2+|v|^2)^j f^R_{n,k}$ in a similar way to $(1+|x|^2+|v|^2)f^R_{n,k}$, by multiplying the truncated Boltzmann equation for $f^R_{n,k}$ by $(1+|x|^2+|v|^2)^j$ and using estimate 2 in Lemma 7.5.4 to bound the collision operator. Using the BDG inequality and Grönwall inequality one can obtain after some tedious, though straight forward, calculations and using the uniform hypothesis 7.5.2 on the noise coefficients,

$$\begin{aligned} \mathbf{E}\|(1+|x|^2+|v|^2)^j f_{n,k} \mathbb{1}_{t\in[0,T\wedge\tau^n_R]}\|^p_{L^\infty_t(L^1_{x,v})} &\le C_{p,T,j}\|(1+|x|^2+|v|^2)^j f^0_n\|^p_{L^1_{x,v}} \\ &+ TC_{p,T,n,j}\mathbf{E}\|(1+|x|^2+|v|^2)^j f_{n,k-1}\mathbb{1}_{t\in[0,T\wedge\tau^n_R]}\|^p_{L^\infty_t(L^1_{x,v})}, \end{aligned}$$

where the constants $C_{p,T,n,j}$ and $C_{p,T,j}$ are independent of k and R and depend on the final time in an increasing way. Since we have made explicit that there is a multiplicative factor in the second term T above (coming from the time integral of the collision operator), we find that, independently of k and the initial data we may choose T small enough so that $TC_{p,T,n,j} < 1$. This means that we may iterate the bound above and sum the geometric series to conclude that for such T, to conclude

$$\mathbf{E}\|(1+|x|^2+|v|^2)^j f_{n,k}\mathbb{1}_{t\in[0,T\wedge\tau_R^n]}\|^p_{L^\infty_t(L^1_{x,v})} \le C_{p,T,n,j}.$$

Again, sending $R \to \infty$ and using monotone convergence we conclude the uniform in k estimate

$$\mathbf{E}\|(1+|x|^2+|v|^2)^j f_{n,k}\|^p_{L^\infty_t(L^1_{x,v})} \le C_{p,T,n,j}.$$

The restriction on T can be removed in the usual way by repeating the above argument a finite number of times.

What remains is to pass the limit in k on these estimates to obtain the estimates on f_n stated in the Lemma. It suffices to show that for each $j \ge 0$, and $p \in [1,\infty)$,

$$\mathbf{E}\|(1+|x|^2+|v|^2)^j f_n\|^p_{L^\infty_t(L^1_{x,v})} \le \sup_{k\in\mathbb{N}} \mathbf{E}\|(1+|x|^2+|v|^2)^j f_{n,k}\|^p_{L^\infty_t(L^1_{x,v})}. \tag{7.52}$$

We do this by cutting off the moment function. Let B_M be the ball of radius $M > 0$ in $\mathbb{R}^{2d}$. Since $f_{n,k} \to f_n$ in $L^2(\Omega; L^\infty_t(L^1_{x,v})))$, upon choosing a further subsequence if necessary, we have that $\mathbf{P}$ almost surely,

$$\|(1+|x|^2+|v|^2)^j f_{n,k}\mathbb{1}_{B_M}\|^p_{L^\infty_t(L^1_{x,v})} \to \|(1+|x|^2+|v|^2)^j f_n\mathbb{1}_{B_M}\|^p_{L^\infty_t(L^1_{x,v})}.$$

Applying Fatou's Lemma, gives

$$\mathbf{E}\|(1+|x|^2+|v|^2)^j f_n\mathbb{1}_{B_M}\|^p_{L^\infty_t(L^1_{x,v})} \le \sup_{k\in\mathbb{N}} \mathbf{E}\|(1+|x|^2+|v|^2)^j f_{n,k}\|^p_{L^\infty_t(L^1_{x,v})}.$$

The inequality (7.52) is then proved by passing the limit in M on the left-hand side by monotone convergence. □

The final step in this section is to realize the a priori estimates obtained from the formal entropy dissipation inequality (7.7). Towards this end, define the approximate entropy dissipation $f \to \mathcal{D}_n(f)$ by the relation

$$\mathcal{D}_n(f) \equiv \frac{1}{4}\big(1 + n^{-1}\langle f, 1\rangle\big)^{-1} \iiint_{\mathbb{R}^{2d}\times\mathbb{S}^{d-1}} d(f)\, b_n(v - v_*, \theta)\, \mathrm{d}\theta \mathrm{d}v_* \mathrm{d}v,$$

where $d(f)$ is defined by (7.8). Similarly, define $\mathcal{D}_n^0(f)$ by

$$\mathcal{D}_n^0(f) \equiv \frac{1}{4}(1 + n^{-1}\langle f, 1\rangle)^{-1} \iint_{\mathbb{R}^{d}\times\mathbb{S}^{d-1}} d(f) b_n(v - v_*, \theta) \mathrm{d}\theta \mathrm{d}v_*.$$

Lemma 7.5.7. *Let $\{f_n\}_{n\in\mathbb{N}}$ be the sequence constructed in Proposition 7.5.6. For each $p \in (1, \infty)$, there exists a constant C_p depending on σ and f_0 such that*

$$\sup_{n\in\mathbb{N}} \mathbf{E}\|f_n \log f_n\|^p_{L^\infty_t(L^1_{x,v})} \le C_p, \quad \sup_{n\in\mathbb{N}} \mathbf{E}\|\mathcal{D}_n(f_n)\|^p_{L^1_{t,x}} \le C_p.$$

Proof. Begin by fixing $n \in \mathbb{N}$. Note that it suffices to verify identity (7.27) from the formal a priori bounds section. For each $\epsilon > 0$, we define the renormalization $\beta_\epsilon(x) = x\log(x+\epsilon)$. Using Proposition 7.3.8 and the fact that f_n belongs to $L^\infty(\Omega \times [0, T] \times \mathbb{R}^{2d})$ and $L^2(\Omega; C_t(L^1_x))$, it can be checked with a truncation argument that $\beta_\epsilon(f_n)$ is a weak solution to the stochastic kinetic equation driven by $\beta'_\epsilon(f_n)\mathcal{B}_n(f_n, f_n)$, starting from $\beta_\epsilon(f^n_0)$. In particular, using the L^1 bounds on f^n and the fact that $\mathcal{B}(f_n, f_n) \in L^1_{t,x,x}$, we can obtain the $\mathbf{P}$ almost sure identity

$$\iint_{\mathbb{R}^{2d}} \beta_\epsilon(f_n(t))\mathrm{d}x\mathrm{d}v = \iint_{\mathbb{R}^{2d}} \beta_\epsilon(f^0_n)\mathrm{d}x\mathrm{d}v + \int_0^t \iint_{\mathbb{R}^{2d}} \beta'_\epsilon(f_n)\mathcal{B}_n(f_n, f_n)\mathrm{d}x\mathrm{d}v\mathrm{d}s. \tag{7.53}$$

Observe that almost everywhere in $\Omega\times[0,T]\times\mathbb{R}^{2d}$, as $\epsilon\to 0$ we have the convergence $\beta_\epsilon(f_n(t))\to f_n\log f_n(t)$ and $\beta'_\epsilon(f_n)\mathcal{B}_n(f_n,f_n)\to[1+\log f_n]\mathcal{B}_n(f_n,f_n)$. Since f_n is in $L^\infty(\Omega\times[0,T]\times\mathbb{R}^{2d})$ and in $L^2(\Omega;C_t(L^1_{x,v}))$, it follows that $\mathbf{P}$ almost surely, for each $t\in[0,T]$

$$\iint_{\mathbb{R}^{2d}}\beta_\epsilon(f_n(t))\mathrm{d}x\mathrm{d}v\to\iint_{\mathbb{R}^{2d}}f_n(t)\log f_n(t)\mathrm{d}x\mathrm{d}v.$$

The initial data are also handled similarly in view of Hypothesis 7.5.1. To pass the limit in the remaining integral on the RHS of (7.53), note that $|\beta'_\epsilon(x)|\leq(2+|\log(x)|)$ for ϵ small. Hence, by the dominated convergence theorem, it suffices to show that $\log f_n\mathcal{B}_n(f_n,f_n)$ belongs to $L^1_{t,x,v}$ with probability one. By Proposition 7.5.6 combined with Hypothesis 7.5.1 we have

$$C_ne^{-\overline{C}_nt}e^{-|\Psi^n_{0,t}|^2}\leq f_n\leq\|f_n\|_{L^\infty(\Omega\times[0,T]\times\mathbb{R}^{2d})}. \tag{7.54}$$

The second estimate on $\Phi^n_{0,t}$ given in Lemma 7.5.5, implies that $\mathbf{P}$ almost surely we have the bound,

$$\sup_{(t,x,v)\in[0,T]\times\mathbb{R}^{2d}}\frac{|\Psi^n_{0,t}(x,v)|}{(1+|x|+|v|)^2}<\infty$$

Combining this with the bounds in (7.54) it follows that $\mathbf{P}$ almost surely

$$\sup_{(t,x,v)\in[0,T]\times\mathbb{R}^{2d}}\frac{|\log f_n(t,x,v)|}{(1+|x|^2+|v|^2)^2}<\infty.$$

Using this, the $\mathbf{P}$ almost sure $L^1_{t,x,v}$ estimate on $\log f_n\mathcal{B}_n(f_n,f_n)$ now follows from property 3 of Lemma 7.5.4 and the fact that $(1+|x|^2+|v|^2)^2f_n\in L^\infty_t(L^1_{x,v})$ with probability one. □

Compactness and Preliminary Renormalization

Let $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ be the sequence of renormalized weak martingale solutions to (7.47) constructed in Proposition 7.5.6. Denote the supporting stochastic basis by $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}}, (\tilde{\mathcal{F}}_t)_{t=0}^T, \{\tilde{\beta}_k\}_{k\in\mathbb{N}})$. In view of Proposition 7.5.6 and Lemma 7.5.7, we have the uniform bounds

$$\begin{aligned} &\sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\|(1+|x|^2+|v|^2+|\log \tilde{f}_n|)\tilde{f}_n\|^p_{L^\infty_t(L^1_{x,v})} < \infty \\ &\sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\|\mathcal{D}_n(\tilde{f}_n)\|^p_{L^1_{t,x}} < \infty. \end{aligned} \tag{7.55}$$

In this section, we will deduce several key tightness results and apply our main stochastic velocity averaging Lemma 7.4.2. We will study the induced laws of the approximations $\{\tilde{f}_n\}_{n\in\mathbb{N}}$, the renormalized approximations $\{\Gamma(\tilde{f}_n)\}_{n\in\mathbb{N}}$, and renormalized collision operators $\{\Gamma'(f_n)\mathcal{B}_n(\tilde{f}_n,\tilde{f}_n)\}_{n\in\mathbb{N}}$. The precise results are stated in Lemmas 7.6.4-7.6.8. Combining our tightness result with a recent extension of the Skorohod Theorem B.1.2 to non-metric spaces, we will obtain our main compactness result Proposition 7.6.1.

Towards this end, we introduce for each $m \in \mathbb{N}$ a truncation type renormalization Γ_m defined by

$$\Gamma_m(z) = \frac{z}{1+m^{-1}z}. \tag{7.56}$$

The space $L^1_{t,x}(\mathcal{M}^*_v)$

In order the apply the velocity averaging results we will find it convenient to turn the tightness results on velocity averages of f of proved in Section 7.4 into tightness results for f on a particular space $L^p_{t,x}(\mathcal{M}^*_v)$ characterizing 'convergence in the sense of velocity averages'. To be more precise, we introduce a topological vector space $L^p_{t,x}(\mathcal{M}^*_v)$ as follows. Let $\mathcal{M}_v$ denote the space of finite Radon measures on $\mathbb{R}^d_v$, which can be identified with the dual of the continuous functions $C_0(\mathbb{R}^d)$ that vanish at ∞, and let $\mathcal{M}^*_v$ be $\mathcal{M}_v$ equipped with it's weak star topology. Consider the collection of equivalence classes (up to Lebesgue $[0,T] \times \mathbb{R}^d_x$ null sets) of measurable maps $f : [0,T] \times \mathbb{R}^d_x \to \mathcal{M}^*_v$, where the Borel sigma algebra is taken on $\mathcal{M}^*_v$. For each equivalence class f, and $\phi \in C_0(\mathbb{R}^d)$ we let $\langle f, \phi \rangle$ denote the pair between $\mathcal{M}_v$ and $C_0(\mathbb{R}^d)$ and for each $\phi \in C_0(\mathbb{R}^d)$, define a corresponding semi-norm ν_ϕ via

$$\nu_\phi(f) = \|\langle f, \phi \rangle\|_{L^p_{t,x}}.$$

We then say that f is in $L^p_{t,x}(\mathcal{M}^*_v)$ provided that for all $\phi \in C_0(\mathbb{R}^d)$, $\nu_\phi(f) < \infty$. Convergence in the space $L^p_{t,x}(\mathcal{M}^*_v)$ can be thought of as *strong* in the variables (t,x) and *weak* in the velocity variable v. The space $L^p_{t,x}(\mathcal{M}^*_v)$ can be identified with $\mathcal{L}(C_0(\mathbb{R}^d), L^p_{t,x})$ the space of bounded linear operators from $C_0(\mathbb{R}^d)$ to $L^1_{t,x}$ under the topology of pointwise convergence (see Lemma B.1.9).

We will also define the space $[L^p_{t,x}(\mathcal{M}^*_v)]_{\text{loc}}$ of locally integrable functions which is the space of equivalence classes of measurable functions $f : [0,T] \times \mathbb{R}^d \to \mathcal{M}_v$ generated by the semi-norms,

$$\nu_{\phi,K}(f) = \|\langle f, \phi \rangle \mathbb{1}_K\|_{L^p_{t,x}}$$

for each $\phi \in C_0(\mathbb{R}^d)$ and each compact set $K \subseteq \mathbb{R}^d$. Again such a space has an identification with $\mathcal{L}(C_0(\mathbb{R}^d), [L^p_{t,x}]_{\text{loc}})$.

Our main tool for obtaining compactness in the space $L^p_{t,x}(\mathcal{M}^*_v)$ are Lemmas B.1.10 and B.1.11, which give necessary and sufficient conditions for compactness and tightness of measure on $L^p_{t,x}(\mathcal{M}^*_v)$.

Statement of the main proposition

The main result of this section is the following compactness result.

Proposition 7.6.1. *There exists a new probability space* $(\Omega, \mathcal{F}, \mathbf{P})$ *and a sequence of maps* $\{\widetilde{T}_n\}_{n\in\mathbb{N}}$ *from* Ω *to* $\tilde{\Omega}$ *with the following properties:*

1. *For each* $n \in \mathbb{N}$*, the map* $\tilde{T}_n$ *is measurable from* $(\Omega, \mathcal{F})$ *to* $(\tilde{\Omega}, \tilde{\mathcal{F}})$ *and* $(\tilde{T}_n)_{\#}\mathbf{P} = \tilde{\mathbf{P}}$.

2. *The new sequence* $\{f_n\}_{n\in\mathbb{N}}$ *defined by* $f_n = \tilde{f}_n \circ \widetilde{T}_n$ *satisfies the uniform bounds* (7.55) *with* $\mathbf{E}$ *replacing* $\tilde{\mathbf{E}}$*. Moreover, for all* $\omega \in \Omega$*, there exists a constant* $C(\omega)$ *such that*
$$\sup_{n\in\mathbb{N}} \|(1 + |x|^2 + |v|^2 + |\log f_n(\omega)|)f_n(\omega)\|_{L^\infty_t(L^1_{x,v})} \leq C(\omega).$$
$$\sup_{n\in\mathbb{N}} \|\mathcal{D}_n(f_n)(\omega)\|_{L^1_{t,x}} \leq C(\omega).$$

3. *The new sequence* $\{\beta^n_k\}_{k\in\mathbb{N}}$ *defined by* $\beta^n_k = \tilde{\beta}^n_k \circ \widetilde{T}_n$ *consists of one-dimensional Brownian motions on* $(\Omega, \mathcal{F}, \mathbf{P})$.

4. *There exist random variables* f *and* $\{\beta_k\}_{k\in\mathbb{N}}$ *with values in* $C_t([L^1_{x,v}]_w)$ *and* $[C_t]^\infty$ *respectively, such that the following convergences hold pointwise on* Ω*:*
$$f_n \to f \quad \text{in} \quad L^1_{t,x}(\mathcal{M}^*_v) \cap C_t([L^1_{x,v}]_w).$$
$$\{\beta^n_k\}_{k\in\mathbb{N}} \to \{\beta_k\}_{k\in\mathbb{N}} \quad \text{in} \quad [C_t]^\infty.$$

5. *For each $m \in \mathbb{N}$, there exist auxiliary random variables $\overline{\Gamma_m(f)}$ and $\overline{\gamma_m(f)}$ in $C_t([L^1_{x,v}]_w)$ along with $\mathcal{B}^-_m$ and $\mathcal{B}^+_m$ in $L^1_{t,x,v}$ and $\overline{\mathcal{D}^0(f)}$ in $\mathcal{M}_{t,x,v}$ such that the following convergences hold pointwise on Ω:*

$$\Gamma_m(f_n) \to \overline{\Gamma_m(f)} \quad in \quad L^1_{t,x}(\mathcal{M}^*_v) \cap C_t([L^1_{x,v}]_w).$$

$$\Gamma'_m(f_n) f_n \to \overline{\gamma_m(f)} \quad in \quad L^1_{t,x}(\mathcal{M}^*_v) \cap C_t([L^1_{x,v}]_w).$$

$$\Gamma'_m(f_n)\mathcal{B}^+_n(f_n, f_n) \to \mathcal{B}^+_m \quad in \quad [L^1_{t,x,v}]_w.$$

$$\Gamma'_m(f_n)\mathcal{B}^-_n(f_n, f_n) \to \mathcal{B}^-_m \quad in \quad [L^1_{t,x,v}]_w$$

$$\mathcal{D}^0_n(f_n) \to \overline{\mathcal{D}^0(f)} \quad in \quad \mathcal{M}^*_{t,x,v}.$$

Remark 7.6.2. For all $n \in \mathbb{N}$, f_n is a weak martingale solution to the stochastic kinetic equation driven by $\mathcal{B}_n(f_n, f_n)$, starting from f_0, with noise coefficients σ^n. The supporting stochastic basis is given by $(\Omega, \mathcal{F}, \mathbf{P}, (\mathcal{F}^n_t)^T_{t=0}, \{\beta^n_k\}_{k\in\mathbb{N}})$, where the Brownian motions are given by $\beta^n_k = \tilde{\beta}^n_k \circ \widetilde{T}_n$ and $\mathcal{F}^n_t = \tilde{T}^{-1}_n \circ \tilde{\mathcal{F}}_t$.

Tightness of renormalized quantities

In this section, we study the compactness properties of the sequences $\{\Gamma(\tilde{f}_n)\}_{n\in\mathbb{N}}$ and

$\{\Gamma'(\tilde{f}_n)\mathcal{B}^+_n(\tilde{f}_n, \tilde{f}_n)\}_{n\in\mathbb{N}}$, where Γ is a renormalization of a particular type.

Definition 7.6.3. Let $\mathcal{R}'$ denote the class of renormalizations $\Gamma \in C^2(\mathbb{R}_+)$, such that $\Gamma(0) = 0$ and

$$\sup_{x\in\mathbb{R}_+} \left(|\Gamma(x)| + (1+x)|\Gamma'(x)| + |\Gamma''(x)|\right) < \infty.$$

Lemma 7.6.4. *For each $\Gamma \in \mathcal{R}'$, the sequences $\{\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)\}_{n\in\mathbb{N}}$ and $\{\Gamma'(\tilde{f}_n)\mathcal{B}_n^+(\tilde{f}_n, \tilde{f}_n)\}_{n\in\mathbb{N}}$ are uniformly bounded in $L^{\infty-}(\tilde{\Omega}; L^1_{t,x,v})$.*

Proof. Let us begin with an estimate for $\{\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)\}_{n\in\mathbb{N}}$. Since $\Gamma \in \mathcal{R}'$, the mapping

$x \to (1+x)|\Gamma'(x)|$ is bounded on $\mathbb{R}_+$ Therefore, the following inequalities hold on $\tilde{\Omega} \times [0,T] \times \mathbb{R}^{2d}$

$$\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n) \lesssim \frac{\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)}{1+f_n} \lesssim \tilde{f}_n * \bar{b}_n,$$

where the convolution is only in the variable v. Recall, by Hypothesis 7.5.3, the sequence $\{\bar{b}_n\}_{n\in\mathbb{N}}$ is uniformly bounded in $L^1(\mathbb{R}^d_v)$. Integrating over $\tilde{\Omega} \times [0,T] \times \mathbb{R}^{2d}$ and applying Young's inequality for convolutions yields for each $p \in [1,\infty)$

$$\sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\|\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)\|^p_{L^1_{t,x,v}} \lesssim \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\|\tilde{f}_n\|^p_{L^1_{t,x,v}}. \tag{7.57}$$

Now we can estimate $\{\Gamma'(\tilde{f}_n)\mathcal{B}_n^+(\tilde{f}_n, \tilde{f}_n)\}_{n\in\mathbb{N}}$ by applying the bound (7.20) pointwise in $\tilde{\Omega} \times [0,T] \times \mathbb{R}^{2d}$ (to the truncated collision operator $\mathcal{B}_n(\tilde{f}_n, \tilde{f}_n)$ instead of $\mathcal{B}(f,f)$), then integrating in all variables to find

$$\begin{aligned}\sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\|\Gamma'(\tilde{f}_n)\mathcal{B}_n^+(\tilde{f}_n, \tilde{f}_n)\|^p_{L^1_{t,x,v}} &\lesssim \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\|\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)\|^p_{L^1_{t,x,v}} + \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\|\mathcal{D}_n(\tilde{f}_n)\|^p_{L^1_{t,x}} \\ &\lesssim \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\|\tilde{f}_n\|^p_{L^1_{t,x,v}} + \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\|\mathcal{D}_n(\tilde{f}_n)\|^p_{L^1_{t,x}},\end{aligned} \tag{7.58}$$

where we used (7.57) in the last line. In view of inequalities (7.57) and (7.58), the Proposition now follows from the uniform bounds (7.55). □

Lemma 7.6.5. *For each $\Gamma \in \mathcal{R}'$, the sequence $\{\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)\}_{n\in\mathbb{N}}$ induces tight laws on $[L^1_{t,x,v}]_w$.*

Proof. Effectively, we have to show that the renormalized collision sequence is bounded, uniformly integrable, and tight in $L^1_{t,x,v}$, with uniformly high probability. Towards this end, let $\Psi(t) = t|\log t|$. By well-known arguments (see Section 3 in [36]), there exists a constant C depending only on Γ and $\|\bar{b}\|_{L^1_v}$ such that the following two inequalities hold. Regarding uniform integrability,

$$\int_0^T \iint_{\mathbb{R}^{2d}} \Psi\left(\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)\right) \mathrm{d}x\mathrm{d}v\mathrm{d}s \le C\Big[\|\tilde{f}_n\|_{L^1_{t,x,v}} + \int_0^T \iint_{\mathbb{R}^{2d}} \Psi(\tilde{f}_n)\mathrm{d}x\mathrm{d}v\mathrm{d}s\Big]. \tag{7.59}$$

Moreover, regarding tightness (in $L^1_{t,x,v}$), for all $R > 0$

$$\begin{aligned}\int_0^T \iint_{\mathbb{R}^{2d}} & 1_{\{|x|+|v|>R\}}\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)\mathrm{d}x\mathrm{d}v\mathrm{d}s \\ &\le C\Big[\|\tilde{f}_n\|_{L^1_{t,x,v}} \int_{\mathbb{R}^d} 1_{\{|v|>\frac{R}{2}\}}\bar{b}_n(v)\mathrm{d}v + R^{-2}\int_0^T \iint_{\mathbb{R}^{2d}} (|x|^2 + |v|^2)\tilde{f}_n\mathrm{d}x\mathrm{d}v\mathrm{d}s\Big].\end{aligned} \tag{7.60}$$

Define the function $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$\lambda(R) = \max\left\{\sup_n \int_{\mathbb{R}^d} \mathbb{1}_{|v|>\frac{R}{2}}\bar{b}_n(v)\mathrm{d}v\,,\ R^{-2}\right\},$$

and note that, by Hypothesis 7.5.3, $\lambda(R) \to 0$ as $R \to \infty$. Combining (7.59) and (7.60) with the uniform bounds on $\tilde{f}_n$,

$$\sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\Big(\big\|\psi\left(\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)\right)\big\|_{L^1_{t,x,v}}\Big) < \infty. \tag{7.61}$$

$$\sup_{n\in\mathbb{N}} \tilde{\mathbf{E}}\bigg(\sup_{R>0}\big[\lambda(R)^{-1}\big\|1_{\{|x|+|v|>R\}}\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n, \tilde{f}_n)\big\|_{L^1_{t,x,v}}\big]\bigg) < \infty. \tag{7.62}$$

To construct our compact sets, note that for all $M > 0$, the set

$$\Big\{f \in L^1_{t,x,v} \mid \|f\|_{L^1_{t,x,v}} + \|\psi(f)\|_{L^1_{t,x,v}} + \sup_{R>0}\big[\lambda(R)^{-1}\|1_{\{|x|+|v|>R\}}f\|_{L^1_{t,x,v}}\big] \le M\Big\}$$

is weakly compact in $L^1_{t,x,v}$. Indeed, every sequence in this set is bounded, uniformly integrable, and tight in $L^1_{t,x,v}$. By Chebyshev, the uniform bounds (7.61), (7.62) and our previous Lemma 7.6.5, it follows that $\{\Gamma'(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n,\tilde{f}_n)\}_{n\in\mathbb{N}}$ induces tight laws on $[L^1_{t,x,v}]_{\mathrm{w}}$. □

Lemma 7.6.6. *For each Γ, the sequence $\{\Gamma'(\tilde{f}_n)\mathcal{B}_n^+(\tilde{f}_n,\tilde{f}_n)\}_{n\in\mathbb{N}}$ induces tight laws on $[L^1_{t,x,v}]_w$.*

Proof. The main ingredient in the proof is a version of inequality (7.20), which we state again in the precise form required. Specifically, for each $j > 1$ the following inequality holds pointwise a.e in $\Omega \times [0,T] \times \mathbb{R}^{2d}$,

$$\Gamma'(\tilde{f}_n)\mathcal{B}_n^+(\tilde{f}_n,\tilde{f}_n) \leq j\Gamma'(f_n)\mathcal{B}_n^-(\tilde{f}_n,\tilde{f}_n) + \frac{1}{\log j}\mathcal{D}_n^0(\tilde{f}_n), \tag{7.63}$$

where we recall that

$$\mathcal{D}_n^0(\tilde{f}_n) = \frac{1}{1+n^{-1}\int_{\mathbb{R}^d}\tilde{f}_n \mathrm{d}v}\int_{\mathbb{R}^d} d_n(\tilde{f}_n)\mathrm{d}v_*.$$

Let $\epsilon > 0$. By Lemma 7.6.5, we may select a weakly compact set K_ϵ^- in $L^1_{t,x,v}$ such that

$$\sup_{n\in\mathbb{N}} \tilde{\mathbf{P}}\big(\Gamma'(f_n)\mathcal{B}_n^-(\tilde{f}_n,\tilde{f}_n) \notin K_\epsilon^-\big) \leq \frac{\epsilon}{2}.$$

Moreover, in view of the uniform bound on the entropy dissipation (7.55), we can select a closed ball, B_{M_ϵ} of size $M_\epsilon > 0$ in $L^1_{t,x,v}$ such that

$$\sup_{n\in\mathbb{N}} \tilde{\mathbf{P}}\big(\mathcal{D}_n^0(\tilde{f}_n) \notin B_{M_\epsilon}\big) \leq \frac{\epsilon}{2}.$$

For each $j \in \mathbb{N}$, we define a set

$$K_{j,\epsilon} = \big\{f \in L^1_{t,x,v} \mid \text{There exists } g \in K_\epsilon^- \text{ and } h \in B_{M_\epsilon} \text{ such that } f \leq jg + (\log j)^{-1}h\big\}.$$

The inequality in the definition of $K_{j,\epsilon}$ is understood to hold a.e. on $[0,T]\times\mathbb{R}^{2d}$. Next define the set K_ϵ via

$$K_\epsilon = \bigcap_{j\in\mathbb{N}} K_{j,\epsilon}.$$

Note that if $\Gamma'(\tilde f_n)\mathcal{B}_n^-(\tilde f_n,\tilde f_n)\in K_\epsilon^-$ and $\mathcal{D}_n^0(\tilde f_n)\in B_{M_\epsilon}$, then inequality (7.63) implies that $\Gamma'(\tilde f_n)\mathcal{B}_n^+(\tilde f_n,\tilde f_n)\in K_\epsilon$. It follows, by the contrapositive, that

$$\sup_{n\in\mathbb{N}}\tilde{\mathbf{P}}\big(\Gamma'(\tilde f_n)\mathcal{B}_n^+(\tilde f_n,\tilde f_n)\notin K_\epsilon\big)\le\sup_{n\in\mathbb{N}}\tilde{\mathbf{P}}\big(\Gamma'(\tilde f_n)\mathcal{B}_n^-(\tilde f_n,\tilde f_n)\notin K_\epsilon^-\big)+\sup_{n\in\mathbb{N}}\tilde{\mathbf{P}}\big(\mathcal{D}_n^0(\tilde f_n,\tilde f_n)\notin B_M\big).$$

Since each term above is of order ϵ, the proof of the Lemma will be complete if we verify that K_ϵ is a weakly compact subset of $L^1_{t,x,v}$. By classical compactness criteria, it suffices to verify the following:

$$\lim_{R\to\infty}\sup_{f\in K_\epsilon}\int_0^T\iint_{\mathbb{R}^{2d}}1_{\{|x|+|v|>R\}}|f|\mathrm{d}x\mathrm{d}v\mathrm{d}t=0. \tag{7.64}$$

$$\lim_{\delta\to0}\sup_{f\in K_\epsilon}\sup_{|E|\le\delta}\int_0^T\iint_{\mathbb{R}^{2d}}1_E|f|\mathrm{d}x\mathrm{d}v\mathrm{d}t=0, \tag{7.65}$$

where in (7.65) the supremum is taken over all measurable $E\subseteq[0,T]\times\mathbb{R}^{2d}$ with Lebesgue measure $|E|<\delta$. To verify (7.64) and (7.65), note that for all $j>1$, by construction of K_ϵ

$$\sup_{f\in K_\epsilon}\int_0^T\iint_{\mathbb{R}^{2d}}1_{\{|x|+|v|>R\}}|f|\mathrm{d}x\mathrm{d}v\mathrm{d}t\le j\sup_{g\in K_\epsilon^-}\int_0^T\iint_{\mathbb{R}^{2d}}1_{\{|x|+|v|>R\}}|g|\mathrm{d}x\mathrm{d}v\mathrm{d}t+\frac{M_\epsilon}{\log j},$$

and

$$\sup_{f\in K_\epsilon}\sup_{|E|<\delta}\int_0^T\iint_{\mathbb{R}^{2d}}1_E|f|\mathrm{d}x\mathrm{d}v\mathrm{d}t\le j\sup_{g\in K_\epsilon^-}\sup_{|E|<\delta}\int_0^T\iint_{\mathbb{R}^{2d}}1_E|g|\mathrm{d}x\mathrm{d}v\mathrm{d}t+\frac{M_\epsilon}{\log j}.$$

First taking $R\to\infty$ and using the $L^1_{t,x,v}$ weak compactness of K_ϵ^- and then sending $j\to\infty$ yields (7.64) and (7.65). □

Lemma 7.6.7. *For each $\Gamma \in \mathcal{R}'$, the laws of $\{\Gamma(\tilde{f}_n)\}_{n\in\mathbb{N}}$ are tight on $C_t([L^1_{x,v}]_w) \cap L^1_{t,x}(\mathcal{M}^*_v)$.*

Proof. We will check that $\{\Gamma(\tilde{f}_n)\}_{n\in\mathbb{N}}$ induces tight laws on the space $L^1_{t,x}(\mathcal{M}^*_v)$ by first verifying the requirements of the L^1 velocity averaging Lemma 7.4.2 and deducing that for each $\varphi \in C^\infty_c(\mathbb{R}^d)$, $\{\langle\Gamma(\tilde{f}_n), \varphi\rangle\}_{n\in\mathbb{N}}$ induces tight laws on $L^1_{t,x}$ and then applying Lemma B.1.11 to conclude that $\{\Gamma(\tilde{f}_n)\}_{n\in\mathbb{N}}$ induces tight law on $L^1_{t,x}(\mathcal{M}^*_v)$.

Observe that for each $n \in \mathbb{N}$, $\Gamma(\tilde{f}_n)$ is a weak martingale solution to the stochastic kinetic equation driven by $\Gamma'(\tilde{f}_n)\mathcal{B}_n(\tilde{f}_n, \tilde{f}_n)$, starting from $\Gamma(\tilde{f}^0_n)$, with noise coefficients σ^n. By Proposition 7.5.6 on the approximating scheme, and the fact that $\Gamma(z) \lesssim |z|$, we can easily conclude that $\Gamma(\tilde{f}_n)$ and $\Gamma'(\tilde{f}_n)\mathcal{B}_n(\tilde{f}_n, \tilde{f}_n)$ belong to $L^{\infty-}(\Omega; L^1_{t,x,v} \cap L^\infty_{t,x,v})$ and $\Gamma(f^0_n)$ is in $L^1_{x,v} \cap L^\infty_{x,v}$. Also, by assumption, $\{\sigma^n\}_{n\in\mathbb{N}}$ satisfy Hypothesis 7.1.2 uniformly.

Next, since $|\Gamma(z)| \lesssim |z|$, and $\{f^0_n\}_{n\in\mathbb{N}}$ is uniformly integrable, then $\{\Gamma(f^0_n)\}_{n\in\mathbb{N}}$ is uniformly integrable. Similarly, the uniform estimates (7.55) imply that for $p \in [1, \infty)$

$$\sup_{n\in\mathbb{N}} \tilde{\mathbf{E}} \big\|(1 + |x|^2 + |v|^2)\Gamma(\tilde{f}_n)\big\|^p_{L^\infty_t(L^1_{x,v})} < \infty. \tag{7.66}$$

Also, Lemma 7.6.4 implies that $\{\Gamma'(\tilde{f}_n)\mathcal{B}_n(\tilde{f}_n, \tilde{f}_n)\}_{n\in\mathbb{N}}$ is uniformly bounded in $L^{\infty-}(\Omega; L^1_{t,x,v})$, while Lemmas 7.6.5 and 7.6.6 imply that $\{\Gamma'(\tilde{f}_n)\mathcal{B}_n(\tilde{f}_n, \tilde{f}_n)\}_{n\in\mathbb{N}}$ also induce tight laws on $[L^p_{t,x,v}]_{\mathrm{w}}$. Finally, we see by Chebyshev that

$$\tilde{\mathbf{P}}(\|\langle\Gamma(\tilde{f}_n), \varphi\rangle \mathbb{1}_{|x|>R}\|_{L^1_{t,x}} > \eta) \lesssim \frac{1}{\eta R^2} \tilde{\mathbf{E}} \big\|(1 + |x|^2 + |v|^2)\Gamma(\tilde{f}_n)\big\|_{L^\infty_t(L^1_{x,v})},$$

and therefore the right-hand side vanishes uniformly in n as $R \to \infty$. Hence, we meet all the requirements of Lemma 7.4.2 to conclude that $\{\langle\Gamma(\tilde{f}_n),\varphi\rangle\}_{n\in\mathbb{N}}$ induces tight laws on $L^1_{t,x}$.

To check that $\{\Gamma(\tilde{f}_n)\}_{n\in\mathbb{N}}$ induces tight laws on the space $C_t([L^1_{x,v}]_w)$, by Lemma B.1.8 it suffices to show that for each $\varphi\in C_c^\infty(\mathbb{R}^{2d})$, the sequence $\{\langle\Gamma(\tilde{f}_n),\varphi\rangle\}_{n\in\mathbb{N}}$ induces tight laws on $C[0,T]$ and

$$\sup_n \tilde{\mathbf{E}}\|\Gamma(\tilde{f}_n)\|_{L^\infty_t(L^1_{x,v})}<\infty,$$

$$\lim_{R\to\infty}\sup_n \tilde{\mathbf{E}}\|\Gamma(\tilde{f}_n)\mathbb{1}_{|x|^2+|v|^2>R}\|_{L^\infty_t(L^1_{x,v})}=0,$$

$$\lim_{L\to\infty}\sup_n \tilde{\mathbf{E}}\|\Gamma(\tilde{f}_n)\mathbb{1}_{|\Gamma(\tilde{f}_n)|>L}\|_{L^\infty_t(L^1_{x,v})}=0.$$

The first two follow from (7.66), while the last follows from the fact that $|\Gamma(z)|\leq C|z|$ for some constant C, implies that

$$|\Gamma(\tilde{f}_n)|\mathbb{1}_{|\Gamma(\tilde{f}_n)|>L}\leq|\tilde{f}_n|\mathbb{1}_{|\tilde{f}_n|>L/C}$$

and therefore

$$\lim_{L\to\infty}\sup_n \tilde{\mathbf{E}}\|\Gamma(\tilde{f}_n)\mathbb{1}_{|\Gamma(\tilde{f}_n)|>L}\|_{L^\infty_t(L^1_{x,v})}\leq\lim_{L\to\infty}\frac{1}{\log L/C}\sup_n\tilde{\mathbf{E}}\|\tilde{f}_n\log\tilde{f}_n\|_{L^\infty_t(L^1_{x,v})}=0.$$

To see this, use the weak form to obtain the decomposition $\langle\Gamma(\tilde{f}_n),\varphi\rangle=I^{n,1}+I^{n,2}$, where the continuous processes $(I^{n,1}_t)_{t=0}^T$ and $(I^{n,2}_t)_{t=0}^T$ are defined via:

$$\begin{aligned}I^{n,1}_t&=\iint_{\mathbb{R}^{2d}}\Gamma(f^0_n)\varphi\mathrm{d}x\mathrm{d}v+\int_0^t\iint_{\mathbb{R}^{2d}}\Gamma(\tilde{f}_n)[v\cdot\nabla_x\varphi+\mathcal{L}_{\sigma^n}\varphi]\mathrm{d}x\mathrm{d}v\mathrm{d}s\\&\qquad+\sum_{k=1}^\infty\int_0^t\iint_{\mathbb{R}^{2d}}\Gamma(\tilde{f}_n)\sigma^n_k\cdot\nabla_v\varphi\,\mathrm{d}x\mathrm{d}v\mathrm{d}\beta_k(s).\\I^{n,2}_t&=\int_0^t\iint_{\mathbb{R}^{2d}}\Gamma'(\tilde{f}_n)\mathcal{B}_n(\tilde{f}_n,\tilde{f}_n)\varphi\mathrm{d}x\mathrm{d}v\mathrm{d}s.\end{aligned}$$

Arguing as in Lemma 7.3.4, using the uniform bounds, and a Sobolev embedding, there exists an an $\alpha>0$ and $p>1$ such that $\{I^{n,1}\}_{n\in\mathbb{N}}$ is a bounded sequence in

$L^p(\tilde{\Omega}; C_t^\alpha)$. Next observe that by Lemmas 7.6.5 and 7.6.6, the sequence $\{\partial_t I^{n,2}\}_{n\in\mathbb{N}}$ induces tight laws on $L^1[0,T]$ endowed with the weak topology.

Let $\epsilon > 0$ and let K^1_ϵ be the closed ball of radius ϵ^{-1} in C_t^α. In addition, choose a uniformly integrable subset of $L^1[0,T]$, denoted $\hat{K}^2_\epsilon$, such that

$$\sup_{n\in\mathbb{N}} \tilde{\mathbf{P}}(\partial_t I^{2,n} \notin \hat{K}^2_\epsilon) < \epsilon.$$

Define K^2_ϵ to be the anti-derivatives of $\hat{K}^2_\epsilon$, that is:

$$K^2_\epsilon = \big\{ f \in C[0,T] \mid f(0) = 0 \text{ and there exists } g \in \hat{K}^2_\epsilon \text{ such that } \partial_t f = g \big\}.$$

Finally, let K_ϵ be the algebraic sum (in $C[0,T]$) of K^1_ϵ and K^2_ϵ. In view of our decomposition, it follows that

$$\sup_{n\in\mathbb{N}} \tilde{\mathbf{P}}(\langle \Gamma(\tilde{f}_n), \varphi\rangle \notin K_\epsilon) \leq \sup_{n\in\mathbb{N}} \tilde{\mathbf{P}}(I^{n,1} \notin K^1_\epsilon) + \sup_{n\in\mathbb{N}} \tilde{\mathbf{P}}(I^{n,2} \notin K^2_\epsilon).$$

Each of the probabilities above are of order ϵ. Since, by construction, K^1_ϵ and K^2_ϵ are compact of $C[0,T]$ (by Arzelà-Ascoli), it follows that K_ϵ is itself compact in $C[0,T]$. This completes the proof. □

Lemma 7.6.8. *The sequence $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ induces tight laws on the space $C_t([L^1_{x,v}]_w) \cap L^1_{t,x}(\mathcal{M}^*_v)$.*

Proof. Let us begin by verifying that $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ induces a tight sequence of laws on $L^1_{t,x}(\mathcal{M}^*_v)$. From the uniform bounds, we know that $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ is uniformly bounded in $L^1(\tilde{\Omega} \times [0,T] \times \mathbb{R}^{2d})$. By the appendix Lemma B.1.11, it suffices to check that for each $\varphi \in C_c^\infty(\mathbb{R}^d_v)$, the sequence $\{\langle \tilde{f}_n, \varphi\rangle\}_{n\in\mathbb{N}}$ induces a tight sequence of laws on $L^1_{t,x}$. For this purpose, we will use the compactness criterion given in appendix

Lemma B.1.4, together with Lemma 7.6.7. Indeed, recall the definition of $\Gamma_m(z)$ in equation (7.56), then for each $m \in \mathbb{N}$, we have the decomposition

$$\langle \tilde{f}_n, \varphi \rangle = \langle \Gamma_m(\tilde{f}_n), \varphi \rangle + \langle f_n - \Gamma_m(\tilde{f}_n), \varphi \rangle.$$

By Lemma 7.6.7, the sequence $\{\langle \Gamma_m(\tilde{f}_n), \varphi \rangle\}_{n\in\mathbb{N}}$ induces a tight sequence of laws on $L^1_{t,x}$. Hence, by Lemma B.1.4, it only remains to verify that

$$\lim_{m\to\infty} \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}} \big\| \langle \tilde{f}_n - \Gamma_m(\tilde{f}_n), \varphi \rangle \big\|_{L^1_{t,x}} = 0.$$

Towards this end, note the elementary inequality: for all $R > 1$ and $z > 0$,

$$|\Gamma_m(z) - z| \leq \frac{R}{m} z + z 1_{z \geq R} \leq \frac{R}{m} z + |\log R|^{-1} z |\log z|.$$

Hence, for all $m \in \mathbb{N}$ and $R > 1$, we have the inequality

$$\begin{aligned} \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}} \big\| \langle \tilde{f}_n - \Gamma_m(\tilde{f}_n), \varphi \rangle \big\|_{L^1_{t,x}} &\leq \frac{R}{m} \|\varphi\|_{L^\infty_v} \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}} \|\tilde{f}_n\|_{L^1_{t,x,v}} \\ &\quad + |\log R|^{-1} \|\varphi\|_{L^\infty_v} \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}} \|\tilde{f}_n \log \tilde{f}_n\|_{L^1_{t,x,v}}. \end{aligned}$$

Taking first $m \to \infty$ and then $R \to \infty$ gives the claim.

The next step is to check that the sequence $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ induces a tight sequence of laws on $C_t([L^1_{x,v}]_w)$. In view of the uniform bounds (7.55) and tightness criterion on $C_t([L^1_{x,v}]_{\mathrm{w}})$ given in Lemma B.1.8, it suffices to verify that for all $\varphi \in C^\infty_c(\mathbb{R}^{2d})$, the sequence $\{\langle \tilde{f}_n, \varphi \rangle\}_{n\in\mathbb{N}}$ induces tight laws on the space $C[0,T]$. Again, for each $m \in \mathbb{N}$ we have the decomposition

$$\langle \tilde{f}_n, \varphi \rangle = \langle \Gamma_m(\tilde{f}_n), \varphi \rangle + \langle \tilde{f}_n - \Gamma_m(\tilde{f}_n), \varphi \rangle.$$

Moreover, the sequence $\{\langle \Gamma_m(\tilde{f}_n), \varphi \rangle\}_{n\in\mathbb{N}}$ induces tight laws on $C[0,T]$ by Lemma 7.6.7. Arguing in a similar way as above, we find that

$$\lim_{m\to\infty} \sup_{n\in\mathbb{N}} \tilde{\mathbf{E}} \big\| \langle \tilde{f}_n - \Gamma_m(\tilde{f}_n), \varphi \rangle \big\|_{L^\infty_t} = 0.$$

Therefore by Lemma B.1.4 $\{\langle \tilde{f}_n, \varphi\rangle\}_{n\in\mathbb{N}}$ is tight on $C[0,T]$. □

Proof of Proposition 7.6.1

For each $n \in \mathbb{N}$, introduce random variables $\tilde{X}_n, \tilde{Y}_n$, and $\tilde{Z}_n$ by setting

$$\tilde{X}_n = \big((1+|x|^2+|v|^2+|\log \tilde{f}_n|)\tilde{f}_n, \mathcal{D}_n^0(\tilde{f}_n)\big)$$

$$\tilde{Y}_n = \big(\tilde{f}_n, \{\tilde{\beta}_k\}_{k\in\mathbb{N}}\big)$$

$$\tilde{Z}_n = \Big\{\big(\Gamma_m(\tilde{f}_n), \gamma_m(\tilde{f}_n), \Gamma'_m(\tilde{f}_n)\mathcal{B}_n^-(\tilde{f}_n,\tilde{f}_n), \Gamma'_m(\tilde{f}_n)\mathcal{B}_n^+(\tilde{f}_n,\tilde{f}_n)\big)\Big\}_{m\in\mathbb{N}}$$

The random variables $\tilde{X}_n, \tilde{Y}_n$, and $\tilde{Z}_n$ induce laws defined on the spaces E, F, and G respectively, where

$$E = [L^1_t(C_0(\mathbb{R}^{2d}))]'_* \times \mathcal{M}^*_{t,x,v}$$

$$F = L^1_{t,x}(\mathcal{M}^*_v) \cap C_t([L^1_{x,v}]_w) \times [C_t]^\infty$$

$$G = \big[[L^1_{t,x}(\mathcal{M}^*_v) \cap C_t([L^1_{x,v}]_w)]^2 \times [L^1_{t,x,v}]^2_w\big]^\infty.$$

To be clear, we use $[L^1_t(C_0(\mathbb{R}^{2d}))]'_*$ to denote the dual of $L^1_t(C_0(\mathbb{R}^{2d}))$ endowed with the weak star topology.

Our first observation is that the sequence $\{\tilde{X}_n\}_{n\in\mathbb{N}}$ induces tight laws on E. For this, we use the fact that $L^\infty_t(L^1_{x,v})$ embeds isometrically into the space $L^\infty_t(\mathcal{M}_{x,v})$, which in turn embeds isometrically into $[L^1_t(C_0(\mathbb{R}^{2d}))]'$ by classical duality results on Lebesgue-Bochner spaces. Also, $L^1_{t,x,v}$ embeds isometrically into $\mathcal{M}_{t,x,v}$. Since bounded sets in $L^\infty_t(L^1_{x,v}) \times L^1_{t,x,v}$ are compact in E, the uniform bounds (7.55) and Banach Alaoglu yield the tightness claim.

Next we observe that $\{\tilde{Y}_n\}_{n\in\mathbb{N}}$ induces tight laws on F. This follows from

Lemma 7.6.8 and classical facts about Brownian motions. Finally, by Lemmas 7.6.5, 7.6.6, and 7.6.7 it follows that the sequence $\{\tilde{Z}_n\}_{n\in\mathbb{N}}$ induces tight laws on G. Combining these observations, we find that the sequence $\{(\tilde{X}_n, \tilde{Y}_n, \tilde{Z}_n)\}_{n\in\mathbb{N}}$ induces tight laws on $E \times F \times G$.

Apply the Jakubowski/Skorohod Theorem B.1.2 (working on a subsequence if necessary) to obtain a new probability space $(\Omega, \mathcal{F}, \mathbf{P})$, random variables (X, Y, Z) on $E \times F \times G$, and a sequence of maps $\{\tilde{T}_n\}$ satisfying Part 1 of Proposition 7.6.1. First observe that the uniform bounds and the explicit representation guarantees that $X_n(\omega) \in L^\infty_t(L^1_{x,v}) \times L^1_{t,x,v}$ for almost all $\omega \in \Omega$ and $n \in \mathbb{N}$. Thus, Part 1 now yields that $\{f_n\}_{n\in\mathbb{N}}$ satisfies the uniform bounds (7.55) with $\mathbf{E}$ in place of $\tilde{\mathbf{E}}$. This gives the first claim in Part 2 of Proposition 7.6.1. Theorem B.1.2 also guarantees that the sequence $\{X_n\}_{n\in\mathbb{N}}$ defined by $X_n = \tilde{X}_n \circ \tilde{T}_n$ converges pointwise on Ω to X in the space E. In particular, there exists a random constant $C(\omega)$ such that

$$\sup_{n\in\mathbb{N}} \|(1 + |x|^2 + |v|^2 + |\log f_n(\omega)|) f_n(\omega)\|_{[L^1_t(C_0(\mathbb{R}^{2d}))]'} \le C(\omega).$$

$$\sup_{n\in\mathbb{N}} \|\mathcal{D}^0_n(f_n)(\omega)\|_{\mathcal{M}_{t,x,v}} \le C(\omega).$$

Using again the isometric embedding of $L^\infty_t(L^1_{x,v})$ into $[L^1_t(C_0(\mathbb{R}^{2d}))]'$ and $L^1_{t,x,v}$ into $\mathcal{M}_{t,x,v}$, together with the fact that $X_n(\omega) \in L^\infty_t(L^1_{x,v}) \times L^1_{t,x,v}$, this completes the proof of Part 2. To obtain the remaining parts of Proposition 7.6.1, let $\overline{D}$ be the second component of X, and denote

$$Y = \big(f, \{\beta_k\}_{k\in\mathbb{N}}\big).$$

$$Z = \Big\{\big(\overline{\Gamma_m(f)}, \overline{\gamma_m(f)}, \mathcal{B}^-_m, \mathcal{B}^+_m\big)\Big\}_{m\in\mathbb{N}}.$$

Part 3 follows easily from Part 1 and the martingale representation theorem. Part 4 follows from the pointwise convergence of $\{Y_n\}_{n\in\mathbb{N}}$ towards Y in the space F. Part 5 follows from the pointwise convergence of $\{Z_n\}_{n\in\mathbb{N}}$ towards Z and $\{X_n\}_{n\in\mathbb{N}}$ to X. This completes the proof of Proposition 7.6.1.

Preliminary identification

As our first application of Proposition 7.6.1, we send $n \to \infty$, but the limit passage is in a preliminary sense. Namely, we do not yet obtain the renormalized form for f, but we obtain a stochastic kinetic equation for a strong approximation $\overline{\Gamma_m(f)}$. In fact, using Proposition 7.6.1, we will prove:

Corollary 7.6.9. *For all $m \in \mathbb{N}$, the process $\overline{\Gamma_m(f)}$ is a renormalized weak martingale solution to the stochastic kinetic equation driven by $\mathcal{B}_m^+ - \mathcal{B}_m^-$, starting from $\Gamma_m(f_0)$, with noise coefficients $\sigma = \{\sigma_k\}_{k\in\mathbb{N}}$. Moreover, $\mathbf{P}$ almost surely, $\overline{\Gamma_m(f)}$ belongs to $L^\infty_{t,x,v}$ and has strongly continuous sample paths in $C_t(L^1_{x,v})$.*

Proof. Fix an $m \in \mathbb{N}$. First, using the uniform bounds and the convergence results obtained in Proposition 7.6.1, we verify the hypotheses of the stability result for martingale solutions of stochastic kinetic equations, Proposition 7.3.5. Namely, we will analyze the sequence $\{\Gamma_m(f_n)\}_{n\in\mathbb{N}}$. Once we verify Parts $1-3$ of Proposition 7.3.5, we may conclude that the process $\overline{\Gamma_m(f)}$ is a weak martingale solution to the stochastic kinetic equation driven by $\mathcal{B}_m^+ - \mathcal{B}_m^-$, starting from $\Gamma_m(f_0)$, with noise coefficients $\sigma = \{\sigma_k\}_{k\in\mathbb{N}}$. The next step will be to show that the solution is actually a renormalized weak martingale solution, applying the renormalization Proposition 7.3.8. Finally we will show strong continuity by applying Lemma 7.3.9 on our renormalized weak martingale solution.

To verify Part 1 of Proposition 7.3.5, let us first check that the process $\Gamma_m(f_n)$ is a weak martingale solution to the stochastic kinetic equation driven by $\Gamma'_m(f_n)\mathcal{B}_n(f_n, f_n)$, starting from $\Gamma_m(f_n^0)$, relative to the noise coefficients σ^n and the Brownian motions $\{\beta_k^n\}_{k\in\mathbb{N}}$ obtained in 7.6.1. Indeed, $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ is a renormalized weak martingale solution to the stochastic kinetic equation driven by $\mathcal{B}_n(\tilde{f}_n, \tilde{f}_n)$, starting from f_n^0, relative to the noise coefficients σ^n and the Brownian motions $\{\tilde{\beta}_k^n\}_{k\in\mathbb{N}}$. The claim can now be checked by using the explicit expression for $\{f_n\}_{n\in\mathbb{N}}$ and $\{\beta_k^n\}_{k\in\mathbb{N}}$ in terms of the maps $\{\tilde{T}_n\}_{n\in\mathbb{N}}$ together with the fact that $\Gamma_m \in \mathcal{R}'$.

To verify Part 2 of Proposition 7.3.5, from the uniform bounds in Proposition 7.6.1 and the fact that $\Gamma_m(z) \leq z$, it follows that the sequence $\{\Gamma_m(f_n)\}_{n\in\mathbb{N}}$ is uniformly bounded in $L^2(\Omega; L_t^\infty(L_{x,v}^1))$. Also, Lemma 7.6.4 and Part 1 of Proposition 7.6.1 imply that $\{\Gamma'_m(f_n)\mathcal{B}_n^-(f_n, f_n)\}_{n\in\mathbb{N}}$ and $\{\Gamma'_m(f_n)\mathcal{B}_n^+(f_n, f_n)\}_{n\in\mathbb{N}}$ are uniformly bounded in $L^2(\Omega; L_{t,x,v}^1)$. Combining this with the pointwise convergences from Part 5 of Proposition 7.6.1, we easily verify (7.31) and (7.32).

Finally Part 3 of Proposition 7.3.5 follows from the convergences from Part 4 of Proposition 7.6.1 together with Hypotheses 7.5.2 and 7.5.1 regarding the sequences $\{\sigma^n\}_{n\in\mathbb{N}}$ and $\{f_n^0\}_{n\in\mathbb{N}}$.

Next we argue that $\overline{\Gamma_m(f)}$ is actually a *renormalized* weak martingale solution. Indeed, by the conditions on the noise coefficients σ in Hypotheses (H3) and (H4) this will follow from Proposition 7.3.8 as soon as $\overline{\Gamma_m(f)} \in L^{\infty-}(\Omega\times[0,T]\times\mathbb{R}^{2d})$. To argue this, we note that since $\Gamma_m(z) \leq m$ and $\Gamma_m(z) \leq z$, this gives the following uniform bounds in $L^\infty_{\omega,t,x,v}$ and $L^1_{\omega,t,x,v}$,

$$\sup_n \|\Gamma_m(f_n)\|_{L^\infty(\Omega\times[0,T]\times\mathbb{R}^{2d})} < m < \infty$$

$$\sup_n \|\Gamma_m(f_n)\|_{L^1(\Omega\times[0,T]\times\mathbb{R}^{2d})} \le T \sup_n \mathbf{E}\|f_n\|_{L^\infty_t(L^1_{x,v})} < \infty.$$

Therefore, by interpolation, $\{\Gamma_m(f_n)\}_{n\in\mathbb{N}} \in L^p(\Omega\times[0,T]\times\mathbb{R}^{2d})$ uniformly in n for each $p \in [1,\infty]$ and $m \ge 1$. Using the weak sequential compactness of $L^p(\Omega\times[0,T]\times\mathbb{R}^{2d})$ for $p \in (1,\infty)$, weak-* sequential compactness of $L^\infty(\Omega\times[0,T]\times\mathbb{R}^{2d})$, and the fact that by Proposition 7.6.1, $\mathbf{P}$ almost surely, $\Gamma_m(f_n) \to \overline{\Gamma_m(f)}$ in $C_t([L^1_{x,v}]_w)$, we can conclude that the limit $\overline{\Gamma_m(f)}$ must belong to $L^p(\Omega\times[0,T]\times\mathbb{R}^{2d})$ for every $p \in [1,\infty]$.

Finally we show that process $t \mapsto \overline{\Gamma_m(f_t)}$ has continuous sample paths in $L^1_{x,v}$ with the strong topology. Observe that any sequence converging strongly in $L^2_{x,v}$ and weakly in $L^1_{x,v}$ also converges strongly in $L^1_{x,v}$. Therefore, since $\overline{\Gamma_m(f)} \in C_t([L^1_{x,v}]_w)$ with probability one, it suffices to show that $\overline{\Gamma_m(f)} \in C_t(L^2_{x,v})$ with probability one. However, since $\overline{\Gamma_m(f)}$ is a renormalized weak martingale solution, by Lemma 7.3.9 it is sufficient to show that $\overline{\Gamma_m(f)}$ belongs to $L^\infty_t(L^2_{x,v})$ $\mathbf{P}$ almost surely. Since Proposition 7.6.1 implies that $\overline{\Gamma_m(f)}$ also belongs to $L^\infty_t(L^1_{x,v})$ $\mathbf{P}$ almost surely and $\overline{\Gamma_m(f)}$ belongs to $L^\infty(\Omega\times[0,T]\times\mathbb{R}^{2d})$, we can conclude, again by interpolation, that $\overline{\Gamma_m(f)}$ belongs to $L^\infty_t(L^2_{x,v})$ $\mathbf{P}$ almost surely. □

In fact, this preliminary identification of $\overline{\Gamma_m(f)}$ allows us to upgrade the continuity properties on f from weakly continuous to strongly continuous. This is the content of the following corollary.

Corollary 7.6.10. *The sample paths of f belong* $\mathbf{P}$ *almost surely to* $C_t(L^1_{x,v})$. *Moreover as* $m \to \infty$, *the sequence* $\{\overline{\Gamma_m(f)}\}_{m\in\mathbb{N}}$ *converges* $\mathbf{P}$ *a.s. to* f *in* $C_t(L^1_{x,v})$.

Proof. Recall, by Corollary 7.6.9, $\overline{\Gamma_m(f)}$ belongs to $C_t(L^1_{x,v})$, hence it suffices to show that $\{\overline{\Gamma_m(f)}\}_{m\in\mathbb{N}}$ converges $\mathbf{P}$ a.s. to f in $L^\infty_t(L^1_{x,v})$. This is accomplished by applying Proposition 7.6.1 to conclude that for each $t \in [0,T]$, $f_n(t) - \Gamma_m(f_n(t)) \to f_t - \overline{\Gamma_m(f)}_t$ weakly in $L^1_{x,v}$, $\mathbf{P}$ almost-surely, then using weak lower semi-continuity of the $L^1_{x,v}$ norm to obtain the $\mathbf{P}$ almost-sure inequality

$$\begin{aligned}\sup_{t\in[0,T]} \|f_t - \overline{\Gamma_m(f)}_t\|_{L^1_{x,v}} &\le \liminf_{n\to\infty} \sup_{t\in[0,T]} \|f_n(t) - \Gamma_m(f_n)(t)\|_{L^1_{x,v}} \\ &\le \frac{1}{\sqrt{m}} \sup_n \|f_n\|_{L^\infty_t(L^1_{x,v})} + \sup_n \|f_n \mathbb{1}_{f_n \ge \sqrt{m}}\|_{L^\infty_t(L^1_{x,v})},\end{aligned}$$

where in the last inequality we used the fact that

$$|x - \Gamma_m(x)| \le \frac{1}{\sqrt{m}} x + x\mathbb{1}_{|x|\ge\sqrt{m}}.$$

In view of Part 2 in Proposition 7.6.1, for $\mathbf{P}$ almost all $\omega \in \Omega$, the sequence $\{f_n(\omega)\}_{n\in\mathbb{N}}$ is uniformly integrable in $L^\infty_t(L^1_{x,v})$. Taking $m \to \infty$ on both sides of the inequality above completes the proof. □

Analysis of the Renormalized Collision Operator

In this section, we prepare for the passage of $m \to \infty$. By applying the renormalization lemma for martingale solutions of stochastic kinetic equations, we obtain the following immediate corollary.

Corollary 7.7.1. *For all* $m \in \mathbb{N}$, *the process* $\log(1 + \overline{\Gamma_m(f)})$ *is a weak martingale* *solution to the stochastic kinetic transport equation driven by* $(1+\overline{\Gamma_m(f)})^{-1}[\mathcal{B}_m^+ - \mathcal{B}_m^-]$, *starting from* $\log(1+\Gamma_m(f_0))$.

Our primary focus is to analyze the limiting behavior of the sequence $\{\mathcal{B}_m^+\}_{m\in\mathbb{N}}$. The main source of difficulty here is that this sequence is not bounded in $L^1(\Omega \times [0,T]\times\mathbb{R}^{2d})$. This is natural in the sense that we expect $\mathcal{B}_m^+$ to be close to $\mathcal{B}^+(f,f)$ as we relax the truncation parameter $m \in \mathbb{N}$. In fact, we know that the main strategy in dealing with $\mathcal{B}^+(f,f)$ is to renormalize with $\Gamma'(f)\mathcal{B}^+(f,f)$ before we can hope for an estimate it in $L^1(\Omega \times [0,T] \times \mathbb{R}^{2d})$. The main result of this section is the following:

Proposition 7.7.2. *For any* $\phi \in L^\infty_{t,x,v}$ *as* $m \to \infty$, *the following convergences hold:*

$$\left\{\left\langle \frac{\mathcal{B}_m^-}{1+\overline{\Gamma_m(f)}}, \phi \right\rangle\right\}_{m\in\mathbb{N}} \to \left\langle \frac{\mathcal{B}^-(f,f)}{1+f}, \phi\right\rangle \quad in \quad L^2(\Omega),$$
$$\left\{\left\langle \frac{\mathcal{B}_m^+}{1+\overline{\Gamma_m(f)}}, \phi \right\rangle\right\}_{m\in\mathbb{N}} \to \left\langle \frac{\mathcal{B}^+(f,f)}{1+f}, \phi\right\rangle \quad in \quad L^2(\Omega).$$

The most challenging part of the analysis is analyzing the positive part of the collision operator. To analyze the $m \to \infty$ limit, we must analyze the consequences of the pointwise (in ω) convergence of $f_n(\omega)$ towards $f(\omega)$ in the space $L^1_{t,x}(\mathcal{M}_v^*)$. In fact, this has not been used so far in the proof.

Lemma 7.7.3. *As* $n \to \infty$, *the following convergence holds* $\mathbf{P}$ *almost surely:*

$$\frac{\mathcal{B}_n^+(f_n,f_n)}{1+\langle f_n, 1\rangle} \to \frac{\mathcal{B}^+(f,f)}{1+\langle f,1\rangle} \quad in \quad L^1_{t,x}(\mathcal{M}_v^*).$$

Proof. The proof follows essentially the same manipulations as in [36] and [65], carried out pointwise in $\omega \in \Omega$. We sketch the proof only to convince the reader that the compactness properties obtained in Proposition 7.6.1 are sufficient to deduce the claim in the same way as for the deterministic theory, without pulling any further subsequences (potentially depending on ω). Let $\varphi \in C_c(\mathbb{R}^d_v)$. We will fix an $\omega \in \Omega$ and mostly omit dependence on this variable throughout the proof. A change of variables from $(v, v_*) \to (v', v'_*)$ and an application of Fubini yields the identities

$$\begin{aligned} \left\langle \frac{\mathcal{B}^+_n(f_n, f_n)}{1 + \langle f_n, 1\rangle}, \varphi \right\rangle &= \left\langle f_n \frac{\mathcal{L}_n f_n}{1 + \langle f_n, 1\rangle}, 1 \right\rangle. \\ \left\langle \frac{\mathcal{B}^+(f, f)}{1 + \langle f, 1\rangle}, \varphi \right\rangle &= \left\langle f \frac{\mathcal{L} f}{1 + \langle f, 1\rangle}, 1 \right\rangle, \end{aligned} \tag{7.67}$$

where $\mathcal{L}_n$ is the linear operator on $L^1(\mathbb{R}^d_v)$ defined by

$$\mathcal{L}_n f(v) = \iint_{\mathbb{R}^d \times \mathbb{S}^{d-1}} f_* b_n(v - v_*) \varphi' \mathrm{d}v_* \mathrm{d}\theta,$$

and $\mathcal{L}$ is defined analogously, but with b replacing b_n. Since $\{f_n(\omega)\}_{n\in\mathbb{N}}$ converges to $f(\omega)$ in $L^1_{t,x}(\mathcal{M}^*_v)$ and is tight as a sequence in $L^1_{t,x,v}$, while b_n converges to b pointwise on $\mathbb{R}^d \times \mathbb{S}^{d-1}$ and is bounded in $L^\infty(\mathbb{R}^d \times \mathbb{S}^{d-1})$ by Hypothesis 7.5.3, one can deduce that, $\mathbf{P}$ almost surely, both $\{\mathcal{L}_n f_n\}_{n\in\mathbb{N}} \to \mathcal{L} f$ and $\{\langle f, 1\rangle\}_{n\in\mathbb{N}} \to \langle f, 1\rangle$ in measure on $[0, T] \times \mathbb{R}^{2d}$. Therefore, $\mathbf{P}$ almost surely

$$\left\{ \frac{\mathcal{L}_n f_n}{1 + \langle f_n, 1\rangle} \right\}_{n\in\mathbb{N}} \to \frac{\mathcal{L} f}{1 + \langle f, 1\rangle} \quad \text{in measure on} \quad [0, T] \times \mathbb{R}^{2d}. \tag{7.68}$$

Using the uniform bounds on $\{b_n\}_{n\in\mathbb{N}}$ in $L^\infty(\mathbb{R}^d \times \mathbb{S}^{d-1})$, the sequence in (7.68) is also uniformly bounded in $L^\infty_{t,x,v}$, pointwise in ω. Applying the second part of the product lemma B.1.12 gives

$$\left\{ f_n \frac{\mathcal{L}_n f_n}{1 + \langle f_n, 1\rangle} \right\}_{n\in\mathbb{N}} \to f \frac{\mathcal{L} f}{1 + \langle f, 1\rangle} \quad \text{in} \quad L^1_{t,x}(\mathcal{M}^*_v).$$

An approximation argument (since 1 does not belong to $C_0(\mathbb{R}^d_v)$) and the pointwise (in ω) uniform bounds on $\{f_n\}_{n\in\mathbb{N}}$ from Proposition 7.6.1 yields the $\mathbf{P}$ almost sure convergence

$$\left\langle f_n \frac{\mathcal{L}_n f_n}{1+\langle f_n, 1\rangle}, 1\right\rangle \to \left\langle f \frac{\mathcal{L} f}{1+\langle f, 1\rangle}, 1\right\rangle \quad \text{in} \quad L^1_{t,x}.$$

In view of the identities (7.67), this completes the proof. □

The purpose of the next lemma is to reduce our analysis of $\mathcal{B}^+_m$ to regions where there are no concentrations in $\{f_n\}_{n\in\mathbb{N}}$.

Lemma 7.7.4. *As $R \to \infty$, the following limit holds $\mathbf{P}$ almost surely:*

$$\frac{\mathcal{B}^+_n(f_n, f_n)}{1+\langle f_n, 1\rangle} \mathbb{1}_{f_n > R} \to 0 \quad \text{in} \quad L^1_{t,x}(\mathcal{M}^*_v),$$

uniformly in $n \in \mathbb{N}$.

Proof. Let $\varphi \in C_0(\mathbb{R}^d_v)$ be a non-negative function. Fix an $\omega \in \Omega$ and mostly omit dependence throughout the proof. The bound (7.20) yields the following inequality on $\Omega \times [0,T] \times \mathbb{R}^{2d}$: for all $K > 1$,

$$\mathcal{B}^+_n(f_n, f_n) \leq (\log K)^{-1} \mathcal{D}^0_n(f_n) + K \mathcal{B}^-_n(f_n, f_n).$$

Hence, for almost every $(\omega, t, x) \in \Omega \times [0,T] \times \mathbb{R}^d$, we find that

$$\begin{aligned}
&\left\langle \frac{\mathcal{B}^+_n(f_n, f_n)}{1+\langle f_n, 1\rangle} \mathbb{1}_{f_n > R}, \varphi \right\rangle \\
&\leq (\log K)^{-1} \|\varphi\|_{L^\infty_v} \mathcal{D}_n(f_n) + K \left\langle \frac{\mathcal{B}^-_n(f_n, f_n)}{1+\langle f_n, 1\rangle} \mathbb{1}_{f_n > R}, \varphi \right\rangle.
\end{aligned}$$

Next we observe that pointwise in Ω,

$$\left\| \left\langle \frac{\mathcal{B}^-_n(f_n, f_n)}{1+\langle f_n, 1\rangle} \mathbb{1}_{f_n > R}, \varphi \right\rangle \right\|_{L^1_{t,x}} \leq \|\bar{b}_n\|_{L^\infty_v} \|\varphi\|_{L^\infty_v} \|f_n \mathbb{1}_{f_n > R}\|_{L^1_{t,x,v}}.$$

By Proposition 7.6.1, $\{f_n(\omega)\}_{n\in\mathbb{N}}$ is uniformly integrable in $L^1_{t,x,v}$ and $\{\bar{b}_n\}_{n\in\mathbb{N}}$ is uniformly bounded in L^∞_v, passing $R\to\infty$ yields

$$\limsup_{R\to\infty}\sup_{n\in\mathbb{N}}\left\|\left\langle\frac{\mathcal{B}^+_n(f_n,f_n)}{1+\langle f_n,1\rangle}\mathbb{1}_{f_n>R},\varphi\right\rangle\right\|_{L^1_{t,x}}\leq(\log K)^{-1}\|\varphi\|_{L^\infty_v}\sup_{n\in\mathbb{N}}\|\mathcal{D}_n(f_n)\|_{L^1_{t,x}},\tag{7.69}$$

pointwise in Ω. By Proposition 7.6.1, there exists a constant $C(\omega)$ such that

$$\sup_{n\in\mathbb{N}}\|\mathcal{D}_n(f_n)(\omega)\|_{L^1_{t,x}}\leq C(\omega).$$

Sending $K\to\infty$ on both sides of (7.69) we find

$$\lim_{R\to\infty}\sup_n\left\|\left\langle\frac{\mathcal{B}^+_n(f_n,f_n)}{1+\langle f_n,1\rangle}\mathbb{1}_{f_n>R},\varphi\right\rangle\right\|_{L^1_{t,x}}\to 0.$$

Since we can always split any $\varphi\in C_0(\mathbb{R}^d_v)$ into positive and negative parts also in $C_0(\mathbb{R}^d_v)$, the above convergence holds for any $\varphi\in C_0(\mathbb{R}^d_v)$, completing the proof. $\square$

The next step is to apply Lemma 7.7.4 to obtain another Lemma written below.

Lemma 7.7.5. *As $m\to\infty$, the following limit holds $\mathbf{P}$ almost surely:*

$$\frac{\mathcal{B}^+_m}{1+\langle f,1\rangle}\to\frac{\mathcal{B}^+(f,f)}{1+\langle f,1\rangle}\quad in\quad L^1_{t,x}(\mathcal{M}^*_v).$$

Proof. Let $\varphi\in C_0(\mathbb{R}^d_v)$ be non-negative. Fix $\omega\in\Omega$ throughout and mostly omit. The first step is to observe that for each fixed $m\in\mathbb{N}$, pointwise in Ω,

$$\begin{aligned}&\left\|\left\langle\frac{\mathcal{B}^+_m-\mathcal{B}^+(f,f)}{1+\langle f,1\rangle},\varphi\right\rangle\right\|_{L^1_{t,x}}\\&\leq\liminf_{n\to\infty}\left\|\left\langle\frac{\Gamma'_m(f_n)\mathcal{B}^+_n(f_n,f_n)-\mathcal{B}^+_n(f_n,f_n)}{1+\langle f_n,1\rangle},\varphi\right\rangle\right\|_{L^1_{t,x}}.\end{aligned}\tag{7.70}$$

Indeed, this follows from the following two observations. In view of Lemma 7.7.4,

$$\left\langle \frac{\mathcal{B}_n^+(f_n, f_n)}{1+\langle f_n, 1\rangle}, \varphi \right\rangle \to \left\langle \frac{\mathcal{B}^+(f, f)}{1+\langle f, 1\rangle}, \varphi \right\rangle \quad \text{strongly in} \quad L^1_{t,x},$$

pointwise in Ω. By Proposition 7.6.1, $\{\Gamma'_m(f_n)\mathcal{B}_n^+(f_n, f_n)(\omega)\}_{n\in\mathbb{N}}$ converges to $\mathcal{B}_m^+(\omega)$ weakly in $L^1_{t,x,v}$ and $\{f_n(\omega)\}_{n\in\mathbb{N}}$ converges to $f(\omega)$ in $L^1_{t,x}(\mathcal{M}_v^*)$. Therefore using the uniform bounds on $\{f_n(\omega)\}$ we conclude that $\langle f_n(\omega), 1\rangle$ converges to $\langle f(\omega), 1\rangle$ in measure on $[0, T]\times\mathbb{R}^{2d}$. Therefore, the product Lemma B.1.12 yields the $\mathbf{P}$ almost sure convergence

$$\frac{\Gamma'_m(f_n)\,\mathcal{B}_n^+(f_n, f_n)}{1+\langle f_n, 1\rangle} \to \frac{\mathcal{B}_m^+}{1+\langle f, 1\rangle} \quad \text{weakly in} \quad L^1_{t,x,v}.$$

Now the desired inequality follows from the lower semi-continuity of the $L^1_{t,x,v}$ norm with respect to weak convergence.

The next step is to observe that for all $R > 1$,

$$\begin{aligned} &\left\| \left\langle \frac{\Gamma'_m(f_n)\mathcal{B}_n^+(f_n, f_n) - \mathcal{B}_n^+(f_n, f_n)}{1+\langle f_n, 1\rangle}, \varphi \right\rangle \right\|_{L^1_{t,x}} \\ &\qquad \le \left[1 - \left(1+\frac{R}{m}\right)^{-2}\right] \left\| \left\langle \frac{\mathcal{B}_n^+(f_n, f_n)}{1+\langle f_n, 1\rangle}, \varphi \right\rangle \right\|_{L^1_{t,x}} + \left\| \left\langle \frac{\mathcal{B}_n^+(f_n, f_n)}{1+\langle f_n, 1\rangle} 1_{f_n > R}, \varphi \right\rangle \right\|_{L^1_{t,x}}. \end{aligned} \tag{7.71}$$

Indeed, writing $1 = 1_{f_n<R} + 1_{f_n\ge R}$ and recalling that $\Gamma'_m(x) = (1+\frac{x}{m})^{-2}$, we find the following upper and lower bounds hold pointwise in $\Omega\times[0, T]\times\mathbb{R}^{2d}$

$$\begin{aligned} \frac{\mathcal{B}_n^+(f_n, f_n)}{(1+\frac{f_n}{m})^2} &\le \mathcal{B}_n^+(f_n, f_n). \\ \frac{\mathcal{B}_n^+(f_n, f_n)}{(1+\frac{f_n}{m})^2} &\ge \frac{\mathcal{B}_n^+(f_n, f_n)}{(1+\frac{R}{m})^2} - \mathcal{B}_n^+(f_n, f_n)\,1_{f_n\ge R}. \end{aligned}$$

Subtracting $\mathcal{B}_n^+(f_n, f_n)$ on both sides, pairing with φ, dividing by $1+\langle f_n, 1\rangle$, and integrating over $[0, T]\times\mathbb{R}^d$ gives the claim.

Using (7.70), we may pass $n \to \infty$ on both side of (7.71), pointwise in Ω. Appealing to Lemma 7.7.3 to pass the limit in the first term on the right-hand side of (7.71), we find that for each $m \in \mathbb{N}$ and $R > 1$, the following inequality holds pointwise in Ω

$$\begin{aligned}
&\left\| \left\langle \frac{\mathcal{B}_m^+ - \mathcal{B}^+(f,f)}{1+\langle f,1\rangle}, \varphi \right\rangle \right\|_{L^1_{t,x}} \\
&\le \left[1 - \left(1 + \frac{R}{m}\right)^{-2}\right] \left\| \left\langle \frac{\mathcal{B}^+(f,f)}{1+\langle f,1\rangle}, \varphi \right\rangle \right\|_{L^1_{t,x}} + \sup_{n\in\mathbb{N}} \left\| \left\langle \frac{\mathcal{B}_n^+(f_n,f_n)}{1+\langle f_n,1\rangle} 1_{f_n > R}, \varphi \right\rangle \right\|_{L^1_{t,x}}.
\end{aligned}$$

Passing $m \to \infty$ yields for each $R > 1$, pointwise in Ω

$$\limsup_{m\to\infty} \left\| \left\langle \frac{\mathcal{B}_m^+ - \mathcal{B}^+(f,f)}{1+\langle f,1\rangle}, \varphi \right\rangle \right\|_{L^1_{t,x}} \le \sup_{n\in\mathbb{N}} \left\| \left\langle \frac{\mathcal{B}_n^+(f_n,f_n)}{1+\langle f_n,1\rangle} 1_{f_n > R}, \varphi \right\rangle \right\|_{L^1_{t,x}}.$$

Finally, sending $R \to \infty$ and applying Lemma 7.7.4 to remove the peaks completes the proof. □

Proof of Proposition 7.7.2

Finally, we can apply our lemmas in order to obtain our main Proposition.

Proof of Proposition 7.7.2. Let us begin with the analysis of the negative part $\mathcal{B}_m^-$. The first point is to observe that for all $\omega \in \Omega$, we may identify $\mathcal{B}_m^-(\omega) = \overline{\gamma_m(f)}\, \overline{b} *_v f(\omega)$. Indeed, recall that $\{\Gamma'_m(f_n)\mathcal{B}_n^-(f_n,f_n)(\omega)\}_{n\in\mathbb{N}}$ converges to $\mathcal{B}_m^-(\omega)$ weakly in $L^1_{t,x,v}$ by Proposition 7.6.1. On one hand, since $\{\overline{b}_n *_v f_n(\omega)\}_{n\in\mathbb{N}}$ is uniformly integrable in $L^1_{t,x,v}$ and converges in measure on $[0,T] \times \mathbb{R}^{2d}$ to $\overline{b} *_v f(\omega)$, then by Vitali convergence

$$\{\overline{b}_n *_v f_n(\omega)\}_{n\in\mathbb{N}} \to \overline{b} *_v f(\omega) \quad \text{in} \quad L^1_{t,x,v}.$$

On the other hand, $\{f_n\Gamma'_m(f_n)(\omega)\}_{n\in\mathbb{N}}$ converges weakly to $\overline{\gamma_m(f)}(\omega)$ in $L^1_{t,x,v}$, and is uniformly (in n) bounded in $L^\infty_{t,x,v}$, then (up to a subsequence) $\{f_n\Gamma'_m(f_n)(\omega)\}_{n\in\mathbb{N}}$ converges to $\overline{\gamma_m(f)}$ in $[L^\infty_{t,x,v}]^*$. Therefore (up to a subsequence), since this is a weak-* L^∞ - strongly L^1 product limit, we obtain

$$\Gamma'_m(f_n)\mathcal{B}^-_n(f_n,f_n)(\omega) \to \overline{\gamma_m(f)}\,\bar{b} *_v f(\omega) \quad \text{in} \quad [L^1_{t,x,v}]_w. \tag{7.72}$$

However, since $\{\Gamma'_m(f_n)\mathcal{B}^-_n(f_n,f_n)(\omega)\}_{n\in\mathbb{N}}$ converges to $\mathcal{B}^-_m(\omega)$ in $[L^1_{t,x,v}]_{\mathrm{w}}$ the above convergence holds for the whole sequence and the claimed identification holds.

Next, by Corollary 7.6.10, $\overline{\Gamma_m(f)}(\omega) \to f(\omega)$ in $L^1_{t,x,v}$, and by an analogous argument one can show $\overline{\gamma_m(f)}(\omega) \to f(\omega)$ in $L^1_{t,x,v}$. This allows us to deduce that $\mathbf{P}$ almost surely,

$$\left\{\frac{\mathcal{B}^-_m}{1+\overline{\Gamma_m(f)}}\right\}_{m\in\mathbb{N}} \to \frac{B^-(f,f)}{1+f} \quad \text{in measure on} \quad [0,T]\times\mathbb{R}^{2d}.$$

Since $\gamma_m(z) = z\Gamma'_m(z) = (1+\frac{z}{m})^{-1}\Gamma_m(z)$, then $\overline{\gamma_m(f)} \le \overline{\Gamma_m(f)}$ pointwise for each $m\in\mathbb{N}$. This yields the pointwise inequality

$$\frac{\mathcal{B}^-_m}{1+\overline{\Gamma_m(f)}} \le \bar{b} *_v f. \tag{7.73}$$

A double application of Lebesgue dominated convergence (first in $[0,T]\times\mathbb{R}^{2d}$ and then in Ω) using the bound above and the fact that $f\in L^2(\Omega;L^1_{t,x,v})$ allows us to complete the first part of the proof (in fact it gives strong convergence in $L^2(\Omega;L^1_{t,x,v})$).

To treat the positive part of the renormalized collision operator, observe that for each $m,n\in\mathbb{N}$, the bound (7.20) gives the pointwise bound

$$\frac{\Gamma'_m(f_n)\mathcal{B}^+_n(f_n,f_n)}{1+\overline{\Gamma_m(f)}} \le \frac{1}{\log K}\mathcal{D}^0_n(f_n) + K\frac{\Gamma'_m(f_n)B^-_n(f_n,f_n)}{1+\overline{\Gamma_m(f)}}.$$

Next we pair with a positive $\phi \in C_0([0,T]\times\mathbb{R}^{2d})$ and pass the $n\to\infty$ limit on both sides of the inequality above and use the convergence of $\mathcal{D}^0_n(f_n)$ to $\overline{\mathcal{D}^0(f)}$ in $\mathcal{M}^*_{t,x,v}$ given in Proposition 7.6.1 and the inequality (7.73) to obtain

$$\left\langle \frac{\mathcal{B}^+_m}{1+\overline{\Gamma_m(f)}},\phi\right\rangle \le \frac{1}{\log K}\left\langle \overline{\mathcal{D}^0(f)},\phi\right\rangle + K\left\langle \bar{b}*_v f,\phi\right\rangle.$$

For the second term on the right-hand side above we used the convergence (7.72) and the poinwise bound $\overline{\gamma_m(f)} \le \overline{\Gamma_m(f)}$. Furthermore, using the fact that $(1+\overline{\Gamma_m(f)})^{-1}\mathcal{B}^+_m$ is in $L^1_{t,x,v}$ and taking ϕ to be a suitable approximation of the identity allows us to conclude the almost everywhere $\Omega\times[0,T]\times\mathbb{R}^{2d}$ inequality

$$\frac{\mathcal{B}^+_m}{1+\overline{\Gamma_m(f)}} \lesssim \overline{\mathcal{D}^0(f)}_{\mathrm{ac}} + \bar{b}*f, \tag{7.74}$$

where $\overline{\mathcal{D}^0(f)}_{\mathrm{ac}}$ is the density of the absolutely continuous part of $\overline{\mathcal{D}^0(f)}$.

To finish the proof, we write

$$\frac{\mathcal{B}^+_m}{1+\overline{\Gamma_m(f)}} = \frac{1+\langle f,1\rangle}{1+\overline{\Gamma_m(f)}}\frac{\mathcal{B}^+_m}{1+\langle f,1\rangle}.$$

By Corollary 7.6.10,

$$\left\{\frac{1}{1+\overline{\Gamma_m(f)}}\right\}_{m\in\mathbb{N}} \to \frac{1}{1+f} \quad \text{in measure on} \quad [0,T]\times\mathbb{R}^{2d},$$

and by Lemma 7.7.5

$$\left\{\frac{\mathcal{B}^+_m}{1+\langle f,1\rangle}\right\}_{m\in\mathbb{N}} \to \frac{\mathcal{B}^+(f,f)}{1+\langle f,1\rangle} \quad \text{in} \quad L^1_{t,x}(\mathcal{M}^*_v).$$

The product limit Lemma B.1.12, gives $\mathbf{P}$ almost surely

$$\left\{\frac{\mathcal{B}^+_m}{(1+\langle f,1\rangle)(1+\overline{\Gamma_m(f)})}\right\}_{m\in\mathbb{N}} \to \frac{B^+(f,f)}{(1+\langle f,1\rangle)(1+f)} \quad \text{in} \quad L^1_{t,x}(\mathcal{M}^*_v),$$

and therefore we can conclude (using the fact that $\langle f, 1\rangle$ is independent of v), for each $\varphi \in C_0(\mathbb{R}^d_v)$,

$$\left\{\left\langle \frac{\mathcal{B}^+_m}{1+\overline{\Gamma_m(f)}}, \varphi\right\rangle\right\}_{m\in\mathbb{N}} \to \left\langle \frac{B^+(f,f)}{1+f}, \varphi\right\rangle \quad \text{in measure on} \quad [0,T]\times\mathbb{R}^d_x.$$

In view of the bound (7.74) we would like to again use a double application of the dominated convergence theorem (first in $[0,T]\times\mathbb{R}^d_x$ and then in ω) to complete the proof. Indeed in order to apply dominated convergence in Ω it suffices to show that $\overline{\mathcal{D}^0(f)}_{ac} \in L^2(\Omega; L^1_{t,x,v})$. To show this, choose $\phi \in C_0([0,T]\times\mathbb{R}^{2d})$ non-negative. By the $\mathbf{P}$ almost sure convergence of $\mathcal{D}^0_n(f_n)$ in Proposition $\mathcal{M}^*_{t,x,v}$, $\{|\langle\mathcal{D}^0_n(f_n), \phi\rangle|^2\}_{n\in\mathbb{N}}$ converges to $|\langle\overline{\mathcal{D}^0(f)}, \phi\rangle|^2$ $\mathbf{P}$ almost surely. It follows by Fatou's Lemma (in Ω) that

$$\mathbf{E}|\langle\overline{\mathcal{D}^0(f)}_{\text{as}}, \phi\rangle|^2 \le \mathbf{E}|\langle\overline{\mathcal{D}^0(f)}, \phi\rangle|^2 \le \sup_n \mathbf{E}|\langle\mathcal{D}^0_n(f_n), \phi\rangle|^2 \le \|\phi\|^2_{L^\infty_{t,x,v}} \sup_n \mathbf{E}\|\mathcal{D}_n(f_n)\|^2_{L^1_{t,x}}.$$

Since $\overline{\mathcal{D}^0(f)}_{\text{as}} \ge 0$, we may replace ϕ by a sequence of non-negative functions $\{\phi_k\}_{k\in\mathbb{N}} \subseteq C_0(\mathbb{R}^d)$, $\phi_k \to 1$ pointwise and monotonically. Then, passing $k\to\infty$ using monotone convergence and using the uniform bounds on $\mathcal{D}_n(f_n)$ yields the result. □

Proof of Main Result

Proof of Theorem 7.1.7. We begin by proving estimates (7.12). Recall that Proposition 7.6.1 implies that $\{f_n\}_{n\in\mathbb{N}}$ converges to f in $C_t([L^1_{x,v}]_w)$ with probability one. We begin by showing the bound on $(1+|x|^2+|v|^2)f$. Let B_R denote the ball of radius $R > 0$ in $\mathbb{R}^{2d}$. It follows that $\mathbf{P}$ almost surely,

$$\|(1+|x|^2+|v|^2)1_{B_R} f_n\|^p_{L^\infty_t(L^1_{x,v})} \to \|(1+|x|^2+|v|^2)1_{B_R} f\|^p_{L^\infty_t(L^1_{x,v})}.$$

By Fatou's Lemma in Ω, we find that

$$\mathbf{E}\|(1+|x|^2+|v|^2)1_{B_R} f\|^p_{L^\infty_t(L^1_{x,v})} \leq \sup_{n\in\mathbb{N}} \mathbf{E}\|(1+|x|^2+|v|^2)1_{B_R} f_n\|^p_{L^\infty_t(L^1_{x,v})} < \infty,$$

in view of Part 2 of Proposition 7.6.1. Sending $R \to \infty$ and applying Fatou's Lemma once more yields

$$\mathbf{E}\|(1+|x|^2+|v|^2) f\|^p_{L^\infty_t(L^1_{x,v})} < \infty.$$

To show the bounds on $f|\log f|$ and $\mathcal{D}(f)$, we recall the proof of Lemma 7.5.7, where we showed that $\{f_n\}_{n\in\mathbb{N}}$ satisfies the following entropy equation $\mathbf{P}$- almost surely for each $t \in [0,T]$,

$$\iint_{\mathbb{R}^{2d}} f_n(t)\log f_n(t)\mathrm{d}x\mathrm{d}v + \int_0^t \int_{\mathbb{R}^d} \mathcal{D}_n(f_n(s))\mathrm{d}x\mathrm{d}s = \iint_{\mathbb{R}^{2d}} f_0 \log f_0 \mathrm{d}x\mathrm{d}v. \quad (7.75)$$

Since $z \mapsto z\log z$ is convex, and $\{f_n\}_{n\in\mathbb{N}} \to f$ in $C_t([L^1_{x,v}]_w)$ $\mathbf{P}$ almost surely, then, by lower semi-continuity and the non-negativity of $\mathcal{D}_n(f_n)$, the following inequality holds pointwise in $\Omega \times [0,T]$,

$$\iint_{\mathbb{R}^{2d}} f\log f \mathrm{d}x\mathrm{d}v \leq \iint_{\mathbb{R}^{2d}} f_0 \log f_0 \mathrm{d}x\mathrm{d}v.$$

From this point on, we may follow the arguments in Section 7.2.4 to conclude

$$\mathbf{E}\|f\log f\|^p_{L^\infty_t(L^1_{x,v})} < \infty.$$

To show the bound on the dissipation $\mathcal{D}(f)$, we remark that a standard modification of the proof of Lemma 7.7.3 allows us to conclude the $\mathbf{P}$ almost surely

$$\frac{f_n' f_{n,*}'}{1+\epsilon\langle f_n, 1\rangle} \to \frac{f' f_*'}{1+\epsilon\langle f, 1\rangle} \quad \text{in} \quad [L^1([0,T]\times\mathbb{R}^{3d}_{x,v,v*}\times\mathbb{S}^{d-1})]_w,$$

for each $\epsilon > 0$. Similarly, by the product limit Lemma B.1.12, we may also conclude that $\mathbf{P}$ almost surely,

$$\frac{f_n f_{n,*}}{1+\epsilon\langle f_n, 1\rangle} \to \frac{f f_*}{1+\epsilon\langle f, 1\rangle} \quad \text{in} \quad [L^1([0,T]\times\mathbb{R}^{3d}_{x,v,v_*})]_w.$$

Notice that the function $(x,y)\mapsto(x-y)(\log x - \log y)$ is convex on $\mathbb{R}^2_+$. Therefore, by lower semi-continuity we may conclude that $\mathbf{P}$ almost surely, for every $t\in[0,T]$ and each $\epsilon > 0$

$$\begin{aligned}\int_0^t \iiint_{\mathbb{R}^{3d}\times\mathbb{S}^{d-1}} &\frac{d(f)b}{1+\epsilon\langle f,1\rangle}\mathrm{d}\theta\mathrm{d}v\mathrm{d}v_*\mathrm{d}x\mathrm{d}s \\ &\le \liminf_n \int_0^t \iiint_{\mathbb{R}^{3d}\times\mathbb{S}^{d-1}} \frac{d(f_n)b}{1+\epsilon\langle f_n,1\rangle}\mathrm{d}\theta\mathrm{d}v\mathrm{d}v_*\mathrm{d}x\mathrm{d}s \\ &\le \liminf_n \int_0^t \int_{\mathbb{R}^d} \mathcal{D}_n(f_n)\mathrm{d}x\mathrm{d}s.\end{aligned}$$

Taking $\epsilon\to 0$, by the monotone convergence theorem, gives

$$\int_0^t\int_{\mathbb{R}^d}\mathcal{D}(f)\mathrm{d}x\mathrm{d}s \le \liminf_n \int_0^t\int_{\mathbb{R}^d}\mathcal{D}_n(f_n)\mathrm{d}x\mathrm{d}s.$$

Passing $n\to\infty$ on both sides of (7.75) yields, the global entropy inequality (7.16),

$$\iint_{\mathbb{R}^{2d}} f(t)\log f(t)\mathrm{d}x\mathrm{d}v + \int_0^t\int_{\mathbb{R}^d}\mathcal{D}(f)(s)\mathrm{d}x\mathrm{d}s \le \int_{\mathbb{R}^{2d}} f_0\log f_0 \mathrm{d}x\mathrm{d}v.$$

Whereby we obtain the bound

$$\|\mathcal{D}(f)\|_{L^1_{t,x,v}} \le \|f\log f\|_{L^\infty_t(L^1_{x,v})} + \|f_0\log f_0\|_{L^1_{x,v}}.$$

Using the bound on $f\log f$ above, gives

$$\mathbf{E}\|\mathcal{D}(f)\|^p_{L^1_{t,x,v}} < \infty.$$

Next we show the conservation laws (7.13-7.16). In fact we have already shown (7.16) in the computation above. To show (7.13-7.15), recall that f_n satisfies for each

$\varphi \in C_c^\infty(\mathbb{R}^{2d})$

$$\begin{aligned}\langle f_n, \varphi\rangle = \langle f_n^0, \varphi\rangle + \int_0^t \langle f_n(s), v\cdot\nabla_x\varphi + \mathcal{L}_{\sigma^n}\varphi\rangle \mathrm{d}s \\ + \int_0^t \langle f_n(s), \sigma_k^n \cdot \nabla_v \varphi\rangle \mathrm{d}\beta_k^n(s) + \int_0^t \langle \mathcal{B}_n(f_n, f_n), \varphi\rangle \mathrm{d}s\end{aligned} \tag{7.76}$$

in distribution in x, v. Using the **P** almost sure moment estimates provided by property 2 in Proposition 7.6.1 and the boundedness of the truncated collision operator $\mathcal{B}_n(f_n, f_n)$

$$\|(1+|x|^2+|v|^2)^k f_n\|_{L_t^\infty(L^1_{x,v})} < \infty, \quad \|(1+|x|^2+|v|^2)^k \mathcal{B}_n(f_n, f_n)\|_{L^1_{t,x,v}} < \infty.$$

It is straight forward to use these estimates to upgrade to a class of test functions $\varphi(x, v)$ with sub quadratic growth

$$\sup_{x,v} \frac{|\varphi(x,v)|}{(1+|x|^2+|v|^2)} < \infty,$$

in equation (7.76). Choosing the test function to be constant in v gives **P** almost surely for each $t \in [0, T]$

$$\int_{\mathbb{R}^d} f_n(t)\mathrm{d}v + \mathrm{div}_x \int_0^t \int_{\mathbb{R}^d} v f_n(s)\mathrm{d}v\mathrm{d}s = \int_{\mathbb{R}^d} f_0 \mathrm{d}v \quad \text{in} \quad \mathcal{D}'_x. \tag{7.77}$$

Likewise taking the test function to approach $\phi(x, v) = v$, and taking expectation, we can obtain for each $t \in [0, T]$

$$\mathbf{E}\iint_{\mathbb{R}^{2d}} v f_n(t)\mathrm{d}v\mathrm{d}x = \mathbf{E}\int_0^t \iint_{\mathbb{R}^{2d}} (\mathcal{L}_{\sigma^n} v) f_n(s)\mathrm{d}v\mathrm{d}x\mathrm{d}s + \int_{\mathbb{R}^{2d}} v f_0^n \mathrm{d}v\mathrm{d}x. \tag{7.78}$$

Finally taking a test function approaching $\varphi(x, v) = \frac{1}{2}|v|^2$, and taking expectation gives for each $t \in [0, T]$

$$\mathbf{E}\iint_{\mathbb{R}^{2d}} \frac{1}{2}|v|^2 f_n(t)\mathrm{d}v\mathrm{d}x = \mathbf{E}\int_0^t \iint_{\mathbb{R}^{2d}} \frac{1}{2}(\mathcal{L}_{\sigma^n}|v|^2) f_n(s)\mathrm{d}v\mathrm{d}x\mathrm{d}s + \int_{\mathbb{R}^{2d}} \frac{1}{2}|v|^2 f_0^n \mathrm{d}v\mathrm{d}x. \tag{7.79}$$

In order the pass the limit in n above, will find it useful to prove the following extension of the product limit Lemma B.1.12 for the sequence $\{f_n\}_{n\in\mathbb{N}}$.

Lemma 7.8.1. *Let $\{\phi_n\}_{n\in\mathbb{N}}$ be a sequence of functions in $[L^\infty_{x,v}]_{\mathrm{loc}}$ converging pointwise a.e to ϕ satisfying the uniform growth assumption*

$$\lim_{R\to\infty}\sup_n \left\| \frac{\phi_n(x,v)}{1+|x|^2+|v|^2}\mathbb{1}_{B_R^c}\right\|_{L^\infty_{x,v}} = 0 \tag{7.80}$$

where $B_R \subset \mathbb{R}^{2d}$ is the ball of radius R. Then,

$$\iint_{\mathbb{R}^{2d}} \phi_n f_n \mathrm{d}v\mathrm{d}x \to \iint_{\mathbb{R}^{2d}} \phi f \mathrm{d}v\mathrm{d}x \quad in \quad [L^2(\Omega\times[0,T])]_{\mathrm{w}}.$$

Proof. Proposition 7.6.1 implies that $\mathbf{P}$ almost surely $\{f_n\}_{n\in\mathbb{N}} \to f$ in $C_t([L^1_{x.v}]_{\mathrm{w}})$. Since $\phi_n\mathbb{1}_{B_R}$ is uniformly bounded in $L^\infty_{x,v}$ and converges in pointwise a.e. to $\phi\mathbb{1}_{B_R}$ the product limit Lemma B.1.12 implies that $\mathbf{P}$ almost surely for each $t\in[0,T]$

$$\int_{\mathbb{R}^{2d}} \phi_n \mathbb{1}_{B_R} f_n(t)\mathrm{d}v\mathrm{d}x \to \int_{\mathbb{R}^{2d}} \phi\mathbb{1}_{B_R} f(t)\mathrm{d}v\mathrm{d}x \tag{7.81}$$

Now, letting $C<\infty$ denote the (random) constant such that

$$\sup_n \|(1+|x|^2+|v|^2)(|f_n|+|f|)\|_{L^\infty_t(L^1_{x,v})} < C,$$

we have

$$\begin{aligned}\int_{\mathbb{R}^{2d}} \phi_n(f-f_n)\mathrm{d}v\mathrm{d}x &\leq \int_{\mathbb{R}^{2d}} \phi_n\mathbb{1}_{B_R}(f-f_n)\mathrm{d}v\mathrm{d}x + \int_{\mathbb{R}^{2d}} \phi_n\mathbb{1}_{B_R^c}(f-f_n)\mathrm{d}v\mathrm{d}x \\ &\leq \int_{\mathbb{R}^{2d}} \phi_n\mathbb{1}_{B_R}(f-f_n)\mathrm{d}v\mathrm{d}x + C\sup_n\left\|\frac{\phi_n(x,v)}{1+|x|^2+|v|^2}\mathbb{1}_{B_R^c}\right\|_{L^\infty_{x,v}}.\end{aligned}$$

First, pass $n\to\infty$ above using (7.81) and then, send $R\to\infty$ above to conclude that for all ϕ satisfying (7.80), $\mathbf{P}$ almost surely, and for each $t\in[0,T]$,

$$\int_{\mathbb{R}^{2d}} \phi_n f_n \mathrm{d}v\mathrm{d}x \to \int_{\mathbb{R}^{2d}} \phi f\mathrm{d}v\mathrm{d}x.$$

Moreover by the average moment estimate on $\{f_n\}_{n\in\mathbb{N}}$,

$$\left\{\iint_{\mathbb{R}^{2d}} \phi_n f_n \mathrm{d}v\mathrm{d}x\right\}_{n\in\mathbb{N}} \quad \text{is bounded in} \quad L^2(\Omega\times[0,T]),$$

and therefore by Vitali convergence we may conclude that

$$\iint_{\mathbb{R}^{2d}} \phi_n f_n \mathrm{d}v\mathrm{d}x \to \iint_{\mathbb{R}^{2d}} \phi f \mathrm{d}v\mathrm{d}x \quad \text{in} \quad [L^2(\Omega\times[0,T])]_{\mathrm{w}}.$$

□

Immediately we can use this Lemma to pass the limit in each term of (7.77). Taking the derivative in time gives the local conservation law (7.13). Also using the fact that $\mathcal{L}_{\sigma^n} v = \sigma_k^n \cdot \nabla_v \sigma_k^n$ is bounded in $L^\infty_{x,v}$ and converges pointwise to $\mathcal{L}_\sigma v$, we may also pass the limit in each term of (7.78) to obtain (7.14).

Now, note that we cannot pass the limit directly in the energy equation (7.79) since $\frac{1}{2}|v|^2$ does not satisfy (7.80). However, $\mathcal{L}_{\sigma^n}|v|^2$ does satisfy (7.80), and so upon cutting of the domain on the left hand side of (7.79) can pass the limit in n and conclude for each $R > 0$,

$$\mathbf{E}\int_{\mathbb{R}^{2d}} \frac{1}{2}\mathbb{1}_{|v|<R}|v|^2 f(t) \leq \mathbf{E}\int_0^t \iint_{\mathbb{R}^{2d}} \frac{1}{2}(\mathcal{L}_\sigma |v|^2) f(s)\mathrm{d}v\mathrm{d}x\mathrm{d}s + \int_{\mathbb{R}^{2d}} \frac{1}{2}|v|^2 f_0 \mathrm{d}v\mathrm{d}x.$$

Apply the monotone convergence theorem to the left-hand side and sending $R \to \infty$ gives the desired inequality (7.15).

Next, we prove that f verifies the conditions of Definition 7.1.4. Begin by observing that for each $n \in \mathbb{N}$, $\tilde{f}_n$ has the property that for each $(t,\omega) \in [0,T]\times\Omega$, the quantity $\tilde{f}_n(t,\omega)$ is a non-negative element of $L^1_{x,v}$. Since f_n is given explicitly as $f_n = \tilde{f}_n \circ \tilde{T}_n$, it inherits this property. Finally, Proposition 7.6.1 implies that

$\{f_n(t,\omega)\}_{n\in\mathbb{N}}$ converges to $f(t,\omega)$ weakly in $L^1_{x,v}$. Since weak convergence is order preserving, this shows that f satisfies Part 1 of Definition 7.1.4. Also, by Corollary 7.6.10, $f : \Omega \times [0,T] \to L^1_{x,v}$ has continuous sample paths.

In view of Definition 7.3.1 and Remark 7.3.2, Parts 2 and 3 of Definition 7.1.4 will follow once we check that for each $\Gamma \in \mathcal{R}$, the process $\Gamma(f)$ is a weak martingale solution to the stochastic kinetic equation driven by $\Gamma'(f)\mathcal{B}(f,f)$, starting from $\Gamma(f^0)$. In fact, the problem can be reduced further.

Let us show that it suffices to verify $\log(1+f)$ is a weak martingale solution driven by $(1+f)^{-1}\mathcal{B}(f,f)$, starting from $\log(1+f^0)$. Assume for the moment this property of $\log(1+f)$ and let $\Gamma \in \mathcal{R}$ be arbitrary. Since we showed $f \in L^2(\Omega; L^\infty_t(L^1_{x,v}))$, it follows that $\log(1+f)$ belongs to $L^2(\Omega \times [0,T] \times \mathbb{R}^{2d})$. Hence, by Proposition 7.3.8, $\log(1+f)$ is a renormalized solution. We would like to renormalize by a β such that $\beta \circ \log(1+x) = \Gamma(x)$, or equivalently $\beta(x) = \Gamma(e^x - 1)$, but this is not quite admissible in the sense of Definition 7.3.6 since $\Gamma \in \mathcal{R}$ need not imply boundedness of β''. Instead, we proceed by a sequence of approximate renormalizations $\{\beta_k\}_{k\in\mathbb{N}}$ where $\beta_k(x) = \Gamma_k(e^x - 1)$ and $\{\Gamma_k\}_{k\in\mathbb{N}}$ have the following properties: for each $k \in \mathbb{N}$, Γ_k is compactly supported (and hence β_k'' is bounded), the pair $(\Gamma_k, \Gamma_k') \to (\Gamma, \Gamma')$ pointwise in $\mathbb{R}_+$, and the following uniform bound holds

$$\sup_{k\in\mathbb{N}} \sup_{x\in\mathbb{R}_+} (1+x)|\Gamma_k'(x)| < \infty.$$

By Proposition 7.3.8, it follows that $\Gamma_k(f)$ is a weak martingale solution driven by $\Gamma_k'(f)\mathcal{B}(f,f)$. Using the properties of $\{\Gamma_k\}_{k\in\mathbb{N}}$ and the fact that $f \in L^2(\Omega; L^\infty_t(L^1_{x,v}))$ and $(1+f)^{-1}\mathcal{B}(f,f) \in L^2(\Omega; L^1_{t,x,v})$, it is straight forward to use the stability result, Proposition 7.3.5, to pass $k \to \infty$ and conclude that $\Gamma(f)$ is a weak martingale solution driven by $\Gamma'(f)\mathcal{B}(f,f)$ starting from $\Gamma(f_0)$.

Thus, it remains to show that $\log(1+f)$ is a weak martingale solution to the stochastic kinetic equation driven by $(1+f)^{-1}\mathcal{B}(f,f)$, starting from $\log(1+f_0)$. For this, we use once more our stability result. Recall that for each $m \in \mathbb{N}$, the process $\log(1+\overline{\Gamma_m(f)})$ is a weak martingale solution to the stochastic kinetic equation driven by $(1+\overline{\Gamma_m(f)})^{-1}[\mathcal{B}_m^+ - \mathcal{B}_m^-]$, starting from $\log(1+\Gamma_m(f_0))$. First observe that that for all $\varphi \in C_c^\infty(\mathbb{R}^{2d})$, the sequence $\{\langle \log(1+\overline{\Gamma_m}), \varphi\rangle\}_{m\in\mathbb{N}}$ converges in $L^2(\Omega; C_t)$ towards $\langle \log(1+f), \varphi\rangle$. Indeed, this follows from Corollary 7.6.10, the almost everywhere inequality $\overline{\Gamma_m} \leq f$, and the estimates (7.12). Next, for each $t \in [0,T]$ we can use Proposition 7.7.2 with $\phi = \mathbb{1}_{[0,t]}\varphi$ to conclude that

$$\int_0^t \left\langle \frac{\mathcal{B}_m}{1+\overline{\Gamma_m}}, \varphi \right\rangle \mathrm{d}s \to \int_0^t \left\langle \frac{\mathcal{B}(f,f)}{1+f}, \varphi \right\rangle \mathrm{d}s \quad \text{in} \quad L^2(\Omega).$$

Using these facts together with the stability result Proposition 7.3.5, we may pass $m \to \infty$ and complete the proof. □

Large Deviations and Local Limit Theorems

Local Limit Theorems

Preliminaries

We begin by assuming that there is a probability measure μ on $\mathbb{R}^d$ with mean m and covariance matrix C given by

$$m = \int_{\mathbb{R}^d} x\,\mu(\mathrm{d}x), \quad C = \int_{\mathbb{R}^d} (x-m)^{\otimes 2}\mu(dx).$$

Let $\phi(\xi)$ be the characteristic function of $\mu(\mathrm{d}x)$

$$\phi(u) = \int_{\mathbb{R}^d} \mathrm{e}^{i\xi\cdot x}\mu(\mathrm{d}x),$$

we will assume the following conditions on $\phi(\xi)$:

Hypothesis A.1.1.

1. *For $|u| > 0$, $|\phi(\xi)| < 1$ (sometimes called non-lattice condition on μ).*

2. *There exists an $N \geq 1$ which is the smallest number such that $|\phi(\xi)|^N$ is an integrable function on $\mathbb{R}^d$.*

Local central limit theorem

We begin by proving the local central limit theorem. For this, we assume that we have a sequence of independent, mean zero random variables $\{X_i\}$ in $\mathbb{R}^d$ each with the same law $\mu(\mathrm{d}x)$. Define the partial sum

$$S_n = \sum_{i=1}^{n} X_i.$$

We would like to study the law $\mu_n(x)$ of S_n, defined by duality for any smooth bounded test function $\varphi(x)$ as

$$\int_{\mathbb{R}^d} \varphi(x)\mu_n(\mathrm{d}x) = \int_{(\mathbb{R}^d)^n} \varphi\left(\sum_{i=1}^{n} x_i\right) \prod_{i=1}^{n} \mu(\mathrm{d}x_i),$$

We note that the integrability condition in Hypothesis A.1.1 implies that $\mu_n(x)$ has a density $f_n(x)$ with respect to Lebesgue for large enough n. Our first step will be to prove the following theorem:

Theorem A.1.2. *Let $\mu(\mathrm{d}x)$ be a measure on $\mathbb{R}^d$ satisfying Hypothesis A.1.1. Then for $n > N$, $\mu_n(\mathrm{d}x)$ has a density $f_n(x)$ and the following limit holds uniformly in x*

$$\lim_{n\to\infty} (\sqrt{n})^d f_n(\sqrt{n}x) = \frac{\exp\left(-\frac{1}{2}(x, C^{-1}x)\right)}{\sqrt{(2\pi)^d \det C}}.$$

Proof. As is typical for the central limit theorem, the proof will study the characteristic function $\phi(\xi)$ of $\mu(\mathrm{d}x)$. The characteristic function of $f_n(x)$ is related to $\phi(\xi)$ by

$$\Phi_n(\xi) = \int_{\mathbb{R}^d} \mathrm{e}^{i\xi\cdot x} f_n(x)\mathrm{d}x = \left(\int_{\mathbb{R}^d} \mathrm{e}^{i\xi\cdot x}\mu(\mathrm{d}x)\right)^n = [\phi(\xi)]^n$$

Taking the inverse Fourier transform,

$$(\sqrt{n})^d f_n(\sqrt{n}x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \mathrm{e}^{-i\xi\cdot x}[\phi(\xi/\sqrt{n})]^n \mathrm{d}\xi,$$

and therefore we conclude that

$$\left|(\sqrt{n})^d f_n(\sqrt{n}x) - \frac{\exp\left(-\frac{1}{2}(x, C^{-1}x)\right)}{\sqrt{(2\pi)^d \det C}}\right|$$
$$\leq \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left|[\phi(\xi/\sqrt{n})]^n - \mathrm{e}^{-\frac{1}{2}(\xi, C\xi)}\right|\mathrm{d}\xi.$$

We will show the first part of the theorem if we can estimate the right hand-side above. By Taylor's theorem, since $\phi(\xi)$ is at least C^2 differentiable, for each $\xi \in \mathbb{R}^d$,

$$[\phi(\xi/\sqrt{n})]^n = \left[1 - \frac{1}{2n}(\xi, C\xi) + o(n^{-1})\right]^n \to \mathrm{e}^{-\frac{1}{2}(\xi, C\xi)}$$

as $n \to \infty$. Furthermore if ξ is sufficiently small, say $|\xi| < \delta$, Taylor's theorem also implies that

$$|\phi(\xi)| = \left|1 - \frac{1}{2}(\xi, C\xi) + \mathcal{O}(\delta^3)\right| \leq \mathrm{e}^{-\frac{1}{4}(\xi, C\xi)}$$

Therefore, if $|\xi| \leq \delta\sqrt{n}$, we have the bound

$$\left|[\phi(\xi/\sqrt{n})]^n - \mathrm{e}^{-\frac{1}{2}(\xi, C\xi)}\right| \leq \mathrm{e}^{-\frac{1}{4}(\xi, C\xi)}$$

It follows by dominated convergence that

$$\int_{|\xi|<\delta\sqrt{n}}\left|[\phi(\xi/\sqrt{n})]^n - \mathrm{e}^{-\frac{1}{2}(\xi, C\xi)}\right|\mathrm{d}\xi \to 0,$$

as $n \to \infty$. To estimate the region where $|\xi| > \delta\sqrt{n}$ we remark that since $|\phi(\xi)|^N$ is integrable and $\phi(\xi)$ is absolutely continuous, then $|\phi(\xi)| \to 0$ as $\xi \to \infty$ therefore the non-lattice condition $|\phi(\xi)| < 1$ implies that $\sup_{|\xi|>\delta}|\phi(\xi)| = \gamma < 1$. Therefore if $n > N$,

$$\int_{|\xi|>\delta\sqrt{n}}\left|[\phi(\xi/\sqrt{n})]^n - \mathrm{e}^{-\frac{1}{2}(\xi, C\xi)}\right| \leq \gamma^{n-N}\int_{\mathbb{R}^d}|\phi(\xi/\sqrt{\xi})|^N\mathrm{d}\xi + \int_{|\xi|>\delta\sqrt{n}}\mathrm{e}^{-\frac{1}{2}(\xi, C\xi)}\mathrm{d}\xi$$
$$= \gamma^{n-N}\sqrt{n}\int_{\mathbb{R}^d}|\phi(\xi)|^N\mathrm{d}\xi + \int_{|\xi|>\delta\sqrt{n}}\mathrm{e}^{-\frac{1}{2}(\xi, C\xi)}\mathrm{d}\xi$$

Sending $n \to \infty$ and using the integrability of $e^{-\frac{1}{2}(\xi, C\xi)}$ and the fact that $\gamma^n \sqrt{n} \to 0$, the above integral converges to 0. $\square$

We can prove a more quantitative version of the theorem above, which is essentially a local version of the Berry-Esseen theorem in the multidimensional setting. However, as we are proving this at the level of the densities, and require an integrability condition on the characteristic function, we are not able to obtain the typical Berry-Esseen estimate that only depends moments of the measure μ. Instead the bound depends on various quantities related to the decay and integrability of the characteristic function.

Theorem A.1.3 (Local Berry-Esseen Theorem)**.** *Assume that μ satisfies Hypothesis A.1.1 and assume the third moment $\int_{\mathbb{R}^d} |x|^3 \mu(\mathrm{d}x)$ is finite. Then there are universal constants $A > 0$ and $\delta > 0$ independent of μ so that*

$$\sup_{x \in \mathbb{R}^d} \left| (\sqrt{n})^d f_n(\sqrt{n} x) - \frac{\exp\left(-\frac{1}{2}(x, C^{-1}x)\right)}{\sqrt{(2\pi)^d \det C}} \right| \le \frac{A\rho}{\sqrt{n \det C}} + \frac{\gamma_\rho^n (\sqrt{n})^d \beta_N}{\sqrt{\det C}}$$

where $R = \sqrt{C}$ is the square root of C, $\rho = \int_{\mathbb{R}^d} |R^{-1}x|^3 \mu(\mathrm{d}x)$, $\gamma_\rho = \sup_{|\xi| > \delta/\rho} |\phi(\xi)| < 1$ and $\beta_N = \|\phi\|_{L^N}^N$, with N being the smallest number so that $|\phi(\xi)|^N$ is integrable (as per the Hypothesis A.1.1).

Proof. To begin, we assume that $C = \mathrm{Id}$, which can be obtained by changing coordinates to $y = R^{-1}x$ so that the measure μ has identity covariance and the third moment is $\rho = \int_{\mathbb{R}^d} |u|^3 \mu(\mathrm{d}x)$. Let $\epsilon = \rho/\sqrt{n}$. We being by noting that the third moment estimate gives

$$\phi(\xi) = 1 - \frac{1}{2}|\xi|^2 - i\frac{1}{6}|\xi|^3 + R(\xi)$$

where the remainder is given by

$$R(\xi) = -i\frac{1}{6}\int_0^1 \int_{\mathbb{R}^d} (\xi\cdot x)^3 \mathrm{e}^{i\xi\cdot x\lambda}\mu(\mathrm{d}x)\mathrm{d}\lambda, \quad |R(\xi)| \leq \frac{1}{6}|\xi|^3\rho$$

and therefore

$$n\log\phi(\xi/\sqrt{n}) = -\frac{1}{2}|\xi|^2 + \mathcal{O}(|\xi|^3\epsilon)$$

Using this, we conclude

$$\begin{aligned}\left|[\phi(\xi/\sqrt{n})]^n - \mathrm{e}^{-\frac{1}{2}|\xi|^2}\right| &\leq \mathrm{e}^{-\frac{1}{2}|\xi|^2}\left|\mathrm{e}^{|\xi|^3\mathcal{O}(\epsilon)} - 1\right| \\ &\leq \mathcal{O}(|\xi|^3\epsilon)\mathrm{e}^{-\frac{1}{2}|\xi|^2+\mathcal{O}(|\xi|^3\epsilon)}\end{aligned}$$

Choose a universal δ ($\delta = 1/24$ is sufficient) so that when $|\xi| < \delta/\epsilon$,

$$-\frac{1}{2}|\xi|^2 + \mathcal{O}(|\xi|^3\epsilon) \leq \frac{1}{4}|\xi|^2.$$

Therefore at frequencies less than δ/ϵ,

$$\int_{|\xi|<\delta/\epsilon}\left|[\phi(\xi/\sqrt{n})]^n - \mathrm{e}^{-\frac{1}{2}|\xi|^2}\right|\mathrm{d}\xi \leq \mathcal{O}(\epsilon)\int_{\mathbb{R}^d}|\xi|^3\mathrm{e}^{-\frac{1}{4}|\xi|^2}\mathrm{d}\xi = \mathcal{O}(\epsilon).$$

For the high frequencies, define $\gamma_\rho = \sup_{|\xi|>\delta/\rho}|\phi(\xi)|$ and $\beta_N = \|\phi\|_{L^N}^N$. Since $\phi(\xi)$ is uniformly continuous, $|\phi(\xi)|^N$ is integrable and $|\phi(\xi)| < 1$ for $\xi \neq 0$, we can conclude that $\gamma_\rho < 1$. Therefore

$$\begin{aligned}\int_{|\xi|>\delta/\epsilon}\left|[\phi(\xi/\sqrt{n})]^n - \mathrm{e}^{-\frac{1}{2}|\xi|^2}\right|\mathrm{d}\xi &\leq (\sqrt{n})^d\int_{|\xi|>\delta/\rho}|\phi(\xi)|^n\mathrm{d}\xi + \int_{|\xi|>\delta/\epsilon}\mathrm{e}^{-\frac{1}{2}|\xi|^2}\mathrm{d}\xi \\ &\leq \gamma_\rho^{n-N}\beta_N(\sqrt{n})^d + \frac{\epsilon}{\delta}\int_{\mathbb{R}^d}|\xi|\mathrm{e}^{-\frac{1}{2}|\xi|^2}\mathrm{d}\xi.\end{aligned}$$

We complete the proof by writing

$$\int_{\mathbb{R}^d}\left|[\phi(\xi/\sqrt{n})]^n - \mathrm{e}^{-\frac{1}{2}|\xi|^2}\right|\mathrm{d}\xi \leq \mathcal{O}(\epsilon) + \gamma_\rho^n(\sqrt{n})^d\beta_N,$$

and noting that we may change coordinates back to the original coordinates □

Local large-deviations on $\mathbb{R}^d$

We now study the large deviations of averages of sums of independent, identically distributed random variables. To begin, we will re-introduce the framework of Section A.2 and The logarithmic moment generating function of μ, $L : \mathbb{R}^d \to \mathbb{R}$ is defined by

$$L(\lambda) = \log(M(\lambda)), \quad \text{where} \quad M(\lambda) = \int_{\mathbb{R}^d} \mathrm{e}^{\lambda\cdot x}\mu(\mathrm{d}x),$$

and define its domain $D_L = \{\lambda \in \mathbb{R}^d \, : \, L(\lambda) < +\infty\}$. We will assume that 0 is contained the interior of D_L, which, of course, implies all moments of μ are finite.

Note that $L(\lambda) > -\infty$ for all λ, since by Jensen's inequality $L(\lambda) \geq \lambda \cdot m$. Indeed, as we are in the same setting as L enjoys some nice convexity and regularity properties summarized in the following Lemma.

Lemma A.1.4. *$L(\lambda)$ is strictly convex and C^∞ on the interior of it's domain. Moreover we have the following formulas for the gradient and the Hessian of $L(\lambda)$,*

$$\nabla L(\lambda) = \int_{\mathbb{R}^d} x\, \mathrm{e}^{\lambda\cdot x - L(\lambda)}\mu(\mathrm{d}x), \tag{A.1}$$

$$\nabla^2 L(\lambda) = \int_{\mathbb{R}^d} x \otimes x\, \mathrm{e}^{\lambda\cdot x - L(\lambda)}\mu(\mathrm{d}x) - \left(\int_{\mathbb{R}^d} x\, \mathrm{e}^{\lambda\cdot x - L(\lambda)}\mu(\mathrm{d}x)\right)^{\otimes 2}. \tag{A.2}$$

The *rate function* $I : D_I \subseteq \mathbb{R}^d \to \mathbb{R}$ associated to $L(\lambda)$ is define by the Legendre-Fenchel transform

$$I(u) = \sup_{\lambda\in D_G} \left(\lambda\cdot u - L(\lambda)\right), \tag{A.3}$$

where $D_I = \{x \in \mathbb{R}^d : I(x) < +\infty\}$.

Lemma A.1.5. *In addition to the consequences of Lemma A.2.2 the rate function I has the additional properties:*

1. *$I(u) \geq 0$ and $I(m) = 0$.*

2. *$\lim_{|u|\to\infty} I(u) = \infty$ and its sub-level sets $\{u \in \operatorname{Int} D_I : I(u) \leq \alpha\}$ are compact.*

Proof. The fact that $I(u) \geq 0$ follows from $L(0) = 0$, since

$$I(u) \geq -L(0) = 0.$$

Moreover, since $\lambda \cdot m - L(\lambda) \leq 0$, at $u = m$ we have the reverse inequality $I(m) \leq 0$, and therefore

$$I(m) = 0.$$

Now, fix a $u \in D_I$ and let r be such that $B_r(0) \subseteq D_L$. Then upon choosing $\lambda = ru/|u|$, we have

$$I(u) \geq r|u| - L(ru/|u|) \geq |u| - C_r, \quad C_r = \max_{\lambda \in \partial B_r(0)} L(\lambda) > 0. \tag{A.4}$$

Sending $|u| \to \infty$ on both sides above gives $\lim_{|u|\to\infty} I(u) = \infty$. Moreover since L is convex, it is continuous, and the sub-level set $\{I(u) \leq \alpha\}$ is closed. Also, by the inequality (A.4) is bounded since

$$\{I(u) \leq \alpha\} \subset \{r|u| \leq \alpha + C_r\} = B_R(0), \quad R = (\alpha + C_r)/r.$$

Therefore $\{I(u) \leq \alpha\}$ is compact. □

Lemma A.1.4 shows that derivatives of L are naturally given in terms of a probability measure $\mu_\lambda(\mathrm{d}x)$,

$$\mu_\lambda(\mathrm{d}x) := \frac{1}{M(\lambda)} \mathrm{e}^{\lambda \cdot x} \mu(\mathrm{d}x) = \mathrm{e}^{\lambda \cdot x - L(\lambda)} \mu(\mathrm{d}x),$$

defined for each $\lambda \in D_L$. $\mu_\lambda(\mathrm{d}x)$ is often called a *tilted measure*. Lemma A.1.4 immediately implies that $\mu_\lambda(\mathrm{d}x)$ has mean $m_\lambda = \nabla L(\lambda)$ and positive definite covariance matrix $C_\lambda = \nabla^2 L(\lambda)$. Therefore, for a given $u \in D_I$, $\mu_{\lambda_u}(\mathrm{d}x) = \mathrm{e}^{\lambda_u \cdot x - L(\lambda_u)}\mu(\mathrm{d}x)$ has mean u and covariance $[\nabla^2 I(m)]^{-1}$. Note that in this case, $\mu_{\lambda_m}(\mathrm{d}x)$ can be written as

$$\mu_{\lambda_m}(\mathrm{d}x) = \mathrm{e}^{\nabla I(m)\cdot(x-m)+I(m)}\mu(\mathrm{d}x).$$

Now we are ready to state the first theorem of this section. As in our discussion of the central limit theorem, we will let $S_n = X_1 + \ldots + X_n$ denote the sum of a family of n independent identically distributed random variables in $\mathbb{R}^d$ with common law $\mu(\mathrm{d}x)$. We would like to study the law $\mu_{\hat{S}_n}(\mathrm{d}x)$ of the sample mean $\hat{S}_n = S_n/n$. By the strong law of large numbers, we know $\hat{S}_n \to m$ almost surely. Therefore we expect $\mu_{\hat{S}_n}(\mathrm{d}x)$ to concentrate on a Dirac measure,

$$\mu_{\hat{S}_n} \to \delta_m, \quad \text{as} \quad n \to \infty,$$

in the tight topology of measures. However, in many applications, one would like more information on the approach of the distribution $\mu_{\hat{S}_n}$ to a Dirac. That is, often one is interested in gaining more information about the probability of deviations of $\hat{S}_n$ from it's limit m when n is large. As it turns out, for large n, if $u \neq m$, then the probability that $\hat{S}_n$ is near u decays exponentially fast with speed determined by the rate function $I(u)$. Roughly speaking,

$$\mathbf{P}(\hat{S}_n \text{ is near } u) \approx \mathrm{e}^{-nI(u)}.$$

Since the rate function $I(u) \geq 0$ and $I(m) = 0$, then when n is large the only event that whose probability doesn't decay exponentially fast is when $\hat{S}_n$ is near m, in accordance with the law of large numbers.

When $n > N$, the integrability condition on the characteristic function, given in Hypothesis A.1.1, implies that $\mu_{\hat{S}_n}(\mathrm{d}x)$ has a density $\hat{f}_n(x)$. The above discussion is made more concrete by the following local deviation theorem.

Theorem A.1.6. *Suppose that u belongs to the interior of D_I. Then*

$$\lim_{n\to\infty} \frac{1}{n} \log \hat{f}_n(u) = -I(u),$$

locally uniformly.

Proof. To prove this, we consider the tilted measure $\mu_{\lambda_u}(\mathrm{d}x)$ with mean $u \in \operatorname{Int} D_I$

$$\mu_{\lambda_u}(\mathrm{d}x) = \mathrm{e}^{\nabla I(u)\cdot(x-u)+I(u)}\mu(\mathrm{d}x).$$

Since the characteristic function $\phi(\xi)$ associated to μ is in $L^N(\mathbb{R}^d)$, then the characteristic function $\phi_{\lambda_u}(\xi)$ associated with $\mu_{\lambda_u}(\mathrm{d}x)$ is also in L^N. For $n > N$, let $\hat{f}_{u,n}(x)$ be the density of the law of the mean zero random variable

$$\hat{S}^u_n - u,$$

where $\hat{S}^u_n = (X^u_1 + X^u_2 + \dots X^u_n)/n$ and each X^u_i is distributed with respect to the tilted measure $\mu_{\lambda_u}(\mathrm{d}x)$. The density can be defined by duality for every continuous bounded function $\varphi(x)$,

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(x)\hat{f}_{u,n}(x)\mathrm{d}x &= \int_{(\mathbb{R}^d)^n} \varphi\Big(n^{-1}\sum_{i=1}^n x_i - u\Big)\prod_{i=1}^n \mu_{\lambda_u}(\mathrm{d}x_i)\\ &= \int_{(\mathbb{R}^d)^n} \varphi\Big(n^{-1}\sum_{i=1}^n x_i - u\Big)\mathrm{e}^{n\nabla I(u)\cdot(n^{-1}\sum_{i=1}^n x_i-u)+nI(u)}\prod_{i=1}^n \mu(\mathrm{d}x_i)\\ &= \int_{\mathbb{R}^d} \varphi(x-u)\mathrm{e}^{n\nabla I(u)\cdot(x-u)+nI(u)}\hat{f}_n(x)\mathrm{d}x.\end{aligned}$$

So that $\hat{f}_n$ and $\hat{f}_{u,n}$ are related by

$$\hat{f}_n(x+u) = \mathrm{e}^{-n\nabla I(u)\cdot x - nI(u)}\hat{f}_{u,n}(x).$$

Setting $x = 0$ in the above equation gives

$$\hat{f}_n(u) = \mathrm{e}^{-nI(u)}\hat{f}_{n,u}(0).$$

Since the change of measure from μ to μ_{λ_u} is done by an absolutely continuous transformation, if μ satisfies the non-lattice condition ($|\phi(\xi)| < 1$ for $|\xi| > 0$), then so does μ_{λ_u}. Therefore the local central limit Theorem A.1.2 applies to the random variable $\hat{S}_n^u - u$, yielding

$$\lim_{n\to\infty}(\sqrt{n})^{-d} f_{u,n}(0) = \sqrt{(2\pi)^{-d}\det\nabla^2 I(u)}.$$

This implies that $f_{u,n}(0) = \mathcal{O}(n^{d/2})$, and therefore

$$\frac{1}{n}\log\hat{f}_n(u) = -I(u) + \frac{1}{n}\log(f_{u,n}(0)) = -I(u) + \mathcal{O}(n^{-1}\log n).$$

Sending $n \to \infty$ completes the proof. □

In fact the previous theorem actually implies the following improved asymptotic, which is actually sharp in the case that the initial distribution is normal or gamma distributed. The following corollary is an easy consequence of the previous theorem the fact that all moments are finite and the local Berry-Esseen inequality proved in Theorem A.1.3.

Corollary A.1.7. *For each $u \in \mathbb{R}^n$, we have the asymptotic,*

$$\hat{f}_n(u) = \frac{\mathrm{e}^{-nI(u)}}{(2\pi)^{d/2}}\sqrt{n^d\det\nabla^2 I(u)}\left(1+\mathcal{O}(n^{-1/2})\right).$$

It is also likely that A.1.6 to C^2 convergence. Namely, the following convergences hold locally uniformly

$$\lim_{n\to\infty} \frac{1}{n} \log \hat{f}_n(u) = -I(u),$$

$$\lim_{n\to\infty} \frac{1}{n} \nabla \log \hat{f}_n(u) = -\nabla I(u),$$

$$\lim_{n\to\infty} \frac{1}{n} \nabla^2 \log \hat{f}_n(u) = -\nabla^2 I(u).$$

However, a proof of this result would rely on improved central limit theorem convergence results, as well as some uniform control on the statistical quantities related to the measure μ_{λ_u} in the parameter u. This is typically rather non-trivial and requires more assumptions on the measure μ.

General framework and abstract Gibbs ensembles

To begin, we will consider a general framework for a class of abstract Gibbs measures. Namely, those that can be written as a product of certain single particle Gibbs measures. The reason for considering such a general abstract approach to Gibbs measures, as opposed to presenting the following results for the more classical definitions of Gibbs measures, is due to the fact that we will not only be considering Gibbs measures corresponding to certain classical particles systems, but will also be considering more general Gibbs measures associated to certain fluid-particle systems. Also, we will find it necessary to change variables

In general, assume that we have $\gamma(\mathrm{d}x)$ a (potentially unbounded) positive, σ-finite Borel measure on a smooth d-dimensional manifold Γ, typically taken to be

$\mathbb{T}^d$, $\mathbb{R}^d$ or some product of the two. Furthermore, suppose we are given a measurable mapping,

$$h : \Gamma \to \mathbb{R}^r,$$

which we interpret as a generalized energy function on a one-particle phase space Γ. Associated with h we have a free energy function $F : D_F \subset \mathbb{R}^r \to \mathbb{R}$ by

$$F(\lambda) = \log Z(\lambda), \quad Z(\lambda) = \int_\Gamma \mathrm{e}^{\lambda \cdot h(x)} \gamma(\mathrm{d}x),$$

and suppose that it's domain $D_F = \{\lambda \in \mathbb{R}^r : |F(\lambda)| < \infty\}$ has non-empty interior. We will assume that for every $v \in \mathbb{R}^r$, $h \cdot v$ is non-constant on Γ, which will be sufficient to obtain strict convexity of F. Specifically, we have the following properties of free energy function $F(\lambda)$:

Lemma A.2.1. *F is strictly convex and C^∞ on $\operatorname{Int} D_F$ and D_F is convex. Moreover we have the following formulas for the gradient and the Hessian of $F(\lambda)$,*

$$\nabla F(\lambda) = \int_\Gamma h(x)\, \mathrm{e}^{\lambda \cdot h(x) - F(\lambda)} \mu(\mathrm{d}x),$$

$$\nabla^2 F(\lambda) = \int_\Gamma h(x)^{\otimes 2}\, \mathrm{e}^{\lambda \cdot h(x) - F(\lambda)} \mu(\mathrm{d}x) - \left(\int_\Gamma h(x)\, \mathrm{e}^{\lambda \cdot h(x) - F(\lambda)} \mu(\mathrm{d}x) \right)^{\otimes 2}.$$

Proof. To prove convexity, let $\alpha \in [0, 1]$ and $\lambda_1, \lambda_2 \in \operatorname{Int} D_F$, then Hölders inequality implies

$$Z(\alpha\lambda_1 + (1-\alpha)\lambda_2) = \int_\Gamma \left(\mathrm{e}^{\lambda_1 \cdot h(x)}\right)^\alpha \left(\mathrm{e}^{\lambda_2 \cdot h(x)}\right)^{1-\alpha} \gamma(\mathrm{d}x) \le Z(\lambda_1)^\alpha Z(\lambda_2)^{(1-\alpha)}.$$

Taking the logarithm of both sides implies convexity of $F(\lambda)$ as well as the convexity of D_F.

To prove C^∞, it suffices to show that $Z(\lambda)$ is C^∞ on $\operatorname{Int} D_F$. To see this, fix $\lambda \in \operatorname{Int} D_F$ and take $v \in \mathbb{R}^d$, $|v| = 1$ and choose ϵ_0 small enough so that $\lambda + \epsilon h \in \operatorname{Int} D_F$ for all $|\epsilon| \leq \epsilon_0$. Then the divided difference

$$D_{\epsilon v} \mathrm{e}^{\lambda \cdot h(x)} = (\mathrm{e}^{(\lambda+\epsilon v)\cdot h(x)} - \mathrm{e}^{\lambda \cdot h(x)})\epsilon^{-1}$$

converges pointwise to $v \cdot h(x)\mathrm{e}^{\lambda \cdot h(x)}$ as $\epsilon \to 0$, and has the bound

$$D_{\epsilon v} \mathrm{e}^{\lambda \cdot h(x)} \leq \mathrm{e}^{\lambda \cdot h(x)}(\mathrm{e}^{\epsilon_0 |v \cdot h(x)|} - 1)\epsilon_0^{-1} \leq \mathrm{e}^{\lambda \cdot h(x)}(\mathrm{e}^{\epsilon_0 v \cdot h(x)} + \mathrm{e}^{-\epsilon_0 v \cdot h(x)})\epsilon_0^{-1}.$$

This means that

$$\int_\Gamma D_{\epsilon v} \mathrm{e}^{\lambda \cdot h(x)} \gamma(\mathrm{d}x) \leq \epsilon_0^{-1} \left[Z(\lambda + \epsilon_0 v) + Z(\lambda - \epsilon_0 v) \right] < \infty.$$

Applying dominated convergence gives

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} Z(\lambda + \epsilon v)|_{\epsilon=0} = \int_{\mathbb{R}^d} v \cdot h(x) \mathrm{e}^{\lambda \cdot h(x)} \gamma(\mathrm{d}x),$$

and therefore

$$\nabla Z(\lambda) = \int_\Gamma h(x)\, \mathrm{e}^{\lambda \cdot h(x)} \gamma(\mathrm{d}x). \tag{A.5}$$

The same argument may be applied to obtain higher derivatives. For instance, taking divided differences $h(x)\, \mathrm{e}^{\lambda \cdot h(x)}$ we can employ the same bound above to conclude

$$\int_\Gamma |D_{\epsilon v}(h(x)\, \mathrm{e}^{\lambda \cdot h(x)})| \gamma(\mathrm{d}x) \leq \epsilon_0^{-1} \left[|\nabla Z(\lambda + \epsilon_0 v)| + |\nabla Z(\lambda - \epsilon_0 v)| \right] < \infty.$$

Again using dominated convergence gives

$$\nabla^2 Z(\lambda) = \int_\Gamma h(x) \otimes h(x)\, \mathrm{e}^{\lambda \cdot h(x)} \gamma(\mathrm{d}x). \tag{A.6}$$

The formulas (A.1) and (A.2) follow immediately by applying the chain rule to $F(\lambda) = \log Z(\lambda)$

$$\nabla F(\lambda) = \frac{\nabla Z(\lambda)}{Z(\lambda)}, \quad \nabla^2 F(\lambda) = \frac{\nabla^2 Z(\lambda)}{Z(\lambda)} - \frac{\nabla Z(\lambda) \otimes \nabla Z(\lambda)}{Z(\lambda)^2}$$

and using the formulas (A.5) and (A.6).

To see strict convexity, note that the Hessian $\nabla^2 F$ can be written

$$\nabla^2 F(\lambda) = \int_\Gamma \left(h(x) - \int_\Gamma h(y) \mathrm{e}^{\lambda \cdot h(y) - F(\lambda)} \gamma(\mathrm{d}y) \right)^{\otimes 2} \mathrm{e}^{\lambda \cdot h(x) - F(\lambda)} \gamma(\mathrm{d}x),$$

and therefore can only be degenerate at a particular $\lambda \in \operatorname{Int} D_F$ if there is a direction $v \in \mathbb{R}^r$ such that for all $x \in \Gamma$

$$v \cdot h(x) = \int_\Gamma v \cdot h(y) \mathrm{e}^{\lambda \cdot h(y) - F(\lambda)} \gamma(\mathrm{d}y).$$

However as we assumed that $h(x) \cdot v$ is non-constant, this cannot be true. □

Next, we define the *entropy function* $S(u) : D_S \subset \mathbb{R}^d \to \mathbb{R}$ associated to F by the Legendre-Fenchel transform

$$S(u) = \sup_{\lambda \in D_F} (\lambda \cdot u - F(\lambda)),$$

where $D_S = \{u \in \mathbb{R}^r : |S(u)| < \infty\}$ is the domain of S.

Lemma A.2.2. *The entropy function S has the following properties:*

1. *D_S is convex and has non-empty interior*

2. *$S(u)$ is strictly convex and smooth on* $\operatorname{Int} D_S$.

3. For each $u \in D_S$, there exists a unique $\lambda_u \in D_F$ which is the minimizer of (A.3)

$$S(u) = u \cdot \lambda_u - F(\lambda_u).$$

Moreover λ_u satisfies,

$$u = \nabla F(\lambda_u), \quad \lambda_u = \nabla S(u), \quad \nabla^2 S(u) = [\nabla^2 F(\lambda_u)]^{-1}.$$

Proof. The fact that $S(u)$ convex follows from the definition of (A.3), since it is the supremum of linear functions. Next, since F is strictly convex and C^∞, $\lambda \mapsto \nabla F(\lambda)$ is a C^∞ diffeomorphism from D_F to D_S. It follows that for each $u \in D_S$, there is a unique $\lambda_u \in D_F$ satisfying $u = \nabla F(\lambda_u)$. Moreover for any $u \in D_S$, the function $f_u(\lambda) = \lambda \cdot u - F(\lambda)$ is strictly concave, and satisfies $\nabla f_u(\lambda_u) = 0$. Therefore $f_u(\lambda)$ has a unique maximum at $\lambda_u \in D_F$ so that

$$S(u) = f_u(\lambda_u) = u \cdot \lambda_u - F(\lambda_u).$$

Since $\lambda_u = (\nabla F)^{-1}(u)$, the mapping $u \mapsto \lambda_u$ is smooth. Therefore the smoothness of $S(u) = u \cdot \lambda_u - F(\lambda_u)$ follows. Moreover, taking the gradient of $S(u)$ yields

$$\nabla S(u) = \lambda_u + u \cdot \nabla_u \lambda_u - \nabla F(\lambda_u) \cdot \nabla_u \lambda_u = \lambda_u.$$

Finally, differentiating both sides of the relation, $u = \nabla F(\nabla S(u))$ gives

$$\nabla^2 S(u) = [\nabla^2 F(\lambda_u)]^{-1}.$$

□

Remark A.2.3. We say that $u \in D_S$ and $\lambda \in D_F$ are Legendre dual to each other if they are related by

$$u = \nabla F(\lambda), \quad \lambda = \nabla S(u).$$

In particular, any two Legendre dual variables u an λ must satisfy

$$\lambda \cdot u = F(\lambda) + S(u).$$

Abstract Canonical and Micro-canonical Ensembles

Next, we introduce the definition of the abstract cannonical and micro-cannonical ensembles associated to the function h. More specifically, these 'ensembles' refer to certain measures on the space of n particle configurations Γ^n.

We begin by defining, for each $\lambda \in D_F$, the *single particle Gibbs measure*

$$\mu_\lambda(\mathrm{d}x) \equiv \frac{1}{Z(\lambda)} \mathrm{e}^{\lambda \cdot h(x)} \gamma(\mathrm{d}x) = \mathrm{e}^{\lambda \cdot h(x) - F(\lambda)} \gamma(\mathrm{d}x).$$

Note that Lemma A.2.1 implies that with respect to $\mu_\lambda(\mathrm{d}x)$, the function $h(x)$ has mean $\nabla F(\lambda)$ and covariance $\nabla^2 F(\lambda)$.

Denote the pushforward of $\gamma(\mathrm{d}x)$ under h by $\nu(\mathrm{d}y) = h_\# \gamma(\mathrm{d}y)$, and consider, for each $\lambda \in D_F$ the *tilted probability measure*

$$\nu_\lambda(\mathrm{d}y) = \mathrm{e}^{\lambda \cdot x - F(\lambda)} \nu(\mathrm{d}y),$$

which is just the push forward of $\mu_\lambda(\mathrm{d}x)$ under h.

For the remainder of this section we will assume:

Hypothesis A.3.1.

1. *h and γ are such that, for each $\lambda \in D_F$, $\nu_\lambda(\mathrm{d}y)$ satisfies the non-lattice and integrability conditions on it's characteristic function stated in Hypothesis A.1.1 of Section A.1.2.*

2. Each component of $h : \Gamma \to \mathbb{R}^r$, h_i has compact superlevel sets.

Remark A.3.2. The first assumption of Hypothesis A.3.1 is to ensure that certain local limit theorems of Section A.1.3 apply. The second condition on h is to ensure that certain conditional measures are well-defined. It is important to note that the choice of compact super-level sets only comes from physical considerations, since h will typically be taken to be the negative of an energy function. However, one could just as easily assume that h has compact sublevel sets without changing any consequences of the theory below.

It will be useful to relate the rate function associated to $\nu_\lambda(\mathrm{d}y)$ to the free energy F and entropy S.

Lemma A.3.3. *Then the rate function $I_\lambda(u)$ associated to $\nu_\lambda(\mathrm{d}y)$ is given by*

$$I_\lambda(u) = S(u) + F(\lambda) - \lambda \cdot u.$$

Proof. Let $\gamma_h(\mathrm{d}y)$ denote the pushforward of $\gamma(\mathrm{d}x)$ under h. It follows that $\mu_{\lambda,h}(\mathrm{d}y)$ is just a tilted version of $\gamma_h(\mathrm{d}y)$,

$$\nu_{\lambda,h}(\mathrm{d}y) = \mathrm{e}^{\lambda \cdot y - F(\lambda)} \gamma_h(\mathrm{d}y).$$

The logarithmic moment generating function $L_\lambda(\alpha)$ associated to $\mu_{\lambda,h}(\mathrm{d}y)$ can then be written as

$$L_\lambda(\alpha) = F(\alpha + \lambda) - F(\lambda)$$

and therefore $D_{L_\lambda} = D_F - \lambda$. Taking the Legendre-Fenchel transform give the rate function

$$I_\lambda(u) = \sup_{\alpha+\lambda \in D_F} (\alpha \cdot u - F(\alpha + \lambda)) + F(\lambda) = S(u) + F(\lambda) - \lambda \cdot u.$$

□

Physically, we will think of h as a one-particle energy function (actually the negative of the energy) associated to a particle in phase space Γ. The measure $\mu_\lambda(\mathrm{d}x)$ is then thought of as an equilibrium measure for that particular particle. If one instead has n particles $\mathbf{x}_n = (x_1, \ldots, x_n) \in \Gamma^n$, then we will consider the average energy function $\hat{h}_n(\mathbf{x}_n)$ given by

$$\hat{h}_n(\mathbf{x}_n) = \frac{1}{n}\sum_{i=1}^{n} h(x_i).$$

The n particle *canonical ensemble* is then defined to be the product measure

$$\mu_\lambda^n(\mathrm{d}\mathbf{x}_n) = \mu_\lambda^{\otimes n}(\mathrm{d}\mathbf{x}_n) = \frac{1}{Z(\lambda)^n}\mathrm{e}^{n\lambda\cdot\hat{h}_n(\mathbf{x}_n)}\gamma^n(\mathrm{d}\mathbf{x}_n),$$

where we have denoted

$$\gamma^n(\mathrm{d}\mathbf{x}_n) = \gamma^{\otimes n}(\mathrm{d}\mathbf{x}_n).$$

We denote level sets of $\hat{h}_n : \Gamma \to \mathbb{R}^r$, for each $y \in \mathbb{R}^r$, by

$$\Sigma_y^n = \left\{\mathbf{x}_n \in (\mathbb{R}^d)^n \;:\; \hat{h}_n(\mathbf{x}_n) = y\right\}.$$

Then the assumption that h has compact sublevel sets implies that Σ_y^n is bounded. For each $y \in \mathbb{R}^r$, define the *micro-canonical measure* $\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)$, to be the probability measure on Σ_y^n produced by conditioning the canonical measure $\mu^n(\mathrm{d}\mathbf{x}_n)$ with respect to $\hat{h}_n$. Such a measure is given uniquely (up to $\hat{\mu}_{\lambda,n}(\mathrm{d}y)$ null sets) by disintegration

$$\mu_\lambda^n(\mathrm{d}\mathbf{x}_n) = \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)\,\hat{\mu}_{\lambda,n}(\mathrm{d}y), \tag{A.7}$$

where $\hat{\mu}_{\lambda,n}(\mathrm{d}y)$ denotes the pushforward of the canonical measure $\mu_\lambda^n(\mathrm{d}\mathbf{x}_n)$ under $\hat{h}_n$. The above decomposition is to be interpreted by integration against the test functions $\varphi \in C_b(\mathbb{R}^r)$ and $\psi \in C_b(\Gamma^n)$,

$$\begin{aligned}\int_{\Gamma^n} \varphi\left(\hat{h}_n(\mathbf{x}_n)\right)\psi(\mathbf{x}_n)\,\mu_\lambda^n(\mathrm{d}\mathbf{x}_n)\\ = \int_{\mathbb{R}^r}\varphi(y)\left(\int_{\Sigma_y^n}\psi(\mathbf{x}_n)\mu^n(\mathrm{d}\mathbf{x}_n\,|\,y)\right)\hat{\mu}_{\lambda,n}(\mathrm{d}y).\end{aligned} \tag{A.8}$$

The subscript λ is intentionally missing from $\mu^n(\mathrm{d}\mathbf{x}_n\,|\,y)$, since as we will see in the following lemma, it does not depend on λ.

Lemma A.3.4. *In addition to (A.7), the following decomposition also holds*

$$\gamma^n(\mathrm{d}\mathbf{x}_n) = \mu^n(\mathrm{d}\mathbf{x}_n\,|\,y)\hat{\gamma}_n(\mathrm{d}y), \tag{A.9}$$

where $\hat{\gamma}_n(\mathrm{d}y)$ is the pushforward of $\gamma^n(\mathrm{d}\mathbf{x}_n)$ under $\hat{h}_n$, and it is to be interpreted in the sense of equation (A.8). As a consequence, the micro-canonical measure $\mu^n(\mathrm{d}\mathbf{x}_n\,|\,y)$ does not depend on λ.

Proof. To see this, recall that $\mu_\lambda^n(\mathrm{d}\mathbf{x}_n)$ has the form

$$\mu_\lambda^n(\mathrm{d}\mathbf{x}_n) = \mathrm{e}^{n(\lambda\cdot\hat{h}_n(\mathbf{x}_n)-F(\lambda))}\gamma^n(\mathrm{d}\mathbf{x}_n),$$

and therefore $\hat{\mu}_{\lambda,n}(\mathrm{d}y)$ is given by

$$\hat{\mu}_{\lambda,n}(\mathrm{d}y) = \mathrm{e}^{n(\lambda\cdot y-F(\lambda))}\hat{\gamma}_n(\mathrm{d}\mathbf{x}_n).$$

Now, consider a test function of the type $\varphi(y) = \phi(y)\mathrm{e}^{-n\lambda\cdot y+nF(\lambda)}$, where ϕ has compact support contained in a ball of some radius R. The condition that h has compact superlevel sets implies that $\varphi(\hat{h}_n(y))$ has compact support and is bounded, since

$$\operatorname{supp}\varphi(\hat{h}_n(y)) = \{x : h(x) \in \operatorname{supp}\phi\} \subseteq \{x : |h(x)| \leq R\}.$$

Using this test function in (A.8) and employing the forms for $\mu_\lambda^n(\mathrm{d}\mathbf{x}_n)$ and $\hat{\mu}_{\lambda,n}(\mathrm{d}y)$ given above, we obtain

$$\begin{aligned}
&\int_{\Gamma^n} \phi\left(\hat{h}_n(\mathbf{x}_n)\right) \psi(\mathbf{x}_n)\, \gamma^n(\mathrm{d}\mathbf{x}_n) \\
&\qquad\qquad = \int_{\mathbb{R}^r} \phi(y) \left(\int_{\Sigma_y^n} \psi(\mathbf{x}_n) \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) \right) \hat{\gamma}_n(\mathrm{d}y).
\end{aligned}$$

□

When $\hat{\mu}_{\lambda,n}(\mathrm{d}y)$ has a positive density $\hat{f}_{\lambda,n}(y)$ then $\hat{\gamma}_n(\mathrm{d}y)$ has a density $g_n(y)$. So, formally one can take $\phi(z) = \delta(y - z)$ and we may write the micro-canonical measure $\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)$ as

$$\begin{aligned}
\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) &= \frac{1}{\hat{f}_{\lambda,n}(y)} \delta(y - \hat{h}_n(\mathbf{x}_n)) \mu_\lambda^n(\mathrm{d}\mathbf{x}_n) \\
&= \frac{1}{g_n(y)} \delta(y - \hat{h}_n(\mathbf{x}_n)) \gamma^n(\mathrm{d}\mathbf{x}_n).
\end{aligned}$$

and $g_n(y)$ is given by

$$g_n(y) = \int_{\Gamma^n} \delta(y - \hat{h}_n(\mathbf{x}_n)) \gamma(\mathrm{d}\mathbf{x}_n).$$

The function $g_n(y)$ is often refered in statistical mechanics literature as the *density of states*.

Equivalence of Ensembles

Since the cannonical measure is a product measure, it is often more convenient to take averages with, than is the case for the micro-cannonical measure

$\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)$. Indeed, this is often the motivation for using the canonical measure over the micro-canonical measure in applications. However, from physical considerations, the micro-canonical measure is the more natural measure to use due to the fact that for most systems of interest, the particle evolution takes place on the level sets of $\hat{h}_n$, and therefore is usually a natural ergodic invariant measure fore the dynamics. Indeed, the disintegration (A.7) implies that μ^n_λ is always a linear combination of the measures $\mu_n(\,\cdot\,|\,y)$ and therefore, for finite n, μ^n_λ cannot be an ergodic invariant measure.

However, one is often interested in studying the large n behavior of the measures $\mu^n_\lambda(\mathrm{d}\mathbf{x}_n)$ and $\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)$. Indeed, it is in this setting that physicists often justify the use of the canonical ensemble in place of the micro-canonical one. This approximation of the micro-canonical ensemble by the canonical one when n is large is often refered to as the *equivalence of ensembles*. It is precisely this equivalence that we will address in this section.

To begin, we will need the following generalization of the local large deviations theorem A.1.6.

Lemma A.3.5. *Let $\hat{\gamma}_n(\mathrm{d}y)$ be the push-forward of $\gamma^n(\mathrm{d}\mathbf{x}_n)$ under $\hat{h}_n(\mathbf{x}_n)$. Then for large enough n, $\hat{\gamma}_n(\mathrm{d}y)$ has a density $\hat{g}_n(y)$ and for each $y \in D_S$, the following limit holds*

$$\lim_{n\to\infty} \frac{1}{n} \log \hat{g}_n(y) = -S(y).$$

Proof. First note that we have the relation

$$\hat{\mu}_{\lambda,n}(\mathrm{d}y) = \mathrm{e}^{n(\lambda\cdot y - F(\lambda))}\hat{\gamma}_n(\mathrm{d}y),$$

and that $\hat{\mu}_{\lambda,n}(\mathrm{d}y)$ is just the pushforward of the product measure $\nu_\lambda(\mathrm{d}y)^{\otimes n}$ under the mapping $\mathbf{x}_n \mapsto n^{-1}(x_1 + \ldots + x_n)$. Since $\nu_\lambda(\mathrm{d}x)$ satisfies Hypothesis A.1.1 and has bounded moments, we can apply Theorem A.1.6 to conclude that, for large enough n, $\hat{\mu}_{\lambda,n}(\mathrm{d}y)$ has a density $\hat{f}_{\lambda,n}(y)$ and

$$\lim_{n\to\infty} \frac{1}{n} \log \hat{f}_{\lambda,n}(y) = -I_\lambda(y) = \lambda \cdot y - S(y) - F(\lambda), \tag{A.10}$$

where we have used Lemma A.3.3 to obtain the form of the rate function $I_\lambda(y)$. Since $\hat{\mu}_{\lambda,n}(\mathrm{d}y)$ and $\hat{\gamma}_n(\mathrm{d}y)$ are related by an absolutely continuous transformation, $\hat{\gamma}_n(\mathrm{d}y)$ has a density $\hat{g}_n(y)$ and satisfies

$$\frac{1}{n} \log \hat{g}_n(y) = F(\lambda) - \lambda \cdot y + \frac{1}{n} \log \hat{f}_{\lambda,n}(y).$$

Taking the limit as $n \to \infty$ and using (A.10) completes the proof. □

Remark A.3.6. In fact, we can do better than the lemma above. If we use the sharp asymptotic provided by Corollary A.1.7, we can obtain

$$\hat{g}_n(y) = \frac{\mathrm{e}^{-nS(y)}}{(2\pi)^{d/2}} \sqrt{n^d \det \nabla^2 S(y)} \left(1 + \mathcal{O}(n^{-1/2})\right).$$

Our goal is to establish an equivalence of ensembles theorem. In general we will show the following theorem

Theorem A.3.7. *Let G be a continuous, bounded function on Γ^k, and for each $y \in D_S$ let $\lambda_y = \nabla S(y)$. We have the following convergence,*

$$\lim_{n\to\infty} \int_{\Gamma^n} G(\mathbf{x}_k) \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) = \int_{\Gamma^k} G(\mathbf{x}_k) \mu^k_{\lambda_y}(\mathrm{d}\mathbf{x}_k),$$

locally uniformly on D_S.

We will need the following lemmas.

Lemma A.3.8. *Let G be a continuous bounded function on Γ, and define for each $\theta \in \mathbb{R}$ the unbounded measure*

$$\gamma_\theta(\mathrm{d}x) = \mathrm{e}^{\theta G}\gamma(\mathrm{d}x),$$

and let $F_\theta(\lambda)$ and $S_\theta(y)$ be it's free energy and entropy functions,

$$F_\theta(\lambda) = \log\left(\int_\Gamma \mathrm{e}^{\theta G + \lambda\cdot h}\gamma(\mathrm{d}x)\right), \quad S_\theta(y) = \sup_{\lambda\in D_{F_\theta}} (\lambda\cdot y - F_\theta(\lambda)).$$

Then we have the following,

1. *The domains of S_θ and F_θ coincide with those of S and F respectively.*
2. *S_θ and F_θ are differentiable in θ and, for any pair of Legendre dual variables $(y,\lambda) \in D_S \times D_F$, they satisfy*

$$\partial_\theta S_\theta(y) = -\partial_\theta F_\theta(\lambda) = -\int_\Gamma G(x)\mathrm{e}^{\lambda\cdot h(x) - F_\theta(\lambda)}\gamma_\theta(\mathrm{d}x).$$

In particular, this implies that

$$\partial_\theta S_\theta(y)\big|_{\theta=0} = -\int_\Gamma G(x)\mu_\lambda(\mathrm{d}x).$$

Proof. The fact that the domains of F_θ and S_θ are the same as those of F and S follows from the fact that G is bounded. Also differentiability of $F_\theta(\lambda)$ in θ follows from the fact that the divided differences

$$(\mathrm{e}^{(\theta+\epsilon)G} - \mathrm{e}^{\theta G})/\epsilon$$

are uniformly bounded in x for small ϵ and, since $\mathrm{e}^{\lambda\cdot h}\gamma(\mathrm{d}x)$ is a finite measure for $\lambda \in D_F$. Applying Lebesgue dominated convergence gives

$$\partial_\theta F_\theta(\lambda) = \int_\Gamma G(x)\mathrm{e}^{\lambda\cdot h(x)-F_\theta(\lambda)}\gamma_\theta(\mathrm{d}x).$$

To conclude the proof fix $y \in D_S$ and let $\lambda = \nabla S_\theta(y)$ be it's Legendre dual. It follows that

$$S_\theta(y) = \nabla S_\theta(y)\cdot y - F_\theta(\nabla S_\theta(y)).$$

Taking the derivatives of both sides in θ yields

$$\partial_\theta S_\theta(y) = \partial_\theta \nabla S_\theta(y)\cdot y - \nabla F(\nabla S_\theta(y))\cdot \partial_\theta\nabla S_\theta - \partial_\theta F_\theta(\nabla S_\theta) = -\partial_\theta F_\theta(\lambda).$$

□

Our main tool will be the following large deviation type theorem.

Lemma A.3.9. *Let G be a continuous bounded function on Γ and let $\hat{G}_n(\mathbf{x}_n) = \frac{1}{k}\sum_{i=1}^n G(x_i)$ be it's average (or sample mean). Then for each $\theta \in \mathbb{R}$ and $y \in D_S$,*

$$\lim_{n\to\infty}\frac{1}{n}\log\left(\int_{\Sigma_y^n}\mathrm{e}^{n\theta\hat{G}_n(\mathbf{x}_n)}\mu^n(\mathrm{d}\mathbf{x}_n\,|\,y)\right) = S(y) - S_\theta(y),$$

where S_θ is defined in Lemma A.3.8.

Proof. We begin by considering the decomposition (A.9) in weak form for test functions $\psi \in C_b(\Gamma^n)$ and $\varphi \in C_c(\mathbb{R}^r)$,

$$\int_{\Gamma^n}\varphi\left(\hat{h}_n(\mathbf{x}_n)\right)\psi(\mathbf{x}_n)\gamma^n(\mathrm{d}\mathbf{x}_n) = \int_{\mathbb{R}^r}\varphi(y)\left(\int_{\Sigma_y^n}\psi(\mathbf{x}_n)\mu^n(\mathrm{d}\mathbf{x}_n\,|\,y)\right)\hat{\gamma}_n(\mathrm{d}y).$$

Upon choosing $\psi(\mathbf{x}_n) = \mathrm{e}^{n\theta\hat{G}_n(\mathbf{x}_n)}$, and denoting $\hat{\gamma}_{\theta,n}(\mathrm{d}y)$ the pushforward of the product measure $\gamma_\theta^{\otimes n}(\mathrm{d}\mathbf{x}_n)$ under $\hat{h}_n(x)$, we find

$$\int_{\mathbb{R}^r}\varphi(y)\hat{\gamma}_{\theta,n}(\mathrm{d}y) = \int_{\mathbb{R}^r}\varphi(y)\left(\int_{\Sigma_y^n}\mathrm{e}^{n\theta\hat{G}_n(\mathbf{x}_n)}\mu^n(\mathrm{d}\mathbf{x}_n\,|\,y)\right)\hat{\gamma}_n(\mathrm{d}y). \tag{A.11}$$

Applying Lemma A.3.5 to both the measures $\hat{\gamma}_n(\mathrm{d}y)$ and $\hat{\gamma}_{\theta,n}(\mathrm{d}y)$ we conclude that they have densities $\hat{g}_n(y)$ and $\hat{g}_{\theta,n}(y)$. Moreover equation (A.11) implies that they are related by

$$\hat{g}_{\theta,n}(y) = \int_{\Sigma_y^n} \mathrm{e}^{n\theta \hat{G}_n(\mathbf{x}_n)} \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)\hat{g}_n(y).$$

Taking the log of both sides, we conclude

$$\frac{1}{n}\log\left(\int_{\Sigma_y^n} \mathrm{e}^{n\theta \hat{G}_n(\mathbf{x}_n)} \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)\right) = \frac{1}{n}\log \hat{g}_{\theta,n} - \frac{1}{n}\log \hat{g}_n.$$

Taking the limit as $n \to \infty$ and appealing to Lemma A.3.5 again yields the result.

□

We are now ready to prove Theorem A.3.7.

Proof of Theorem A.3.7. By the density of linear combinations of factored functions in $C_b(\Gamma^k)$, it suffices to prove the Theorem A.3.7 for functions of the form

$$G(x_1, x_2, \ldots, x_k) = G_1(x_1)G_1(x_2)\ldots G_k(x_k)$$

for $\{G_j\}_{j=1}^k$ a collection of continuous bounded functions on Γ. Without loss of generality, we may assume that

$$\int_{\Gamma^k} G(\mathbf{x}_k)\mu_{\lambda_y}^k(\mathrm{d}\mathbf{x}_k) = 0,$$

and therefore at least one of the functions $\{G_j\}_{j=1}^k$ is mean zero with respect to $\mu_{\lambda_y}(\mathrm{d}x)$. By the symmetry of the measure $\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)$ under permutations of the indices of $\mathbf{x}_n = (x_1, \ldots, x_n)$, we may assume that G_1 is mean zero, that is

$$\int_{\Gamma} G_1(x)\mu_{\lambda_y}(\mathrm{d}x) = 0.$$

Again using the permutation symmetry, we find

$$\int_{\Gamma^n} G_1(x_1)G_2(x_2)\dots G_k(x_k)\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)$$
$$= \int_{\Gamma^n} \left(\frac{1}{n-k+1} \sum_{j=1}^{n-k+1} G_1(x_j) \right) G_2(x_{n-k+2})\dots G_k(x_n)\, \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y).$$

Using the boundedness of $\{G_j\}_{j=1}^k$ we conclude that

$$\left| \int_{\Gamma^n} G_1(x_1)G_2(x_2)\dots G_k(x_k)\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) \right| \lesssim \int_{\Gamma^n} |\hat{G}_{1,n}(\mathbf{x}_n)|\, \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) + \frac{k}{n} \quad \text{(A.12)}$$

where

$$\hat{G}_{1,n}(\mathbf{x}_n) = \frac{1}{n} \sum_{i=1}^{n} G_1(x_i).$$

The proof will be complete if we can show that the first term on the right-hand-side of (A.12) vanishes as $n \to \infty$. With this in mind, by Jensen's inequality we may estimate

$$\begin{aligned} \int_{\Gamma^n} |\hat{G}_{1,n}(\mathbf{x}_n)|\, \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) &\le \frac{1}{n\theta} \log \left(\int_{\Gamma^n} \mathrm{e}^{n\theta|\hat{G}_{1,n}(\mathbf{x}_n)|}\, \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) \right) \\ &\le \frac{1}{n\theta} \log \left(\int_{\Gamma^n} \left[\mathrm{e}^{n\theta\hat{G}_{1,n}(\mathbf{x}_n)} + \mathrm{e}^{-n\theta\hat{G}_{1,n}(\mathbf{x}_n)} \right] \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) \right) \end{aligned} \quad \text{(A.13)}$$

Using the elementary fact that if $\{a_n\}$ and $\{b_n\}$ are two real sequences converging to a and b respectively, then

$$\lim_{n\to\infty} \frac{1}{n} \log \left(\mathrm{e}^{na_n} + \mathrm{e}^{nb_n} \right) = \max\{a, b\}.$$

We find upon sending $n \to \infty$ in (A.13) and applying Lemma A.3.9 that

$$\lim_{n\to\infty} \int_{\Gamma^n} |\hat{G}_{1,n}(\mathbf{x}_n)|\, \mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) \le \frac{1}{\theta} \max \left\{ S(y) - S_\theta(y), S(y) - S_{-\theta}(y) \right\}.$$

Where $S_\theta(y)$ is entropy corresponding to the measure $\mathrm{e}^{\theta G_1(x)}\gamma(\mathrm{d}x)$. Sending $\theta \to 0$ and appealing to Lemma A.3.8 we find

$$\lim_{\theta\to 0} \frac{1}{\theta} \max \left\{ S(y) - S_\theta(y), S(y) - S_{-\theta}(y) \right\} = \partial_\theta S_\theta(y)\big|_{\theta=0} = -\int_{\Gamma} G_1(x)\mu_{\lambda_y}(\mathrm{d}x) = 0.$$

□

In general, the symmetry of the measure $\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)$ and Theorem A.3.7 implies that

$$\lim_{n\to\infty} \int_{\Gamma^n} \hat{G}^n(\mathbf{x}_n)\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y) = \int_{\Gamma^k} G(\mathbf{x}_k)\mu^k_{\lambda_y}(\mathrm{d}\mathbf{x}_k),$$

where $\hat{G}^n(\mathbf{x}_n)$ is a sum of shifts of the function G given by

$$\hat{G}_n(\mathbf{x}_n) = \frac{1}{n-k+1}\sum_{i=0}^{n-k} G(x_{i+1}, x_{i+2}, \ldots, x_{i+k}). \tag{A.14}$$

Indeed, this resembles an ergodic theorem, giving convergence of the averages $\hat{G}_n$ to their canonical average $\int_{\Gamma^k} G\,\mathrm{d}\mu^n_\lambda$ with respect to the micro-canonical ensemble $\mu^n(\mathrm{d}\mathbf{x}_n \,|\, y)$. In fact, one can show the following stronger result, taken from Guo-Papanicolau-Varadhan [70].

Theorem A.3.10. *Let F be a bounded continuous function on Γ^k for some $1 \le k \le n$ and let $\hat{G}_n(\mathbf{x}_n)$ it's average given by (A.14). For each $y \in D_S$ let $\lambda_y = \nabla S(y) \in D_F$ be its Legendre dual variable. Define for each $\delta > 0$ and $y \in D_S$ the set*

$$A_{\delta,y} = \left\{ \mathbf{x}_n \in \Gamma^n \;:\; \left| \hat{G}_n(\mathbf{x}_n) - \int_{\Gamma^k} G(\mathbf{x}_k)\mu^k_{\lambda_y}(\mathrm{d}\mathbf{x}_k) \right| > \delta \right\}.$$

Then for each $y \in D_S$ and $\delta > 0$, there is a constant C, independent of n, y, δ, so that

$$\lim_{n\to\infty} \frac{1}{n} \log \mu^n(A_{\delta,y} \,|\, y) \le -C\delta^2,$$

uniformly on compact sets in y.

Stochastic Processes and Functional Analysis

Compactness and tightness criterion

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and $(E, \tau, \mathcal{B}_\tau)$ be a topological space endowed with its Borel sigma algebra. A mapping $X : \Omega \to (E, \tau)$ is called an "E valued random variable" provided it is a measurable mapping between these spaces. Every E valued valued random variable induces a probability measure on $(E, \tau, \mathcal{B}_\tau)$ by pushforward. A sequence of probability measures $\{\mathbf{P}_n\}_{n\in\mathbb{N}}$ on $\mathcal{B}_\tau$ is said to be "tight" provided that for each $\epsilon > 0$ there exists a τ compact set K_ϵ such that $\mathbf{P}_n(K_\epsilon) \geq 1 - \epsilon$ for all $n \in \mathbb{N}$.

Definition B.1.1. A topological space (E, τ) is called a Jakubowski space provided it admits a countable sequence continuous functionals which separate points.

Our main interest in such spaces is the following fundamental result given in [74].

Theorem B.1.2. *Let (E, τ) be a Jakubowski space. Suppose $\{\tilde{X}_n\}_{n\in\mathbb{N}}$ is a sequence of E valued random variables on a probability spaces $(\tilde{\Omega}, \mathcal{F}, \mathbf{P})$ inducing tight laws with respect to the topology τ. Then there exists a new probability space $(\Omega, \mathcal{F}, \mathbf{P})$ endowed with an E valued random variable X and a sequence of measurable maps*

$\{\widetilde{T}_n\}_{n\in\mathbb{N}}$

$$\widetilde{T}_n : (\Omega, \mathcal{F}, \mathbf{P}) \to (\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$$

with the following two properties:

1. *For each $n \in \mathbb{N}$, the measure $\tilde{\mathbf{P}}_n$ is the pushforward of $\mathbf{P}$ by $\widetilde{T}_n$.*

2. *The new sequence $\{X_n\}_{n\in\mathbb{N}}$ defined via $X_n = \widetilde{X}_n \circ \widetilde{T}_n$ converges $\mathbf{P}$ a.s. to X (with respect to the topology τ).*

We begin by recalling the following 'compact plus small ball" criterion for compactness in Frechet spaces.

Lemma B.1.3. *Let F be a Fréchet space. Then $U \subset F$ is precompact in F if for every $\epsilon > 0$, there exists a compact set $K_\epsilon \subset F$, such that*

$$U \subset K_\epsilon + B_\epsilon,$$

where B_ϵ is a ρ-ball centered at 0 of radius ϵ, for a given metric ρ.

Proof. Fix $\epsilon > 0$ and let K_ϵ be the compact set defined as above. Since K_ϵ is compact and F is a metric space, it is totally bounded. Therefore there exists a finite collection of points $\{x_i\}_{i=1}^N$ so that $K_\epsilon \subseteq \bigcup_{i=1}^N B_\epsilon(x_i)$. However, since

$$K \subseteq \bigcup_{i=1}^{N} B_\epsilon(x_i) + B_\epsilon(0) \subseteq \bigcup_{i=1}^{N} B_{2\epsilon}(x_i),$$

then K is totally bounded and therefore precompact in F. □

In the stochastic setting, we make use of the analogous version as a tightness criterion.

Lemma B.1.4. *Let F be a Frechet space and $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of F-valued random variables. Assume that for all $L \in \mathbb{R}_+$ there exists a decomposition*

$$X_n = Y_n^L + Z_n^L,$$

where $\{Y_n^L\}_{n\in\mathbb{N}}$ induces a tight sequence of laws on F. If in addition, Z_n^L satisfies for every $\eta > 0$,

$$\lim_{L\to\infty} \sup_n \mathbf{P}\left(Z_n^L \notin B_\eta\right) = 0.$$

Then X_n induces tight laws on F.

Proof. Fix $\epsilon > 0$ and choose a sequence $\{L_j\}_{j\in\mathbb{N}}$ so that

$$\sup_n \mathbf{P}\left(Z_n^{L_j} \notin B_{1/j}\right) < \epsilon/2^j.$$

By the tightness of Y_n^L, for each $j \in \mathbb{N}$ there is a compact set $K_j \subseteq F$ such that

$$\sup_n \mathbf{P}\left(Y_n^{L_j} \in K_j\right) < \epsilon/2^j$$

By the classical compactness criterion, Lemma B.1.3, the set

$$K = \bigcap_j (K_j + B_{1/j}).$$

is compact in E. It follows that

$$\sup_n \mathbf{P}(X_n \notin K) \leq \sum_j \left(\sup_n \mathbf{P}\left(Y_n^{L_j} \notin K_j\right) + \sup_n \mathbf{P}\left(Z_n^{L_j} \notin B_{1/j}\right)\right) < 2\epsilon.$$

Therefore $\{X_n\}_{n\in\mathbb{N}}$ induce tight laws on F. □

Next, we recall the classical Dunford-Pettis compactness criterion on $[L^1]_{\mathrm{w,loc}}$.

Lemma B.1.5. *Let K be a bounded subset of $[L^1(\mathbb{R}^d)]_{\mathrm{loc}}$, then K is precompact in $[L^1(\mathbb{R}^d)]_{\mathrm{w,loc}}$ if and only if the following limit holds*

$$\lim_{L\to\infty} \sup_{f\in K} \|f\mathbb{1}_{|f|>L}\|_{L^1} = 0.$$

In the stochastic setting, the corresponding tightness condition is:

Lemma B.1.6. *Let μ_n be a sequence of probability measures on $L^1(\mathbb{R}^d)_{\mathrm{loc}}$, then $\{\mu_n\}_{n\in\mathbb{N}}$ are tight on $[L^1(\mathbb{R}^d)]_{\mathrm{w,loc}}$ if and only if for every $\eta > 0$ the following limit hold*

$$\lim_{L\to\infty} \sup_{n} \mu_n \left\{ f : \|f\mathbb{1}_{|f|>L}\|_{L^1} > \eta \right\} = 0. \tag{B.1}$$

Proof. First suppose that the limits (B.1) hold. Let $\epsilon > 0$ and choose a sequence $\{L_k\}$ such that

$$\sup_{n} \mu_n \left\{ f : \|f\mathbb{1}_{|f|>L_k}\|_{L^1} > 1/k \right\} < \epsilon 2^{-k}.$$

Define the closed set

$$A_k = \left\{ f : \|f\mathbb{1}_{|f|>L_k}\|_{L^1} \le 1/k \right\}.$$

Then by the classical compactness criterion in Lemma B.1.5,

$$K = \bigcap_k A_k$$

is a compact set in $[L^1(\mathbb{R}^d)]_{\mathrm{w,loc}}$. Furthermore, we have

$$\sup_n \mu_n(K) \le \sum_k \sup_n \mu_n(A_k) < \epsilon.$$

Therefore $\{\mu_n\}$ are tight on $[L^1(\mathbb{R}^d)]_{\mathrm{w,loc}}$.

Next suppose that $\{\mu_n\}$ are tight on $[L^1(\mathbb{R}^d)]_{\mathrm{w}}$. And let K be a compact subset of $[L^1]_{\mathrm{w}}$ such that $\sup_n \mu_n(K^c) < \epsilon$. For each $\eta > 0$ it follows by the compactness criterion in Lemma B.1.5 that for large enough L (depending on η), the following set is empty

$$\left\{f \in K : \|f \mathbb{1}_{|f|>L}\|_{L^1} > \eta\right\} = \emptyset.$$

Therefore for large enough L we have

$$\sup_n \mu_n \left\{f : \|f \mathbb{1}_{|f|>L}\|_{L^1} > \eta\right\} \leq \sup_n \mu_n(K^c) < \epsilon.$$

□

We now introduce a useful tightness criterion for probability measures on $C_t([L^1_x]_w)$. First we will need a basic criterion for compactness in $C_t([L^1_x]_w)$.

Lemma B.1.7. *Let $K \subseteq C([0,T];[L^1(\mathbb{R}^d)]_w)$ and denote for each $\varphi \in C_c^\infty(\mathbb{R}^d)$, the set*

$$K_\varphi = \{\langle f, \varphi\rangle \,:\, f \in K\} \subseteq C([0,T]).$$

Then K is precompact in $C([0,T]\,;\,[L^1(\mathbb{R}^d)]_w)$ if any only if K is a weakly precompact subset of $L^\infty([0,T];L^1(\mathbb{R}^d))$ and K_φ equicontinuous in $C([0,T])$ for each $\varphi \in C_c^\infty(\mathbb{R}^d)$.

This gives rise to the following tightness criterion on $C_t([L^1_x]_w)$.

Lemma B.1.8. *Let $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures on $C([0,T],[L^1(\mathbb{R}^d)]_w)$, and for any $\varphi \in C_c^\infty(\mathbb{R}^d)$, let $\{\nu_n^\varphi\}_{n\in\mathbb{N}}$ be the sequence of measures on $C([0,T])$ induced by the mapping $f \mapsto \langle f, \varphi\rangle$. Then the measures $\{\mu_n\}_{n\in\mathbb{N}}$ are tight if and only if $\{\nu_n^\varphi\}_{n\in\mathbb{N}}$ are tight for every $\varphi \in C_c^\infty$ and for every $\eta > 0$ we have*

$$\lim_{M\to\infty} \sup_n \mu_n \left\{f : \|f\|_{L_t^\infty(L^1)} > M\right\} = 0,$$

$$\lim_{L\to\infty}\sup_n \mu_n\Big\{f : \|f\mathbb{1}_{|f|>L}\|_{L_t^\infty(L^1)} > \eta\Big\} = 0,$$

and

$$\lim_{R\to\infty}\sup_n \mu_n\Big\{f : \|f\mathbb{1}_{B_R^c}\|_{L_t^\infty(L^1)} > \eta\Big\} = 0,$$

Proof. Define for any function $f \in C([0,T])$ and $\delta > 0$ the modulus of continuity

$$\omega_\delta(f) := \sup_{|t-s|<\delta} |f(t) - f(s)|.$$

We prove sufficiency first. Let $\epsilon > 0$, and let $\{\varphi_j\}$ be a dense subset of $C_c^\infty(\mathbb{R}^d)$. Then by the classical tightness criterion for functions in $C([0,T])$, we can conclude that for each $\eta > 0$ and φ_j, we have

$$\lim_{\delta\to 0}\sup_n \mu_n\Big\{f : \omega_\delta(\langle f, \varphi_j\rangle) > \eta\Big\} = 0.$$

Therefore for each $j, k \geq 0$ we may choose values $(M_k, L_k, R_k, \delta_{k,j})$ so that

$$\sup_n \mu_n\Big\{f : \|f\|_{L_t^\infty(L^1)} > M_k\Big\} < \epsilon 2^{-k}$$

$$\sup_n \mu_n\Big\{f : \|f\mathbb{1}_{|f|>L_k}\|_{L_t^\infty(L^1)} > 1/k\Big\} < \epsilon 2^{-k}$$

$$\sup_n \mu_n\Big\{f : \|f\mathbb{1}_{B_{R_k}^c}\|_{L_t^\infty(L^1)} > 1/k\Big\} < \epsilon 2^{-k}$$

$$\sup_n \mu_n\Big\{f : \omega_{\delta_{k,j}}(\langle f, \varphi_j\rangle) > 1/k\Big\} < \epsilon 2^{-k-j}.$$

Define the closed sets,

$$A_k = \Big\{f : \|f\|_{L_t^\infty(L^1)} \leq M_k\Big\} \qquad B_k = \Big\{f : \|f\mathbb{1}_{|f|>L_k}\|_{L_t^\infty(L^1)} \leq 1/k\Big\}$$

$$C_k = \Big\{f : \|f\mathbb{1}_{B_{R_k}^c}\|_{L_t^\infty(L^1)} \leq 1/k\Big\} \qquad D_{k,j} = \Big\{f : \omega_{\delta_{k,j}}(\langle f, \varphi_j\rangle) \leq 1/k\Big\}.$$

and let

$$K = \bigcap_{j,k} A_k \cap B_k \cap C_k \cap D_{k,j}.$$

By the compactness criterion in Lemma B.1.7 it is straight forward to verify that K is a compact subset of $C([0,T]\,;\,[L^1]_w)$. Furthermore, we have

$$\mu_n(K^c) \leq \sum_k \mu_n(A_k^c) + \sum_k \mu_k(B_k^c) + \sum_k \mu_k(C_k^c) + \sum_{k,j} \mu_n(D_{k,j}^c) < 4\epsilon,$$

whereby tightness follows.

To prove necessity. We remark that since $f \mapsto \langle f, \varphi \rangle$ is continuous from $C([0,T]\,;\,[L^1]_w)$ to $C([0,T])$ for every $\varphi \in C_c^\infty(\mathbb{R}^d)$, then tightness of $\{\mu_n\}_{n\in\mathbb{N}}$ automatically implies tightness of $\{\nu_n^\varphi\}_{n\in\mathbb{N}}$. Now let $\epsilon > 0$ and let K be the compact subset of $C([0,T]\,;\,[L^1]_w)$ such that $\sup_n \mu_n(K^c) < \epsilon$. Fix an $\eta > 0$. The compactness criterion in Lemma B.1.7 implies that there exist (M', L', R') such that for and $M > M'$, $L > L'$, $R > R'$ the following sets are empty

$$\Big\{ f \in K : \|f\|_{L_t^\infty(L^1)} > M \Big\} = \emptyset,$$

$$\Big\{ f \in K : \|f \mathbb{1}_{|f|>L}\|_{L_t^\infty(L^1)} > \eta \Big\} = \emptyset,$$

$$\Big\{ f \in K : \|f \mathbb{1}_{B_R^c}\|_{L_t^\infty(L^1)} > \eta \Big\} = \emptyset,$$

Therefore, for such M, L and R large enough, we have

$$\mu_n \Big\{ f : \|f\|_{L_t^\infty(L^1)} > M \Big\} \leq \mu_n(K^c) < \epsilon,$$

$$\mu_n \Big\{ f : \|f \mathbb{1}_{|f|>L}\|_{L_t^\infty(L^1)} > \eta \Big\} \leq \mu_n(K^c) < \epsilon$$

$$\mu_n \Big\{ f : \|f \mathbb{1}_{B_R^c}\|_{L_t^\infty(L^1)} > \eta \Big\} \leq \mu_n(K^c) < \epsilon.$$

This completes the proof. □

We have the following representation and compactness criterion for $L^p_{t,x}(\mathcal{M}_v^*)$.

Lemma B.1.9. *The space $L^p_{t,x}(\mathcal{M}_v^*)$ $p \in [1,\infty]$ is continuously linearly isomorphic to $\mathcal{L}(C_0(\mathbb{R}^d), L^p_{t,x})$ the space of continuous linear operators from $C_0(\mathbb{R}^d)$ to $L^p_{t,x}$ under* *the topology of pointwise convergence. Similarly* $[L^p_{t,x}(\mathcal{M}^*_v)]_{\rm loc}$ *is continuously linearly isomorphic to* $\mathcal{L}(C_0(\mathbb{R}^d), [L^p_{t,x}]_{\rm loc})$.

Proof. For each $f \in L^p_{t,x}(\mathcal{M}^*_v)$ we can trivially associate a bounded linear operator $S_f : C_0(\mathbb{R}^d) \to L^p_{t,x}$, by $S_f\phi = \langle f, \phi\rangle$, clearly the map $f \mapsto S_f$ is one-to-one, linear, and continuous from $L^p_{t,x}(\mathcal{M}^*_v)$ to $\mathcal{L}(C_0(\mathbb{R}^d), L^p_{t,x})$ with it's pointwise topology.

Conversely for each bounded linear operator $S \in \mathcal{L}(C_0(\mathbb{R}^d), L^p_{t,x})$ one may define for each $g \in L^q_{t,x}$, $q = p/(p-1)$, the bounded linear functional $h_g : C_0(\mathbb{R}^d) \to \mathbb{R}$, by $h_g\phi = \langle S\phi, g\rangle$ which, by the Riesz-Markov theorem can be represented by a measure $f_g \in \mathcal{M}_v$, satisfying

$$h_g\phi = \langle f_g, \phi\rangle = \langle S\phi, g\rangle.$$

Since the mapping $g \mapsto f_g$ is clearly a continuous linear mapping from $L^q_{t,x}$ to $\mathcal{M}^*_v$, one can readily prove that for any bounded Borel $E \subset [0,T] \times \mathbb{R}^d$, that $\nu(E) = f_{\mathbb{1}_E}$ defines an $\mathcal{M}_v$ valued measure that $\mathrm{d}t\mathrm{d}x$ absolutely continuous and of σ finite variation. Since $\mathcal{M}_v$ is a dual space, it has the weak-* Radon-Nikodym property (see [97] Theorem 9.1) and therefore there is a measurable function $f_S : [0,T] \times \Omega \to \mathcal{M}^*_v$ such that $|\langle f_S, \phi\rangle| \in [L^1_{t,x}]_{\rm loc}$ and

$$\langle S\phi, 1_E\rangle = \langle \nu(E), \phi\rangle = \iint_E \langle f_S, \phi\rangle \mathrm{d}x\mathrm{d}t.$$

Using density of simple functions in $L^q_{t,x}$ we can conclude

$$\langle S\phi, g\rangle = \int_0^T \int_{\mathbb{R}^d} \langle f_S, \phi\rangle g \mathrm{d}x\mathrm{d}t, \tag{B.2}$$

for any $g \in L^q_{t,x}$. Taking the sup in $g \in L^q_{t,x}$, $\|g\|_{L^q} = 1$, on both sides of (B.2) we find

$$\|\langle f,\phi\rangle\|_{L^p_{t,x}} = \|T\phi\|_{L^p_{t,x}} < \infty$$

and therefore $f \in L^p_{t,x}(\mathcal{M}_v^*)$. Moreover this identity implies that the mapping $S \mapsto f_S$ is continuous from $\mathcal{L}(C_0(\mathbb{R}^d), L^p_{t,x})$ with it's pointwise topology to $L^p_{t,x}(\mathcal{M}_v^*)$, while identity (B.2) implies that $S \mapsto f_S$ is linear and one-to-one.

The proof on $[L^p_{t,x}(\mathcal{M}_v^*)]_{\text{loc}}$ is similar and can be proved by the above argument on compact sets of $[0,T]\times\mathbb{R}^d$. □

Lemma B.1.10. *Let K be subset of $[L^p_{t,x}(\mathcal{M}_v^*)]_{\text{loc}}$, $p\in[1,\infty]$, and let $\{\phi_k\}_{k=1}^\infty \subseteq C_c^\infty(\mathbb{R}^d_v)$ be a countable dense subset of $C_0(\mathbb{R}^d)$. Define the map $\Pi_{\phi_k} : [L^p_{t,x}(\mathcal{M}_v^*)]_{\text{loc}} \to [L^p_{t,x}]_{\text{loc}}$ by*

$$\Pi_{\phi_k}(f) = \langle f,\phi_k\rangle.$$

Then K is a compact subset of $[L^p_{t,x}(\mathcal{M}_v^*)]_{\text{loc}}$ if and only if K is bounded in $L^p_{t,x}(\mathcal{M}_v^*)$ and $\Pi_{\phi_j}K$ is compact in $[L^p_{t,x}]_{\text{loc}}$ for all $j\geq 1$.

Proof. Let $\{f_n\}_{n=1}^\infty \subseteq K$, and assume that $j\geq 1$, $\{\langle f_n,\phi_j\rangle\}_{n=1}^\infty$ is compact in $[L^p_{t,x}]_{\text{loc}}$. By a standard argument we may produce a diagonal subsequence, still denoted $\{f_n\}_{n=1}^\infty$, such that $\langle f_n,\phi_j\rangle$ converges as $n\to\infty$ for each $j\geq 1$. Identify $[L^p_{t,x}(\mathcal{M}_v)]_{\text{loc}}$ with $\mathcal{L}(C_0(\mathbb{R}^d);[L^p_{t,x}]_{\text{loc}})$ as in Lemma B.1.9, and for each $f\in[L^p_{t,x}(\mathcal{M}_v)]_{\text{loc}}$ let T_f denote the corresponding element of $\mathcal{L}(C_0(\mathbb{R}^d);[L^p_{t,x}]_{\text{loc}})$. Since $\{f_n\}_{n=1}^\infty$ is bounded in $[L^p_{t,x}(\mathcal{M}_v^*)]_{\text{loc}}$, we have for any compact set $C\subset[0,T]\times\mathbb{R}^d$,

$$\sup_n \|T_{f_n}\phi\|_{L^p_{t,x}(C)} = \|\langle f_n,\phi\rangle\|_{L^p_{t,x}(C)} < \infty.$$

By the uniform boundedness principle,

$$\sup_n \|\mathbb{1}_C T_{f_n}\|_{\mathcal{L}(C_0(\mathbb{R}^d);[L^p_{t,x}]_{\text{loc}})} < \infty.$$

Therefore the mappings $\phi \mapsto \mathbb{1}_C T_{f_n}\phi = \mathbb{1}_C\langle f_n, \phi\rangle$ are equicontinuous. Since $\{\phi_j\}_{j=1}^\infty$ is dense in $C_0(\mathbb{R}^d)$, this equicontinuity implies that for each $\phi \in C_0(\mathbb{R}^d)$, $\{\mathbb{1}_C\langle f_n, \phi\rangle\}_{n=1}^\infty$ is Cauchy in $L^p_{t,x}(C)$ and therefore $\{\langle f_n, \phi\rangle\}_{n=1}^\infty$ is convergent in $[L^p_{t,x}]_{\text{loc}}$. This limit defines a mapping $f : C_0(\mathbb{R}^d) \to [L^p_{t,x}]_{\text{loc}}$, by

$$f(\phi) \equiv \lim_{k\to\infty} \langle f_n, \phi\rangle.$$

It is a simple consequence of the linearity of $\langle f_n, \cdot\rangle$ and the boundedness of $\{f_n\}_{n=1}^\infty$, that the limiting f belongs to $\mathcal{L}(C_0(\mathbb{R}^d), [L^p_{t,x}]_{\text{loc}})$, and therefore belongs to $[L^p_{t,x}(\mathcal{M}^*_v)]_{\text{loc}}$. Therefore K is sequentially compact. Compactness of K now follows from the fact that $[L^p_{t,x}(\mathcal{M}^*_v)]_{\text{loc}}$ is a sequential space.

The converse is simple. If K is compact, since Π_{ϕ_j} are continuous, $\Pi_{\phi_j}K$ are compact in $[L^p_{t,x}]_{\text{loc}}$. □

Lemma B.1.11. *Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space, and let $\{f_n\}_{n\in\mathbb{N}}$ be a bounded sequence in $L^p(\Omega \times [0,T] \times \mathbb{R}^{2d})$ for some $p \in [1,\infty]$. Then $\{f_n\}_{n\in\mathbb{N}}$ induces a tight family of laws on $[L^p_{t,x}(\mathcal{M}^*_v)]_{\text{loc}}$ if and only if for all $\varphi \in C^\infty_c(\mathbb{R}^d_v)$, the sequence $\{\langle f_n, \varphi\rangle\}_{n\in\mathbb{N}}$ induces a tight family of laws on $[L^p_{t,x}]_{\text{loc}}$.*

Proof. Clearly if $\{f_n\}_{n\in\mathbb{N}}$ induce tight laws on $[L^p_{t,x}(\mathcal{M}^*_v)]_{\text{loc}}$ then for each $\varphi \in C^\infty_c(\mathbb{R}^d_v)$, since the mapping $f \mapsto \langle f, \varphi\rangle$ is continuous from $[L^p_{t,x}(\mathcal{M}^*_v)]_{\text{loc}}$ to $[L^p_{t,x}]_{\text{loc}}$, $\{\langle f_n, \varphi\rangle\}_{n\in\mathbb{N}}$ is tight on $[L^p_{t,x}]_{\text{loc}}$.

We proceed in the other direction by explicitly constructing a set K which is compact in $[L^p_{t,x}(\mathcal{M}^*_v)]_{\text{loc}}$ which has uniformly small probability. Fix and $\epsilon > 0$ and let $\{\varphi_j\}_{j=1}^{\infty} \subseteq C_c^\infty(\mathbb{R}_v^d)$ be a dense subset of $C_0(\mathbb{R}_v^d)$. Since $\{\langle f_n, \varphi_j\rangle\}_{n\in\mathbb{N}}$ induce tight laws in $[L^p_{t,x}]_{\mathrm{loc}}$, then for each $j \in \mathbb{N}$ there exist a compact set K_j in $[L^p_{t,x}]_{\mathrm{loc}}$ such that

$$\sup_n \mathbf{P}\{\langle f_n, \varphi_j\rangle \notin K_j\} < \epsilon 2^{-j}.$$

Define, as in Lemma B.1.10, $\Pi_{\varphi_j} f = \langle f, \varphi_j\rangle$. Since Π_{φ_j} is continuous from $[L^p_{t,x}(\mathcal{M}_v^*)]_{\mathrm{loc}}$ to $[L^p_{t,x}]_{\mathrm{loc}}$, the pre-images $\Pi_{\phi_j}^{-1} K_j$ are closed in $[L^p_{t,x}(\mathcal{M}_v^*)]_{\mathrm{loc}}$. Let $C = \sup_n \mathbf{E}\|f_n\|_{L^p_{t,x,v}}$ and define

$$B = \left\{ f \in L^p_{t,x,v} \,:\, \|f\|_{L^p_{t,x,v}} \leq C\epsilon^{-1} \right\}$$

and note that B is a bounded subset of $[L^p_{t,x}(\mathcal{M}_v^*)]_{\mathrm{loc}}$. Now, define the closed set

$$K = \bigcap_{j=1}^{\infty} \left(B \cap \Pi_{\varphi_j}^{-1} K_j \right),$$

and note that $K \subseteq B$ is a bounded subset of $[L^p_{t,x}(\mathcal{M}_v^*)]_{\mathrm{loc}}$, and for each $j \in \mathbb{N}$, $\Pi_{\varphi_j} K$ is a closed subset of K_j, so the set $\Pi_{\varphi_j} K$ is compact in $[L^p_{t,x}]_{\mathrm{loc}}$. Therefore Lemma B.1.10 implies that K is compact in $[L^p_{t,x}(\mathcal{M}_v^*)]_{\mathrm{loc}}$. We conclude the proof with

$$\mathbf{P}\{f_n \notin K\} \leq \mathbf{P}\left\{ \|f\|_{L^p_{t,x,v}} > C\epsilon^{-1} \right\} + \sum_{j=1}^{\infty} \mathbf{P}\{\langle f_n, \varphi_j\rangle \notin K_j\} < 2\epsilon.$$

□

The following product-limit lemma can be established in a classical way, using Egorov's theorem.

Lemma B.1.12. *Let $\{g_n\}_{n\in\mathbb{N}}$ and $\{h_n\}_{n\in\mathbb{N}}$ be sequences in $L^1_{t,x,v}$. Assume that $\{g_n\}_{n\in\mathbb{N}}$ is uniformly bounded in $L^\infty_{t,x,v}$ and converges to g in measure on $[0,T]\times\mathbb{R}^{2d}$. Then we have the following:*

1. If the sequence $\{h_n\}_{n\in\mathbb{N}}$ converges to h in $[L^1_{t,x,v}]_w$, then the sequence of products $\{g_nh_n\}_{n\in\mathbb{N}}$ converge to gh in $[L^1_{t,x,v}]_w$.

2. If the sequence $\{h_n\}_{n\in\mathbb{N}}$ converges to h in $[L^1_{t,x,v}]_w \cap L^1_{t,x}(\mathcal{M}^*_v)$, then the sequence of products $\{g_nh_n\}_{n\in\mathbb{N}}$ converge to gh in $L^1_{t,x}(\mathcal{M}^*_v)$.

The next lemma provides a procedure for identifying a continuous, adapted process as a series of one dimensional stochastic integrals.

Lemma B.1.13. *Let $(\Omega, \mathcal{F}, \mathbf{P}, \{\mathcal{F}^t\}_{t=0}^T, \{\beta_k\}_{k=1}^\infty)$ be a stochastic basis and let $(M_t)_{t=0}^T$ be a continuous $(\mathcal{F}_t)_{t=0}^T$ martingale with the quadratic variation process $(\int_0^t |f_s|^2_{\ell^2(\mathbb{N})})_{t=0}^T$. Moreover, assume that for each $k \in \mathbb{N}$ the cross variation of $(M_t)_{t=0}^T$ with β_k is given by the process $(\int_0^t f_k(s)ds)_{t=0}^T$. Under these hypotheses, the martingale may be identified as*

$$M_t = \sum_{k=1}^\infty \int_0^t f_k(s) \mathrm{d}\beta_k(s).$$

L^2 Stochastic Velocity Averaging

Proof of Lemma 7.4.4. For convenience we denote the velocity averaged process by

$$\rho^\phi(t, x; \omega) = \langle f, \varphi\rangle(t, x; \omega).$$

To begin, we assume that f is regular enough for all the following computations to be well defined. Let $\mathcal{F}_x$ denote the Fourier transform in x and let ξ be the corresponding Fourier variable, for simplicity denote $\widehat{f} = \mathcal{F}_x(f)$ and $\widehat{g} = \mathcal{F}_x(g)$. Taking the Fourier transform of both sides of (7.28) in Itô form gives

$$\partial_t \widehat{f} + iv\cdot\xi\widehat{f} + \mathcal{F}_x(\mathrm{div}_v(f\sigma_k\dot{\beta}_k)) = \mathcal{F}_x(\mathcal{L}_\sigma f) + \widehat{g}.$$

If $|\xi| \leq 1$ we have the simple estimate

$$\mathbf{E}\int_0^T \int_{\mathbb{R}^d} |\xi|^{1/3} |\widehat{\rho}^{\phi}|^2 \, \mathbb{1}_{|\xi|\leq 1} \, \mathrm{d}\xi \mathrm{d}t \leq \|\phi\|^2_{L^\infty_v} \mathbf{E}\|f\|^2_{L^2_{t,x,v}}.$$

To show the $H^{1/6}_x$ estimate, it suffices to consider $|\xi| \geq 1$. We will find it useful to solve this equation with the addition of a damping term on both sides (corresponding to a pseudo-differential operator acting on f in x). Let $\lambda \in C^\infty(\mathbb{R}^d_\xi)$, we now consider

$$\partial_t \widehat{f} + iv\cdot\xi \widehat{f} + \mathcal{F}_x(\mathrm{div}_v(f\sigma_k \dot{\beta}_k)) + \lambda \widehat{f} = \mathcal{F}_x(\mathcal{L}_\sigma f) + \widehat{g} + \lambda \widehat{f}.$$

Solving this via Duhammel, we find

$$\begin{aligned} \widehat{f}(t,\xi,v) &= e^{-(\lambda(\xi)+iv\cdot\xi)t}\widehat{f}_0(\xi,v) + \lambda(\xi)\int_0^t e^{-(\lambda(\xi)+iv\cdot\xi)(t-s)}\widehat{f}(s,\xi,v)\,\mathrm{d}s \\ &\quad + \int_0^t e^{-(\lambda(\xi)+iv\cdot\xi)(t-s)}\widehat{g}(s,\xi,v)\,\mathrm{d}s + \int_0^t e^{-(\lambda(\xi)+iv\cdot\xi)(t-s)}\mathcal{F}_x(\mathcal{L}_\sigma f)(s,\xi,v)\,\mathrm{d}s \\ &\quad - \sum_{k=1}^\infty \int_0^t e^{-(\lambda(\xi)+iv\cdot\xi)(t-s)}\mathcal{F}_x(\mathrm{div}_v(\sigma_k f))(s,\xi,v)\mathrm{d}\beta_k(s). \end{aligned} \tag{B.3}$$

Let $\phi \in C^\infty_c(\mathbb{R}^d_v)$, upon multiplying both sides of (B.3) by ϕ and integrating in v, we see that the velocity average $\widehat{\rho}^{\,\phi}$ satisfies

$$\begin{aligned} \widehat{\rho}^{\,\phi}(t,\xi) &= \int_{\mathbb{R}^d} e^{-(\lambda(\xi)+iv\cdot\xi)t}\phi(v)\widehat{f}_0(\xi,v)\,\mathrm{d}v \\ &\quad + \int_0^t \left(\int_{\mathbb{R}^d} e^{-(\lambda(\xi)+iv\cdot\xi)(t-s)}\widehat{\Gamma}_0(s,\xi,v)\,\mathrm{d}v\right)\mathrm{d}s \\ &\quad - \sum_{k=1}^\infty \int_0^t \left(\int_{\mathbb{R}^d} e^{-(\lambda(\xi)+iv\cdot\xi)(t-s)}\mathcal{F}_x(\phi\,\mathrm{div}_v(\sigma_k f))(s,\xi,v)\,\mathrm{d}v\right)\mathrm{d}\beta_k(s) \end{aligned} \tag{B.4}$$

Where Γ_0 is defined so that

$$\widehat{\Gamma_0}(t,\xi,v) = \phi(v)\left(\lambda(\xi)\widehat{f}(t,\xi,v) + \widehat{g}(t,\xi,v) + \mathcal{F}_x(\mathcal{L}_\sigma f)(t,\xi,v)\right). \tag{B.5}$$

Note that the v integrals in equation (B.4), can be written as a Fourier transform in v. We will denote such a Fourier transform in both x and v as $\mathcal{F}_{x,v}$, and denote by η the Fourier variable dual to v. We find

$$\begin{aligned}\widehat{\rho}^{\phi}(t,\xi) = e^{-\lambda(\xi)t}\mathcal{F}_{x,v}(\phi(v)f_0)(\xi,\xi t) + \int_0^t e^{-\lambda(\xi)(t-s)}\mathcal{F}_{x,v}(\Gamma_0)(s,\xi,\xi(t-s))\mathrm{d}s \\ - \sum_{k=1}^{\infty}\int_0^t e^{-\lambda(\xi)(t-s)}\mathcal{F}_{x,v}(\phi\,\mathrm{div}_v(\sigma_k f))(s,\xi,\xi(t-s))\mathrm{d}\beta_k(s) \\ = I_1 + I_2 + I_3.\end{aligned}$$

The first term, I_1, we can bound

$$|I_1|(t,\xi) \le |\mathcal{F}_{x,v}(\phi\, f_0)(\xi,\xi t)|.$$

For the second term, I_2, we have by Cauchy-Schwartz

$$\begin{aligned}|I_2|^2(t,\xi) &\le \left(\int_0^t e^{-2\lambda(\xi)(t-s)}\,\mathrm{d}s\right)\left(\int_0^t \Big(e^{-\lambda(\xi)(t-s)}|\mathcal{F}_{x,v}(\Gamma_0)(s,\xi,\xi(t-s))|\Big)^2\mathrm{d}s\right) \\ &\le \frac{1}{2\lambda(\xi)}\int_0^t \Big|e^{-\lambda(\xi)(t-s)}\mathcal{F}_{x,v}(\Gamma_0)(s,\xi,\xi(t-s))\Big|^2\mathrm{d}s.\end{aligned}$$

The term, $I_3(t,\xi)$ is a Martingale with quadratic variation

$$\int_0^t \sum_{k=1}^{\infty}\Big(e^{-\lambda(\xi)(t-s)}|\mathcal{F}_{x,v}(\Gamma_k)(s,\xi,\xi(t-s))|\Big)^2\mathrm{d}s,$$

where $\Gamma_k(t,x,v) = \phi\,\mathrm{div}_v(\sigma_k f)(t,x,v)$. We conclude by the BDG inequality that

$$\mathbf{E}|I_3|^2(t,\xi) \le \mathbf{E}\int_0^t \sum_{k=1}^{\infty}\Big(e^{-\lambda(\xi)(t-s)}|\mathcal{F}_{x,v}(\Gamma_k)(s,\xi,\xi(t-s))|\Big)^2\mathrm{d}s.$$

and therefore

$$\begin{aligned}\mathbf{E}|\widehat{\rho}^{\phi}(t,\xi)|^2 \le\ & \mathbf{E}|\mathcal{F}_{x,v}(\phi\, f_0)(\xi,\xi t)|^2 \\ &+ \frac{1}{2\lambda(\xi)}\mathbf{E}\int_0^t \Big(e^{-\lambda(\xi)(t-s)}|\mathcal{F}_{x,v}(\Gamma_0)(s,\xi,\xi(t-s))|\Big)^2\mathrm{d}s \\ &+ \mathbf{E}\int_0^t \sum_{k=1}^{\infty}\Big(e^{-\lambda(\xi)(t-s)}|\mathcal{F}_{x,v}(\Gamma_k)(s,\xi,\xi(t-s))|\Big)^2\mathrm{d}s.\end{aligned}$$

The following identities can be readily verified

$$\Gamma_k = \phi\ \mathrm{div}_v(\sigma_k f) = \mathrm{div}_v(\phi\,\sigma_k f) - \nabla\phi\cdot\sigma_k f,$$

and

$$\phi\,\mathcal{L}_\sigma f = \nabla_v^2 : (D_\sigma\,\phi\,f) - 2\,\mathrm{div}_v(D_\sigma \nabla\phi f) + \nabla_v^2\phi : D_\sigma\,f - \mathrm{div}_v(G_\sigma\,\phi\,f) + \nabla\phi\cdot G_\sigma f,$$

where we have denoted for convenience

$$D_\sigma = \sum_{k=1}^{\infty} \sigma_k \otimes \sigma_k \quad \text{and} \quad G_\sigma = \sum_{k=1}^{\infty} \sigma_k \cdot \nabla\sigma_k.$$

This implies

$$\mathcal{F}_{x,v}(\Gamma_k) = i\eta\cdot\mathcal{F}_{x,v}(\phi\,\sigma_k f) - \mathcal{F}_{x,v}(\nabla\phi\cdot\sigma_k f).$$

and

$$\begin{aligned}\mathcal{F}_{x,v}(\phi\,\mathcal{L}_\sigma f) = -\eta\otimes\eta : \mathcal{F}_{x,v}(D_\sigma\,\phi\,f) - 2i\eta\cdot\mathcal{F}_{x,v}(D_\sigma\nabla\phi\,f) + \mathcal{F}_{x,v}(\nabla_v^2\phi : D_\sigma\,f) \\ - i\eta\cdot\mathcal{F}_{x,v}(G_\sigma\,\phi\,f) + \mathcal{F}_{x,v}(\nabla\phi\cdot G_\sigma f).\end{aligned} \tag{B.6}$$

Using that $z^p e^{-\lambda z} \le C_p\lambda^{-p}$, where C_p is constant depending on p, we may bound

$$e^{-\lambda z}|\mathcal{F}_{x,v}(\Gamma_k)(s,\xi,z\,\xi)| \lesssim \lambda^{-1}|\xi|\,|\mathcal{F}_{x,v}(\phi\,\sigma_k f)(s,\xi,z\,\xi)| + |\mathcal{F}_{x,v}(\nabla\phi\cdot\sigma_k f)(s,\xi,z\,\xi)|$$

and using the definition of Γ_0, (B.5), and (B.6) we can bound

$$\begin{aligned}&e^{-\lambda z}|\mathcal{F}_{x,v}(\Gamma_0)(s,\xi,z\,\xi)| \\ &\qquad\lesssim \lambda|\mathcal{F}_{x,v}(\phi\,f)(s,\xi,z\xi)| + |\mathcal{F}_{x,v}(\phi\,g)(s,\xi,z\,\xi)| + \lambda^{-2}|\xi|^2|\mathcal{F}_{x,v}(\phi D_\sigma f)(s,\xi,z\,\xi)| \\ &\qquad\quad + \lambda^{-1}|\xi||\mathcal{F}_{x,v}(D_\sigma\nabla\phi\,f)(s,\xi,z\,\xi)| + |\mathcal{F}_{x,v}(\nabla_v^2\phi\,:D_\sigma f)(s,\xi,z\,\xi)| \\ &\qquad\quad + \lambda^{-1}|\xi||\mathcal{F}_{x,v}(\phi G_\sigma f)(s,\xi,z\,\xi)| + |\mathcal{F}_{x,v}(\nabla\phi\cdot G_\sigma f)(s,\xi,z\,\xi)|.\end{aligned}$$

Integrating $\mathbf{E}|\widehat{\rho}^\phi(t,\xi)|^2$ over $[0,T]$ and using the previous two bounds we get for a.e

$\xi \in \mathbb{R}^d$,

$$
\begin{aligned}
\mathbf{E}\int_0^T |\widehat{\rho}^{\phi}(t,\xi)|^2 \mathrm{d}t \lesssim \mathbf{E}\int_0^T |\mathcal{F}_{x,v}(\phi\, f_0)(\xi,\xi t)|^2 \mathrm{d}t \\
&+ \mathbf{E}\int_0^T\int_0^t \Big\{ \lambda|\mathcal{F}_{x,v}(\phi\, f)(s,\xi,(t-s)\xi)|^2 + \lambda^{-1}|\mathcal{F}_{x,v}(\phi\, g)(s,\xi,(t-s)\,\xi)|^2 \\
&+ \lambda^{-5}|\xi|^4|\mathcal{F}_{x,v}(\phi D_\sigma f)(s,\xi,(t-s)\,\xi)|^2 + \lambda^{-3}|\xi|^2|\mathcal{F}_{x,v}(D_\sigma \nabla\phi\, f)(s,\xi,(t-s)\,\xi)|^2 \\
&+ \lambda^{-1}|\mathcal{F}_{x,v}(\nabla_v^2\phi : D_\sigma f)(s,\xi,(t-s)\,\xi)|^2 + \lambda^{-3}|\xi|^2|\mathcal{F}_{x,v}(\phi G_\sigma f)(s,\xi,(t-s)\,\xi)|^2 \\
&+ \lambda^{-1}|\mathcal{F}_{x,v}(\nabla\phi\cdot G_\sigma f)(s,\xi,(t-s)\,\xi)|^2 + \sum_{k=1}^{\infty}\lambda^{-2}|\xi|^2\,|\mathcal{F}_{x,v}(\phi\,\sigma_k f)(s,\xi,(t-s)\,\xi)|^2 \\
&+ \sum_{k=1}^{\infty}|\mathcal{F}_{x,v}(\nabla\phi\cdot\sigma_k f)(s,\xi,(t-s)\,\xi)|^2 \Big\} \mathrm{d}s\mathrm{d}t.
\end{aligned}
\tag{B.7}
$$

Let's remark that, apart from the initial data, the above estimate is comprised entirely of integrals of the form

$$
\int_0^T\int_0^t |\mathcal{F}_{x,v}(h)(s,\xi,(t-s)\xi)|^2 \mathrm{d}s\mathrm{d}t.
$$

Following the technique in [18], such integrals can be estimated by changing variables to $(z,s) = (|\xi|(t-s),s)$, using Fubini, applying the classical trace theorem on the one dimensional integral in the z variable, and applying Plancharel. We find that for any $\gamma > (d-1)/2$,

$$
\begin{aligned}
\int_0^T\int_0^t |\mathcal{F}_{x,v}(h)(s,\xi,(t-s)\xi)|^2 \mathrm{d}s\mathrm{d}t &\le |\xi|^{-1}\int_0^T\int_{-\infty}^{\infty}\left|\mathcal{F}_{x,v}(h)\left(s,\xi,z\frac{\xi}{|\xi|}\right)\right|^2 \mathrm{d}z\mathrm{d}s \\
&\lesssim |\xi|^{-1}\int_0^T\int_{\mathbb{R}^d}(1+|v|^2)^{\gamma}|\mathcal{F}_x(h)(s,\xi,v)|^2 \mathrm{d}v\mathrm{d}s,
\end{aligned}
$$

and for the initial data,

$$
\int_0^T |\mathcal{F}_{x,v}(\phi\, f_0)(\xi,\xi t)|^2 \mathrm{d}t \lesssim |\xi|^{-1}\int_{\mathbb{R}^d}(1+|v|^2)^{\gamma}|\mathcal{F}_x(h)(\xi,v)|^2 \mathrm{d}v.
$$

Applying the above two estimates term by term to (B.7), we can readily estimate for a.e. ξ,

$$\mathbf{E}\int_0^T |\rho^\phi(t,\xi)|^2 \mathrm{d}t \le C_{\sigma,\phi} M(\xi)\Big(\int_{\mathbb{R}^d} |\widehat{f}_0(\xi,v)|^2 + \mathbf{E}\int_0^T\int_{\mathbb{R}^d} |\widehat{f}(\xi,v,s)|^2 \mathrm{d}v\mathrm{d}s$$
$$+ \mathbf{E}\int_0^T\int_{\mathbb{R}^d} |\widehat{g}(\xi,v,s)|^2 \mathrm{d}v\mathrm{d}s\Big),$$

where

$$M(\xi) = \frac{|\xi|^3}{\lambda(\xi)^5} + \frac{|\xi|}{\lambda(\xi)^3} + \frac{|\xi|}{\lambda(\xi)^2} + \frac{1}{|\xi|\lambda(\xi)} + \frac{\lambda(\xi)}{|\xi|} + \frac{1}{|\xi|},$$

and

$$C_{\sigma,\phi} \lesssim \|(|\phi|^2 + |\nabla\phi|^2 + |\nabla^2\phi|^2)(1+|v|^2)^\gamma\|_{L^\infty_v} \Big\|\sum_{k=1}^\infty (|\sigma_k|^2 + |\sigma\cdot\nabla\sigma_k|)\Big\|_{L^\infty_v}$$

Choosing $\lambda(\xi) = |\xi|^{2/3}$, (really take $\lambda(\xi) = (\epsilon + |\xi|^2)^{1/3}$ and take $\epsilon \to 0$) we conclude that

$$M(\xi) = 3|\xi|^{-1/3} + 2|\xi|^{-1} + |\xi|^{-5/3} =\le 6|\xi|^{-1/3} \text{ if } |\xi| \ge 1.$$

Therefore

$$\mathbf{E}\int_0^T\int_{\mathbb{R}^d} |\xi|^{1/3}|\widehat{\rho}^\phi|^2 \mathbb{1}_{|\xi|\ge 1}\, \mathrm{d}\xi\mathrm{d}s \le C_{\sigma,\phi}\Big(\|f_0\|^2_{L^2_{x,v}} + \mathbf{E}\|f\|^2_{L^2_{t,x,v}} + \mathbf{E}\|g\|^2_{L^2_{t,x,v}}\Big),$$

whereby we have the desired inequality using the Fourier characterization of $H^{1/6}_x$.

The above proof can be extended to weak solutions $f \in L^2_{\omega,t,x,v}$, by first mollifying the equation in (x,v) as in the proof of theorem 7.3.8 and including the commutators with the term g (along with another stochastic integral). The above computation, with the addition of a stochastic integral to the right-hand-side, still apply and the resulting estimates are computed in terms of the $L^2_{\omega,t,x,v}$ norm of the right-hand-side, the commutator contribution will then vanish as the mollification parameter goes to 0. Furthermore we may pass the limit in each term on the right-hand side using the properties of mollifiers. The resulting $H^{1/6}$ estimate on the mollified velocity average can be easily used conclude the associated $H^{1/6}$ estimate on the limiting f by a monotone convergence argument on the Fourier side.

$\square$

Bibliography

[1] B. J. Alder and T. E. Wainwright. “Velocity autocorrelations for hard spheres”. In: *Phys. Rev. Lett.* 18.23 (1967), p. 988.

[2] B. Alder and T. Wainwright. “Decay of the velocity autocorrelation function”. In: *Physical review A* 1.1 (1970), p. 18.

[3] R. Alexandre and C. Villani. “On the Boltzmann equation for long-range interactions”. In: *Communications on pure and applied mathematics* 55.1 (2002), pp. 30–70.

[4] L. Arkeryd. “Loeb solutions of the Boltzmann equation”. In: *Archive for rational mechanics and analysis* 86.1 (1984), pp. 85–97.

[5] L. Arkeryd et al. “Stability for Rayleigh–Benard convective solutions of the Boltzmann equation”. In: *Archive for rational mechanics and analysis* 198.1 (2010), pp. 125–187.

[6] V. I. Arnol’d. *Mathematical methods of classical mechanics*. Vol. 60. Springer Science & Business Media, 2013.

[7] K. Asano. “On the global solutions of the initial boundary value problem for the Boltzmann equation with an external force”. In: *Transp. Theory Stat. Phys.* 16.4-6 (1987), pp. 735–761.

[8] I. Bailleul and M. Gubinelli. “Unbounded rough drivers”. In: *arXiv preprint arXiv:1501.02074* (2015).

[9] C. Bardos, F. Golse, and D. Levermore. “Fluid dynamic limits of kinetic equations. I. Formal derivations”. In: *Journal of Statistical Physics* 63.1-2 (1991), pp. 323–344.

[10] G. Basile, C. Bernardin, and S. Olla. “Thermal Conductivity for a Momentum Conservative Model”. en. In: *Commun. Math. Phys.* 287.1 (Jan. 2009), pp. 67–98.

[11] G. Basile et al. “Thermal conductivity in harmonic lattices with random collisions”. In: *Thermal Transport in Low Dimensions*. Springer, 2016, pp. 215–237.

[12] N. Bellomo et al. “On the initial value problem for the Boltzmann equation with a force term”. In: *Transp. Theory Stat. Phys.* 18.1 (1989), pp. 87–102.

[13] A. Bensoussan and R. Temam. “Equations stochastiques du type Navier-Stokes”. In: *J. Funct. Anal.* 13.2 (Jan. 1973), pp. 195–222.

[14] C. Bernardin and S. Olla. “Non-equilibrium macroscopic dynamics of chains of anharmonic oscillators”. In: *available at www. ceremade. dauphine. fr/...* (2010).

[15] R. B. Bird et al. *Dynamics of polymeric liquids*. Vol. 1. Wiley New York, 1977.

[16] M. Bixon and R. Zwanzig. “Boltzmann-Langevin equation and hydrodynamic fluctuations”. In: *Physical Review* 187.1 (1969), p. 267.

[17] N. N. Bogolyubov and Y. A. Mitropolʹskiĭ. *Asimptotičeskie metody v teorii nelineĭnyh kolebaniĭ*. Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow, 1955, p. 449.

[18] F. Bouchut and L. Desvillettes. “Averaging lemmas without time Fourier transform and application to discretized kinetic equations”. In: *Proceedings of the Royal Society of Edinburgh, Section: A Mathematics* 129.01 (1999), pp. 19–36.

[19] C. L. Bris and P.-L. Lions. “Existence and uniqueness of solutions to Fokker–Planck type equations with irregular coefficients”. In: *Communications in Partial Differential Equations* 33.7 (2008), pp. 1272–1317.

[20] Z. Brzeźniak, F. Flandoli, and M. Maurelli. “Existence and uniqueness for stochastic 2D Euler flows with bounded vorticity”. In: *Archive for Rational Mechanics and Analysis* 221.1 (2016), pp. 107–142.

[21] F. Calogero. “Solution of the One-Dimensional N-Body Problems with Quadratic and/or Inversely Quadratic Pair Potentials”. In: *J. Math. Phys.* 12.3 (1971), pp. 419–436.

[22] P. Catuogno and C. Olivera. “Lp-solutions of the stochastic transport equation”. In: *Random Operators and Stochastic Equations* 21.2 (2013), pp. 125–134.

[23] C. Cercignani. *The Boltzmann Equation and Its Applications:* Applied Mathematical Sciences. Springer New York, 1988.

[24] C. Cercignani, R. Illner, and M. Pulvirenti. *The mathematical theory of dilute gases*. Vol. 106. Springer Science & Business Media, 2013.

[25] N. Champagnat and P.-E. Jabin. “Strong solutions to stochastic differential equations with rough coefficients”. working paper or preprint. Sept. 2015.

[26] A. J. Chorin, O. H. Hald, and R. Kupferman. “Optimal prediction and the Mori–Zwanzig representation of irreversible processes”. In: *Proceedings of the National Academy of Sciences* 97.7 (2000), pp. 2968–2973.

[27] M. Coghi and F. Flandoli. “Propagation of chaos for interacting particles subject to environmental noise”. In: *arXiv preprint arXiv:1403.1981* (2014).

[28] P. V. Coveney and O. Penrose. “On the validity of the Brussels formalism in statistical mechanics”. In: *J. Phys. A Math. Gen.* 25.19 (1992), p. 4947.

[29] E. B. Davies. “Markovian master equations”. en. In: *Commun. Math. Phys.* 39.2 (Jan. 1974), pp. 91–110.

[30] E. B. Davies. “Markovian master equations. II”. en. In: *Math. Ann.* 219.2 (Jan. 1976), pp. 147–158.

[31] E. B. Davies. “Markovian master equations. III”. In: *Annales de l’IHP Probabilités et statistiques*. Vol. 11. 1975, pp. 265–273.

[32] A. Debussche and J. Vovelle. “Invariant measure of scalar first-order conservation laws with stochastic forcing”. In: *Probability Theory and Related Fields* 163.3-4 (2015), pp. 575–611.

[33] F. Delarue, F. Flandoli, and D. Vincenzi. “Noise Prevents Collapse of Vlasov-Poisson Point Charges”. In: *Communications on Pure and Applied Mathematics* 67.10 (2014), pp. 1700–1736.

[34] R. J. Di Perna and P. L. Lions. “Ordinary differential equations, transport theory and Sobolev spaces”. In: *Invent. Math.* 98 (1989), pp. 511–547.

[35] R. J. DiPerna and P.-L. Lions. “Global solutions of Boltzmann’s equation and the entropy inequality”. In: *Arch. Rational Mech. Anal.* 114.1 (1991), pp. 47–55.

[36] R. J. DiPerna and P.-L. Lions. “On the Cauchy problem for Boltzmann equations: global existence and weak stability”. In: *Ann. of Math. (2)* 130.2 (1989), pp. 321–366.

[37] R. Duan, T. Yang, and C. Zhu. “L1 and BV-type stability of the Boltzmann equation with external forces”. In: *Journal of Differential Equations* 227.1 (2006), pp. 1–28.

[38] J. Dunkel and S. Hilbert. “Consistent thermostatistics forbids negative absolute temperatures”. en. In: *Nat. Phys.* 10.1 (Aug. 2013), pp. 67–72.

[39] W. E and Y. G. Sinai. “Recent results on mathematical and statistical hydrodynamics”. In: *Russian Mathematical Surveys* 55.4 (2000), p. 635.

[40] P. Ehrenfest and T. Ehrenfest. “The Conceptual Foundations of the Statistical Approach in Mechanics”. In: (1959).

[41] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*. Vol. 194. Graduate Texts in Mathematics. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt. Springer-Verlag, New York, 2000, pp. xxii+586.

[42] P. Español. “Dissipative particle dynamics with energy conservation”. In: *EPL* 40.6 (1997), p. 631.

[43] P. Español and P. Warren. “Statistical Mechanics of Dissipative Particle Dynamics”. In: *EPL* 30.4 (1995), p. 191.

[44] P. Español. “Einstein-Helfand form for transport coefficients from coarse-grained descriptions”. en. In: *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 80.6 Pt 1 (Dec. 2009), p. 061113.

[45] P. Español, J. G. Anero, and I. Zúñiga. "Microscopic derivation of discrete hydrodynamics". en. In: *J. Chem. Phys.* 131.24 (28 12 2009), p. 244117.

[46] P. Español and F. Vázquez. "Coarse graining from coarse-grained descriptions". en. In: *Philos. Trans. A Math. Phys. Eng. Sci.* 360.1792 (15 03 2002), pp. 383–394.

[47] R. Esposito, R. Marra, and J. Lebowitz. "Solutions to the Boltzmann equation in the Boussinesq regime". In: *Journal of statistical physics* 90.5-6 (1998), pp. 1129–1178.

[48] E. Fedrizzi et al. "Regularity of Stochastic Kinetic Equations". In: (Mar. 2016). arXiv: 1606.01088 [math.PR].

[49] A. Figalli. "Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients". In: *J. Funct. Anal.* 254.1 (Jan. 2008), pp. 109–153.

[50] F. Flandoli, M. Gubinelli, and E. Priola. "Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift". In: 134.4 (June 2010), pp. 405–422.

[51] F. Flandoli, M. Gubinelli, and E. Priola. "Well-posedness of the transport equation by stochastic perturbation". In: *Invent. Math.* 180.1 (24 11 2009), pp. 1–53.

[52] F. Flandoli. "The interaction between noise and transport mechanisms in PDEs". In: *Milan Journal of Mathematics* 79.2 (2011), pp. 543–560.

[53] F. Flandoli, M. Gubinelli, and E. Priola. "Well-posedness of the transport equation by stochastic perturbation". In: *Inventiones mathematicae* 180.1 (2010), pp. 1–53.

[54] E. G. Flekkøy and P. V. Coveney. "From Molecular Dynamics to Dissipative Particle Dynamics". In: *Phys. Rev. Lett.* 83.9 (30 08 1999), pp. 1775–1778.

[55] E. G. Flekkoy, P. V. Coveney, and G. De Fabritiis. "Foundations of Dissipative Particle Dynamics". In: (Nov. 2000). arXiv: cond-mat/0002174 [cond-mat.soft].

[56] R. F. Fox and G. E. Uhlenbeck. "Contributions to nonequilibrium thermodynamics. II. Fluctuation theory for the Boltzmann equation". In: *Physics of Fluids (1958-1988)* 13.12 (1970), pp. 2881–2890.

[57] M. I. Freidlin and A. D. Wentzell. "Diffusion Processes on Graphs and the Averaging Principle". In: *Ann. Probab.* 21.4 (1993), pp. 2215–2245.

[58] M. I. Freidlin. "The averaging principle and theorems on large deviations". In: *Russian mathematical surveys* 33.5 (1978), pp. 117–176.

[59] J. Fritz, T. Funaki, and J. L. Lebowitz. "Stationary states of random Hamiltonian systems". en. In: *Probab. Theory Related Fields* 99.2 (Jan. 1994), pp. 211–236.

[60] I. Gallagher, L. Saint-Raymond, and B. Texier. *From Newton to Boltzmann: hard spheres and short-range potentials*. European mathematical society, 2013.

[61] B. Gess and P. E. Souganidis. “Long-time behavior, invariant measures and regularizing effects for stochastic scalar conservation laws”. In: *arXiv preprint arXiv:1411.3939* (2014).

[62] J. Gibbs. “Elementary principles in statistical physics”. In: *The Collected Works of JW Gibbs (Yale University, New Haven, CT, 1957)* 2 (1902).

[63] R. A. Gingold and J. J. Monaghan. “Smoothed particle hydrodynamics: theory and application to non-spherical stars”. In: *Monthly notices of the royal astronomical society* 181.3 (1977), pp. 375–389.

[64] F. Golse, B. Perthame, and R. Sentis. “Un résultat de compacité pour les équations de transport et application au calcul de la limite de la valeur propre principale d’un opérateur de transport”. In: *C. R. Acad. Sci. Paris Sér. I Math.* 301.7 (1985), pp. 341–344.

[65] F. Golse and L. Saint-Raymond. “Hydrodynamic Limits for the Boltzmann Equation”. In: *RIVISTA DI MATEMATICA DELLA UNIVERSITÀ DI PARMA* 4** (31 12 2005), pp. 1–144.

[66] F. Golse and L. Saint-Raymond. “Velocity averaging in L1 for the transport equation”. In: *Comptes Rendus Mathematique* 334.7 (2002), pp. 557–562.

[67] F. Golse et al. “Regularity of the moments of the solution of a Transport Equation”. In: *J. Funct. Anal.* 76.1 (Jan. 1988), pp. 110–125.

[68] A. N. Gorban. “Basic types of coarse-graining”. In: *Model Reduction and Coarse-Graining Approaches for Multiscale Phenomena*. Springer, 2006, pp. 117–176.

[69] H. Grabert. “Projection operator techniques in nonequilibrium statistical mechanics”. In: (1982).

[70] M. Z. Guo, G. C. Papanicolaou, and S. R. S. Varadhan. “Nonlinear diffusion limit for a system with nearest neighbor interactions”. In: *Commun. Math. Phys.* 118.1 (1988), pp. 31–59.

[71] C. Hijón et al. “Mori–Zwanzig formalism as a practical computational tool”. In: *Faraday Discuss.* 144 (2010), pp. 301–322.

[72] P. J. Hoogerbrugge and J. M. V. A. Koelman. “Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics”. In: *EPL* 19.3 (1992), p. 155.

[73] P.-E. Jabin and L. Vega. “A real space method for averaging lemmas”. In: *Journal de mathématiques pures et appliquées* 83.11 (2004), pp. 1309–1351.

[74] A. Jakubowski. “The almost sure Skorokhod representation for subsequences in nonmetric spaces”. In: *Teor. Veroyatnost. i Primenen.* 42.1 (1997), pp. 209–216.

[75] M. Jara, T. Komorowski, and S. Olla. "Superdiffusion of energy in a chain of harmonic oscillators with noise". In: *Communications in Mathematical Physics* 339.2 (2015), pp. 407–453.

[76] E. Johnsen. "John G. Kemey, Arthur Schleifer, Jr., J. Laurie Snell and Gerald L. Thornpson: Finite Mathematics with Business Applications. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1962. 482 s." In: *Ledelse og Erhvervsøkonomi* 27 ().

[77] L. P. Kadanoff. "Scaling laws for Ising models near Tc". In: *From Order To Chaos: Essays: Critical, Chaotic and Otherwise*. World Scientific, 1993, pp. 165–174.

[78] J. M. V. A. Koelman and P. J. Hoogerbrugge. "Dynamic Simulations of Hard-Sphere Suspensions Under Steady Shear". In: *EPL* 21.3 (1993), p. 363.

[79] N. V. Krylov and M. Röckner. "Strong solutions of stochastic equations with singular time dependent drift". en. In: *Probab. Theory Related Fields* 131.2 (25 05 2004), pp. 154–196.

[80] H. Kunita. *Stochastic flows and stochastic differential equations*. Vol. 24. Cambridge university press, 1997.

[81] H. Kunita. "Stochastic flows and stochastic partial differential equations". In: *Proceedings of the International Congress of Mathematicians*. Vol. 1. 1986, p. 2.

[82] W. Lamb, I. Murdoch, and J. Stewart. "On an operator identity central to projection operator methodology". In: *Physica A: Statistical Mechanics and its Applications* 298.1–2 (Jan. 2001), pp. 121–139.

[83] L. Landau and E. Lifshitz. "Fluid mechanics, 1959". In: *Course of Theoretical Physics* (1959).

[84] C. Le Bris and P.-L. Lions. "Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients". In: *Comm. Partial Differential Equations* 33.7-9 (2008), pp. 1272–1317.

[85] M. Lesieur, O. Métais, and P. Comte. *Large-eddy simulations of turbulence*. Cambridge University Press, 2005.

[86] P. L. Lions. "Compactness in Boltzmann's equation via Fourier integral operators and applications. III". In: *Journal of Mathematics of Kyoto University* 34.3 (1994), pp. 539–584.

[87] P.-L. Lions, B. Perthame, and P. E. Souganidis. "Stochastic averaging lemmas for kinetic equations". eng. In: *Séminaire Laurent Schwartz – EDP et applications* 2011-2012 (2011-2012), pp. 1–17.

[88] P.-L. Lions. "Compactness in Boltzmann's equation via Fourier integral operators and applications. I, II". In: *J. Math. Kyoto Univ.* 34.2 (1994), pp. 391–427, 429–461.

[89] P.-L. Lions. "Compactness in Boltzmann's equation via Fourier integral operators and applications. III". In: *J. Math. Kyoto Univ.* 34.3 (1994), pp. 539–584.

[90] C. Liverani and S. Olla. "Ergodicity in infinite Hamiltonian systems with conservative noise". en. In: *Probab. Theory Related Fields* 106.3 (Jan. 1996), pp. 401–445.

[91] A. D. Mackie, J. B. Avalos, and V. Navas. "Dissipative particle dynamics with energy conservation: Modelling of heat flow". en. In: *Phys. Chem. Chem. Phys.* 1.9 (1999), pp. 2039–2049.

[92] A. J. Majda, I. Timofeyev, and Vanden Eijnden E. "Models for stochastic climate prediction". en. In: *Proc. Natl. Acad. Sci. U. S. A.* 96.26 (21 12 1999), pp. 14687–14691.

[93] H. Martina. "A Bhatnagar-Gross-Krook approximation to stochastic scalar conservation laws". In: *Annales de L Institut Henri Poincare-Probabilites Et Statistiques*. Vol. 51. 2015, p. 4.

[94] E. Montroll. *Fluctuation phenomena*. Vol. 7. Elsevier, 2012.

[95] H. Mori. "Transport, collective motion, and Brownian motion". In: *Progress of theoretical physics* 33.3 (1965), pp. 423–455.

[96] J. Moser. "Three integrable Hamiltonian systems connected with isospectral deformations". In: *Adv. Math.* 16.2 (Jan. 1975), pp. 197–220.

[97] K. Musial. "Pettis integral". In: *Handbook of measure theory* 1 (2002), pp. 531–586.

[98] T. Naitoh et al. "Velocity correlations in a one-dimensional lattice gas: Theory and simulations". en. In: *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* 47.6 (June 1993), pp. 4098–4103.

[99] S. Nakajima. "On quantum theory of transport phenomena: steady diffusion". In: *Progr. Theoret. Phys.* 20 (1958), pp. 948–959.

[100] W. Neves and C. Olivera. "Wellposedness for stochastic continuity equations with Ladyzhenskaya–Prodi–Serrin condition". In: *Nonlinear Differential Equations and Applications NoDEA* 22.5 (2015), pp. 1247–1258.

[101] S. Olla, S. R. S. Varadhan, and H.-T. Yau. "Hydrodynamical limit for a Hamiltonian system with weak noise". In: *Comm. Math. Phys.* 155.3 (1993), pp. 523–560.

[102] H. C. Öttinger. *Beyond equilibrium thermodynamics*. John Wiley & Sons, 2005.

[103] H. C. Öttinger. "General projection operator formalism for the dynamics and thermodynamics of complex fluids". In: *Phys. Rev. E* 57.2 (Jan. 1998), pp. 1416–1420.

[104] É. Pardoux. "Equations aux dérivées partielles stochastiques non lineaires monotones: Etude de solutions fortes de type Ito". PhD thesis. 1975.

[105] I. Prigogine, C. George, and F. Henin. "Dynamical and statistical descriptions of N-body systems". In: *Physica* 45.3 (1969), pp. 418–434.

[106] S. Punshon-Smith and S. Smith. *On the Boltzmann Equation with Stochastic Kinetic Transport: Global Existence of Renormalized Martingale Solutions.* submitted. June 2016. arXiv: 1606.07451 [math.PR].

[107] F. Rezakhanlou. "Regular flows for diffusions with rough drifts". In: *arXiv preprint arXiv:1405.5856* (2014).

[108] M. Ripoll, P. Español, and M. H. Ernst. "Dissipative Particle Dynamics with Energy Conservation: Heat Conduction". In: *Int. J. Mod. Phys. C* 09.08 (1998), pp. 1329–1338.

[109] H. Spohn. "Fluctuation theory for the Boltzmann equation". In: *Nonequilibrium Phenomena I: The Boltzmann Equation* 1 (1983), pp. 225–251.

[110] H. Spohn. "Fluctuations around the Boltzman equation". In: *Journal of Statistical Physics* 26.2 (1981), pp. 285–305.

[111] H. Spohn. *Large scale dynamics of interacting particles.* Texts and Monographs in Physics. Springer, 1991.

[112] M. Toda. "Vibration of a Chain with Nonlinear Interaction". In: *J. Phys. Soc. Jpn.* 22.2 (1967), pp. 431–436.

[113] M. Toda et al. *Statistical Physics II: Nonequilibrium Statistical Mechanics.* Vol. 30. Springer Science & Business Media, 1992.

[114] H. Ueyama. "The stochastic Boltzmann equation and hydrodynamic fluctuations". In: *Journal of Statistical Physics* 22.1 (1980), pp. 1–26.

[115] S. Ukai, T. Yang, and H. Zhao. "Global solutions to the Boltzmann equation with external forces". In: *Analysis and Applications* 3.02 (2005), pp. 157–193.

[116] A. W. Van Der Vaart and J. A. Wellner. *Weak Convergence.* Springer, 1996.

[117] A. J. Veretennikov and N. V. Krylov. "Explicit formulae for the solutions of stochastic equations". In: *Mat. Sb. (N.S.)* 100(142).2 (1976), pp. 266–284, 336.

[118] A. Y. Veretennikov. "Criteria for the existence of a strong solution of a stochastic equation". In: *Teor. Veroyatnost. i Primenen.* 27.3 (1982), pp. 417–424.

[119] C. Villani. "A review of mathematical topics in collisional kinetic theory". In: *Handbook of mathematical fluid dynamics* 1 (2002), pp. 71–305.

[120] E. Weinan and B. Engquist. "The Heterognous Multiscale Methods". In: *Commun. Math. Sci.* 1.1 (Mar. 2003), pp. 87–132.

[121] K. G. Wilson. "The renormalization group: Critical phenomena and the Kondo problem". In: *Rev. Mod. Phys.* 47.4 (Jan. 1975), pp. 773–840.

[122] X. Zhang. "Strong solutions of SDES with singular drift and Sobolev diffusion coefficients". In: *Stochastic Process. Appl.* 115.11 (Nov. 2005), pp. 1805–1818.

[123] R. Zwanzig. "Ensemble Method in the Theory of Irreversibility". In: *J. Chem. Phys.* 33.5 (Jan. 1960), pp. 1338–1341.

[124] R. Zwanzig. "Problems in nonlinear transport theory". In: *Systems Far from Equilibrium: Sitges Conference on Statistical Mechanics, June 1980, Sitges, Barcelona/Spain.* Ed. by L. Garrido. Berlin, Heidelberg: Springer Berlin Heidelberg, 1980, pp. 198–225.