ABSTRACT

Title of dissertation: A FOKKER-PLANCK STUDY MOTIVATED BY A PROBLEM IN FLUID-PARTICLE INTERACTIONS:

Ioannis Markou, Doctor of Philosophy, 2014

Dissertation directed by: Professor C. David Levermore
Department of Mathematics

This dissertation is a study of problems that relate to a Fokker-Planck (Klein-Kramers) equation with hypoelliptic structure. The equation describes the statistics of motion of an ensemble of particles in a viscous fluid that follows the Stokes' equations of fluid motion. The significance in this problem is that it relates to a variety of phenomena besides its obvious connection to the study of macromolecular chains that are composed by particle "units" in creeping flows. Such phenomena range from Kramers escape probability (for a particle trapped in a potential well), to stellar dynamics. The problem can also be seen as a simplified version of the Vlasov-Poisson-Fokker-Planck system that mainly describes electrostatic models in plasma physics and gravitational forces between galaxies.

Well-posedeness of the equation has been studied by many authors, including the case of irregular coefficients (Lions-Le Bris). The study of Sobolev regularity is interesting in its own right and can be performed with fairly elementary tools (Hérau,Villani,...). We are interested here with short time estimates and with how

smoothing proceeds in time. Different types of Lyapunov functionals can be constructed depending on the type of initial data to show regularization. Of particular interest is a recent technique developed by C.Villani that builds upon a system of differential inequalities and is being implemented here for the slightly more involved case of non constant friction. The question of asymptotic convergence to a stationary state is also discussed, with techniques that are similar to certain extend to the ones used in regularization but which in general involve more computations.

Finally, we examine the hydrodynamic (zero mass) limit of the parametrized version of the Fokker-Planck equation. We discuss two different approaches of hydrodynamic convergence. The first uses weak compactness principles of extracting subsequences that are shown to converge to a solution of the limit problem, and works with initial data in weighted L^2 setting. The second is based on the study of relative entropy, gives L^1 convergence to a solution of the limit problem, and uses entropic initial data.

A FOKKER-PLANCK STUDY MOTIVATED BY A PROBLEM IN FLUID-PARTICLE INTERACTIONS

by

Ioannis Markou

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
2014

Advisory Committee:
Professor C. David Levermore, Chair/Advisor
Professor Pierre-Emmanuel Jabin
Professor Manoussos Grillakis
Professor Matei Machedon
Professor Theodore L. Einstein

Dedication

To my family for their constant support over the years. To everyone who was present during times of sorrow, fear and doubt.

Acknowledgments

It has always been a difficult task writing a letter to acknowledge the importance that people have played in any type of personal achievement. The fear of forgetting important figures as well as the temptation of going to the other extreme with superfluous expressions of gratitude is a possibility. Yet, I've tried as much as possible to keep a balance between the two.

I would like to begin with my first academic mentor, Athanasios Tzavaras. Thanos was definitely one of the biggest motivating factors for me coming to US, and definitely the biggest one for me landing here at UMCP. I owe him my interest on the subject of this dissertation which began with questions we had way back. Unfortunately, our paths had to part because of unforseen family matters five years ago, but his mark on this work is priceless.

Special thanks to Pierre-Emmanuel Jabin for the fruitful discussions we had over time at UMCP. His professionalism and remarks were always very insightful, no nonsense, and as a result extremely helpful. Without his contribution, Theorem 2 would not be possible, as it was born out of one of his ideas.

Manoussos Grillakis and Konstantina Trivisa share my gratitude for being willing to read, hear, and comment on part of my research. Their feedback was very helpful and morally supportive. Additional thanks to Ted Einstein and Matei Machedon for their kindness in accepting the role of being committee members, and their willingness in reviewing this manuscript.

Cedric Villani was generous enough to give me feedback on questions related to part of his work. He was also kind to suggest a possible future path. Besides this fact, this dissertation relies heavily on many of his research ideas and contributions. I couldn't possibly forget to mention my thankfulness to him.

Many thanks to the staff of UMCP for their constant help in academic and graduate support related issues. Very special mention to Hidalgo Haydee (I will never forget her kind e-mails prior to my arriving in College Park), Regalado Celeste, Welton Sharon and Schildknecht William. The staff is literally the backbone of this department.

My friends Yulia Dobrosotskaya and Arseny Zakharov for their companion, help in times of trouble, tech support, academic advice, and the honor of letting me babysit their cats while they were away. They are my family here.

Last, but most certainly not least, I want to express my gratitude to my academic advisor David Levermore. David has a very distinctive ability of asking the questions that made me start searching for answers in much deeper places than I previously thought I could dig. His guidance was never patronizing, and I have to thank him for the freedom he gave me to pick the research questions more close to my interests. Without his continuing support, his contributions, and his ability to make solid plans for me, this work would never be finished.

Table of Contents

Chapter 1: Introduction

This dissertation studies relationships between different math models of systems in which a large number of macromolecules are sparsely immersed within a far larger number of micromolecules. Examples of such systems include dilute solutions of polymers such as arise in many industrial settings (see [3, 18, 19, 38, 40, 41]). We model the macromolecules (or more precisely the monomer parts they are comprised of) as idealized particles (spheres) whose interactions are solely mediated by interactions with the micromolecules. We model the micromolecules as an incompressible fluid governed by Stokes flow. The interactions of these idealized particles with the fluid are modeled by admissible boundary conditions, Brownian noise and a damping term. This leads to a high-dimensional Markov process, whose probability density function is governed by a Fokker-Planck equation in N particle phase-space. We introduce a re-scaling that separates scales, in a way that the now fast scale leads the system to relaxation to local Gibbs states. The main goal is to present a rigorous derivation of the macroscopic density equation, through the study of the hydrodynamic limit as we let the particle mass $m \to 0$.

The importance of the these models lies in the huge diversity of industrial applications of polymeric solutions and materials. Such applications include the use of polymers to thicken and raise the viscosity of industrial products. Examples may range from the thickening of motor oils, which is extremely important in high temperatures (by diblock copolymers), to the reduction of turbulent flows in water (e.g. by polyethylene oxide) and proteins used to remove substances from food. In general, applications involve fields as diverse as biology, medicine, food processing, oil industry, pharmacology and many others.

Despite the simplicity of fluid/particle models, complications arise that cannot be ignored. The first and foremost is that as a model it is extremely special and therefore it is not ideal for describing every single individual polymer architecture. Assuming that monomers can be modeled as spherical particles (this is already a big assumption), the specifics of many polymer structures (how monomers bond together) requires in many cases a reduction to a simpler model. For instance, many macromolecules can be modeled quite satisfactory by rigid rod models, worm-like polymers, thread or tube models etc. Another important problem is related to particle/fluid interactions. In order to have a physically meaningful model, we need to be specific on how to couple the system of particles/fluid so that we get a well-posed mathematical problem. As we are going to show soon, in our model, particles and fluid interact via a set of boundary conditions (the damping force is a result of these BCs). This will allow us to forget the specifics of fluid motion and focus on the kinetic description of the particle system alone, since the equations of particle motion form a closed system now. On the upside, there are also advantages. The particles in the model system imitate the monomer constituents that are the building blocks of the macromolecule. These monomers (particles in our model) form covalent bonds and interact with each other with various forces, e.g. bond forces, electrostatic and van der Waals. In many cases, we can capture the nature of these interactions with the help of a potential $\mathcal{U}(x)$. Regardless of all the complexities, there are still many important cases where this basic model stands as a good approximation.

As we put together all the important parts that follow, we begin with a mathematical description of the particle system and how we go from a microscopic to a kinetic (mesoscopic) level of description. We then derive a kinetic many particle Fokker-Planck equation which will be studied in detail. This equation, re-scaled, has the important property that it reaches local equilibration very fast, thus making the study of the equation that corresponds to the Gibbs states more relevant for small masses or larger time scales. This motivates us to discuss the specifics of the macroscopic (hydrodynamic) limit and deduce a Smoluchowski type of equation that describes the evolution of the particle cloud at a macroscopic level. Doing so, we shall not ignore the importance of an independent study of the regularity of the Fokker-Planck equation.

1.1 From Particle Systems to a Kinetic Formulation

The starting point is a system of N identical, spherical particles of mass m emersed in a fluid governed by Stokes flow. The particles are assumed to be confined by an external potential field $\mathcal{U}(x)$. We shall make the assumption that the particles are not allowed to slip in the fluid medium, and therefore consider the fluid velocity constant on the surface of the particle and equal to the particle's velocity. This assumption is consistent in the sense that it leads to a well-posed fluid/particle mathematical model. Because of the linearity of the Stokes system and the conditions on the boundary interface between fluid/particles, the hydrodynamic (fluid) forces that particles experience depend linearly on the particle velocities and this dependence is manifested with the use of a $3N \times 3N$ friction tensor $G(x)$. The friction is a function of the configuration of the center of particles. As mentioned already, the equations of particle motion can be studied independently from the Stokes flow. The last major assumption that we make is to include a Brownian forcing term in the equations of particle motion to account for the presence of random collisions of small molecules in the fluid with the large particles.

To use mathematical language, the kinematic equations of particle motion are described by the phase-space vector $(x, v) \in \mathbb{R}_x^{3N} \times \mathbb{R}_v^{3N}$, where x describes the position of the center of mass of the N spheres and v their velocities. The resulting system of $6N$ equations of motion is

$$\begin{aligned} \frac{dx}{dt} &= v, \\ m\frac{dv}{dt} &= -G(x)v - \nabla\mathcal{U}(x) + \sqrt{2}G^{1/2}(x)\frac{dW}{dt}, \end{aligned} \tag{1.1}$$

where $W(t)$ is the standard Brownian vector in $\mathbb{R}^{3N}$.

Before we continue with the study of the system of equations, we shall describe how the friction $G(x)$ can be computed or at least approximated. Let us mention for the time being that $G(x)$ is the only part of the equations of motion where information about the interaction of particles is contained. In fact the friction tensor contains all the information for the interaction between spheres through the fluid, making the equations of fluid motion obsolete. The case of a diagonal constant friction $G(x) = \gamma\mathrm{I}$ (for $\gamma > 0$) corresponds to particles that move freely in the fluid without interacting with each other. It is a case of special interest because it is a good approximation for dilute regimes of particles which will be described in Chapter 2. The case of diagonal, non constant, friction $G(x) = \gamma(x)\mathrm{I}$ is also of special interest, since it helps simplify calculations in the analysis.

Unfortunately, the exact computation of the friction would require solution of the Stokes problem for every possible profile of particles in the fluid. We can nonetheless approximate it by various techniques making the business of constructing meaningful approximations an important topic on its own. We are going to present the two approximations that are most prominent and the problems related to them.

The first approximation encountered is the Stokeslet approximation. The Stokeslet (Oseen tensor) is the Green's function to the Stokes problem associated with point particles that experience a singular force. This would actually be a good first order approximation of the hydrodynamic mobility tensor (inverse of friction tensor) if not for its failure to be nonnegative. Nonnegativity of $G(x)$ is a property essential for constructing meaningful energies for the particle system and thus fails an important test.

The construction of a nonnegative tensor can be achieved via a variational (energy) formulation, first developed in [65] & independently as a perturbation expansion in [73]. This approximation takes into account particle size and can be seen as the second order (correction to Oseen tensor) term of an expansion with respect to a parameter which is typical inter-particle distance over particle radius. A closer study of the Rotne-Prager-Yamakawa tensor will reveal, that although the tensor is nonnegative by construction, we cannot avoid degeneracy of $G(x)$ for certain configurations. These configurations as we will examine in the case of a two particle system happen when the particles coincide.

We now go back to the kinetic equations of motion and consider the problem for unit mass (non-parametrized version). The evolution of the particle system can be described statistically with the use of a probability density $f(t,x,v)$. The equation for f is the forward N-particle Fokker-Planck equation $\partial_t f + Lf = 0$, where L is the operator

$$Lf = v \cdot \nabla_x f - \nabla \mathcal{U}(x) \cdot \nabla_v f - \nabla_v \cdot (G(x)(\nabla_v f + vf)).$$

The Cauchy problem

$$\partial_t f + Lf = 0, \qquad f(0,x,v) = f_0(x,v) \tag{1.2}$$

will be the object of study for a large part of this dissertation. The ideal is for initial data f_0 to lie in $L^1(\mathbb{R}^{3N,3N}_{x,v})$. Unfortunately, for many results it will be required that initial data belong to the less natural functional space $L^2_{\mathcal{M}_{eq}} = \mathcal{M}_{eq} L^2(\mathcal{M}_{eq}\, dx\, dv)$, where $\mathcal{M}_{eq}(x,v) = e^{-\mathcal{U}(x)} e^{-\frac{v^2}{2}}/Z$ (with $Z = \iint e^{-\mathcal{U}(x)} e^{-\frac{v^2}{2}}\, dv\, dx$) is the global stationary state solution.

Besides the motivation for the particle/fluid problem given in the beginning of this introduction, equation (1.2) is interesting by itself as it relates to a wider class of physical phenomena.

It is worth mentioning, for example, that it serves as a model case for the Vlasov-Poisson-Fokker-Planck system described by the couple of equations

$$\partial_t f + v \cdot \nabla_x f - \nabla \mathcal{U}(t,x) \cdot \nabla_v f = \nabla_v \cdot (\nabla_v f + v f),$$

$$\triangle \mathcal{U}(t,x) = \int f\, dv \qquad x, v \in \mathbb{R}^d,$$

in higher dimensions $d > 3$. The theory of this system of equations is still in relatively early stage, with results on well-posedness (see [5,9,13,69] ...), regularity (see [5,6,13,62,70] ...), and hydrodynamic limit (see [24,25,61] ...), to name only a few. Other mathematical problems with similar structure are the Vlasov-F-P equation, the Vlasov-Navier-Stokes system (see e.g. [26,27]) etc.

1.2 Macroscopic Limit

Hydrodynamic limits are extremely important in connecting between two different levels of description (a kinetic or in some cases microscopic and a macroscopic one). Historically, this goes back to the work of Maxwell, Boltzmann and Hilbert (famous sixth problem, Hilbert expansion etc) in the foundations of the kinetic theory of gases. Typically, macroscopic variables and the conservation laws they obey are far more helpful in describing observable quantities of a physical system than the rapidly changing phase-space densities. This makes their study necessary not just for theoretical but also experimental reasons.

Going back to the equations of particle dynamics with mass m (1.1), we want to study the behavior of solutions of the new Cauchy problem as $m \to 0$. We shall first consider the parabolic scaling $\epsilon = \sqrt{m}$, $\epsilon v \to v$, $x \to x$, which leads to the parametrized FP equation

$$\partial_t f_\epsilon + L_\epsilon f_\epsilon = 0, \qquad f_\epsilon(0,x,v) = f_{0,\epsilon}(x,v), \tag{1.3}$$

with

$$L_\epsilon = \frac{1}{\epsilon}\left(v \cdot \nabla_x f_\epsilon - \nabla \mathcal{U}(x) \cdot \nabla_v f_\epsilon\right) - \frac{1}{\epsilon^2}\nabla_v \cdot \left(G(x)(\nabla_v f_\epsilon + v f_\epsilon)\right).$$

Defining the hydrodynamic density and flux vector by

$$\rho_\epsilon = \int f_\epsilon \, dv, \qquad J_\epsilon = \frac{1}{\epsilon}\int v f_\epsilon \, dv$$

and taking the limit as $\epsilon \to 0$ we should at least formally obtain the system of equations

$$\begin{aligned} &\partial_t \rho + \nabla_x \cdot J = 0, \\ &J = -G^{-1}(x)(\nabla \rho + \rho \nabla \mathcal{U}(x)). \end{aligned} \tag{1.4}$$

In the above system ρ, J are the limits of ρ_ϵ and J_ϵ with convergence understood in some appropriate setting. At the same time, the formal limit for f_ϵ is $f_\epsilon \to \rho(t,x)\mathcal{M}(v)$, where $\mathcal{M}(v)$ is the canonical Maxwellian $\mathcal{M}(v) = e^{-\frac{v^2}{2}}/(2\pi)^{n/2}$ (where $n = 3N$). We give two results proving the rigorous limit based on different a priori estimates to the solutions of (1.3).

The **first** result establishes weak convergence for the hydrodynamic variable $\rho_\epsilon(t,x)$ based on weak compactness arguments. The proof is actually quite elementary and can be outlined here. We assume that operator L_ϵ generates a continuous semigroup and that a solution $f_\epsilon(t,x,v)$ to the equation (1.3), with initial data $f_\epsilon(0,x,v)$, is $f_\epsilon(t,x,v) = e^{-tL_\epsilon} f_\epsilon(0,x,v)$. We shall also make the assumption of finite initial energy in $L^2_{\mathcal{M}_{eq}}$, i.e. $\|f_\epsilon(0,x,v)\|_{L^2_{\mathcal{M}_{eq}}} < C, \quad \forall \epsilon > 0, C > 0$. Next, we decompose $f_\epsilon(t,x,v)$ into a local equilibrium state $\mathcal{M}(v)\rho_\epsilon(t,x)$, and a deviation $\mathcal{M}(v)\tilde{g}_\epsilon(t,x,v)$. With the help of the a priori estimate, we can extract convergent subsequences for $\rho_\epsilon(t,x)$, $\tilde{g}_\epsilon(t,x,v)$ and $\frac{1}{\epsilon}G^{1/2}(x)\nabla_v \tilde{g}_\epsilon(t,x,v)$. By a simple application of Arzela-Ascoli lemma, one can show that ρ_ϵ is compact in $C([0,T], \text{w}-L^2(\mathbb{R}^n_x))$, for any $T > 0$. Next, we write an evolution equation for $\tilde{g}_\epsilon$ (in distributional sense) and pass to the limit in $\epsilon \to 0$. To do this, since we are dealing with a weak formulation, one has to find the order in ϵ of each integral in the formulation and ignore lower order terms in ϵ. The final step is to couple the limiting equation for $\tilde{g}_\epsilon$ with the limiting equation for ρ_ϵ. This coupling results in the Smoluchowski limit equation. The exact statement of the theorem is:

Theorem 1. *Let ρ_ϵ be the hydrodynamic density defined above, where f_ϵ is a mild-weak solution of* (1.3) *with bounded initial energy $\|f_\epsilon(0,.,.)\|_{L^2_{\mathcal{M}_{eq}}(\mathbb{R}^{n,n}_{x,v})} < \infty$ (uniformly in ϵ). Assume that the divergence-free tensor $G^{-1}(x)$ is non-degenerate a.e., with $G^{-1}(x) \in L^1_{loc}(\mathbb{R}^n_x)$. In the limit $\epsilon \to 0$,*

$$\rho_\epsilon \rightharpoonup \rho \qquad in \quad C([0,T], w - L^2(dx)),$$

where $\rho(t,x)$ satisfies the Smoluchowski equation

$$\partial_t \rho = \nabla_x \cdot \left(G^{-1}(x) \left(\nabla_x \rho + \nabla \mathcal{U}(x) \rho \right) \right) \quad in \quad C([0,T], \mathcal{D}'(\mathbb{R}^n_x)).$$

The **second** method was actually born after some fruitful discussions with P-E Jabin [35], who indicated to me how relative entropy can be used to establish hydrodynamic limits. His experience on this method is contained among other sources in his own work, e.g. see [27].

Since the relative entropy functional $H(f|g) = \iint f \log \frac{f}{g} \, dv \, dx$ between two probability densities f, g is a measure of the distance between them (see Csiszár-Kullback-Pinsker inequality), by finding $\lim_{\epsilon \to 0} H(f_\epsilon | \rho \mathcal{M})$ we can actually control the square of the L^1 distance between f_ϵ and $\rho\mathcal{M}$ in the limit $\epsilon \to 0$. We can show for our problem that the dissipation of $H(f_\epsilon|\rho\mathcal{M})$ contains a non negative part and residual terms. It will be our purpose to show rigorously that the residual terms vanish as $\epsilon \to 0$, and this is exactly where we have to be specific on the regularity required for the solutions $f_\epsilon(t,x,v)$ of (1.3). Once we show that in the limit the relative entropy is strictly dissipative, it will be enough to consider initial data "prepared" in a way s.t. $H(f_\epsilon(0,.,.)|\rho(0,.)\mathcal{M}) \to 0$ (as $\epsilon \to 0$) so that it is implied that $H(f_\epsilon(t,.,.)|\rho(t,.)\mathcal{M}) \to 0$ for t in some finite interval $[0,T]$, for any $T > 0$.

We now give details on the requirements for the statement to be proven. The first assumption is the existence of a unique stationary state $\mathcal{M}_{eq}(x,v)$, and a potential $\mathcal{U}(x)$ that satisfies $e^{-\mathcal{U}(x)} \in L^1(\mathbb{R}^n_x)$. The assumption on the potential alone, is in fact enough, (as long as $G(x)$ is non degenerate) to imply the existence of a unique stationary state. We work with weak solutions to (1.3) in the sense given in [50], which allows for irregular coefficients that satisfy certain growth assumptions. The computations involving the relative entropy are performed first at a formal level, e.g. by assuming solutions that belong in Schwartz class with all derivatives vanishing polynomially fast. We continue with a standard regularization argument that approximates a solution f_ϵ with a mollified one $f_{\epsilon,\delta} \in C^\infty$. All the formal computations performed earlier will hold for $f_{\epsilon,\delta}$, with extra terms that will be shown to vanish as $\delta \to 0$. In the limit $\delta \to 0$, $H(f_\epsilon|\rho\mathcal{M})$ is non increasing. We then take $\epsilon \to 0$ and the result follows readily.

Under these assumptions we get theorem:

Theorem 2. *Let $f_\epsilon(0,x,v)$ be initial data for the FP equation s.t. $f_\epsilon(0,x,v) \geq 0$ and*

$$\sup_{\epsilon>0} \iint f_\epsilon(0,x,v)(1 + \mathcal{U}(x) + |v|^2 + \log f_\epsilon(0,x,v))\, dv\, dx \leq C < \infty.$$

Let $\rho(0,x) \in \mathcal{D}'(\mathbb{R}^n_x)$ be initial data to the limit equation that satisfies

$$\int \rho(0,x)\, dx = \iint f_\epsilon(0,x,v)\, dv\, dx = 1.$$

We finally make the assumption for the initial data that

$$H(f_\epsilon(0,x,v)|\rho(0,x)\mathcal{M}) \to 0 \quad as \quad \epsilon \to 0.$$

Then, if we assume a solution $\rho(t,x) \in C([0,T], \mathcal{D}'(\mathbb{R}^n_x))$ *to the limit equation for* $T > 0$*, we have*

$$\sup_{0 \le t \le T} H(f_\epsilon(t,x,v)|\rho(t,x)\mathcal{M}) \to 0 \quad as \quad \epsilon \to 0.$$

The two theorem presented above answer the question of relaxation of the hydrodynamic density ρ_ϵ to a limiting density ρ, which follows the prescribed macroscopic equation. These two results are to the best of my knowledge novel and assume weak solutions to equation (1.3).

1.3 Regularity

In this Section, we mention some of the partial contributions which will be presented in Chapter 4 related to the regularity of solutions, and review other known results.

We limit ourselves to the study of two types of regularity estimates, namely local and short time estimates. Much emphasis will be given to the second category.

A local regularity estimate can be obtained as a straightforward application of Hörmander's hypoellipticity theory. The theory suggests an estimate of the type

$$\|u\|_{H^r} \leq C(\|Lu\|_{L^2} + \|u\|_{L^2}) \qquad \text{for some } r > 0,$$

as long as the operator L can be expressed in the form $L = A^*A + B$ and the operators A, B which are smooth differential operators generate all the directions of differentiation (here x and v). Since this fact is rather established, we only review the main ingredients of a proof presented by J.Kohn (see e.g. [29, 43]) which uses the language of pseudo-differential operators.

The local regularity estimates can be extended with appropriate tools to regularity estimates that are valid on a short time interval, typically $0 \leq t \leq 1$. The main estimate for the Cauchy problem is

$$\|\nabla_x h\|_{L^2(\mu)} + \|\nabla_v^3 h\|_{L^2(\mu)} \leq \frac{C}{t^{3/2}}\|h_0\|_{L^2(\mu)} \qquad 0 \leq t \leq 1,$$

for a measure μ having density the stationary state $\mathcal{M}_{eq}(x, v)$, and $h = \frac{f}{\mathcal{M}_{eq}}$.

A good motivation for starting the study of short time estimates is to find the exact estimates for a quadratic potential. Here, we use a semi-explicit representation of the solution to the FP equation (given in [15]) that allows for a quadratic potential plus a smooth perturbation of it (as long as the perturbation decays fast enough in space). Based on estimates of this solution we present an exact short time estimate for initial data in $f_0 \in L^1(\mathbb{R}^{n,n}_{x,v})$, i.e.

$$\|f\|_{H^{k,l}_{x,v}(\mathbb{R}^{n,n}_{x,v})} \leq \frac{C}{t^{n+\frac{3}{2}k+\frac{1}{2}l}} \qquad 0 < t < t_0, \quad \text{for some} \quad t_0 > 0.$$

Short time regularity estimates will be obtained with the help of two different approaches. The first approach is a method originally used by F. Hérau which is quite elementary in nature. The method employs the use of the functional

$$\mathcal{E}(t,h) := \int h^2\, d\mu + at \int |\nabla_v h|^2\, d\mu + 2bt^2 \int \nabla_v h \cdot \nabla_x h\, d\mu + ct^3 \int |\nabla_x h|^2\, d\mu,$$

that depends on three parameters a, b, c (in general aligned as $1 \gg a \gg b \gg c$), which is proven to be dissipative for carefully selected values of those parameters, when $0 \leq t \leq 1$. It assumes a C^2 potential $\mathcal{U}(x)$ and a smooth friction $G(x) = \gamma(x)\mathrm{I}$ with $\gamma(x)$ bounded by $\Lambda_0 > \gamma(x) > \lambda_0 > 0$. An important feature of this technique is that it can be applied to the treatment of $L \log L(\mu)$ initial data, and give short time estimates for log-Sobolev type of norms. As will be explained in more detail, the Hérau technique is quite similar to techniques used for proving hypocoercivity of the operator L. This result is quite standard, and it will be reviewed in detail in Sections 4.2.2 - 4.2.3.

A different approach is provided by C.Villani and explained in detail in his excellent monograph [71]. It is, generally speaking, stronger than the one presented by Hérau but not quite as elementary. The objective of his approach is the construction of a system of differential inequalities that can be studied independently and provide the desired short time estimate. Although the Hérau method treats $L^2(\mu)$ & $L \log L(\mu)$ data, the power of the Villani approach is that it can be used for the treatment of L^1 initial data in the case of constant friction $\gamma > 0$. The interesting twist in this approach is that when we tried to implement it for a diagonal, non constant, smooth friction $\gamma(x)$ with bounds like above, we were only able to prove the desired inequalities in $L^2(\mu)$ for the quadratic potential (allowing L^∞ perturbations of it). An explanation of this restriction is based on the fact that a certain type of Sobolev interpolation inequalities that are part of the differential system, can only be proven "by hand" in flat and in the Gaussian measure. It is exactly for this reason that the method itself leaves some very interesting extensions and open problems. These results will be presented in Sections 4.2.4 - 4.2.6.

The exact Sobolev estimate for the quadratic potential with constant friction $\gamma > 0$, appears to be new, at least as it is derived in the case with an added perturbation. Although there is no doubt that the result for the purely quadratic case is known for quite some time, our case is more interesting in that it reaffirms the robustness of the estimate under a smooth enough perturbation. This result can be found in Section 4.2.1. In the same spirit, Theorem 14, presented in Section 4.2.5 is another minor addition to theory.

1.4 Outline

Chapter 2 gives a detailed description of the fluid/particle system. In Section 2.2 we provide an iterative technique of solving the Stokes' equations of motion which converges under certain assumptions. A characterization of a regime of convergence (dilute regime) is given in Section 2.2.2. Section 2.3 gives the most important approximations of the friction/hydrodynamic tensors.

In chapter 3, we give a definition of a weak solution to the Cauchy problem. We shall present a theory for L^2 & L^1 initial data. This theory is based on the study of renormalized solutions by Di-Perna & Lions (see [17, 49, 50]), for solutions of transport equations with irregular coefficients. The existence of solution is a straightforward consequence of the energy estimate, but some extra effort is required for showing uniqueness.

Chapter 4 follows with a discussion of results on regularity. As we explained already, two types of estimates are more common. Local estimates are given in Section 4.1, with the presentation of Hörmander's theorem on hypoellipticity and an exact estimate based on [32]. In Section 4.2 we list all the details of the results mentioned in Section 1.3.

Chapter 5 follows with the study of convergence to the unique stationary state solution. We review an elementary method (presented in detail in [71]) very close in spirit to the Hérau regularization method, giving convergence in $L^2(\mu)$ (Section 5.1). A different result following [15] is given in Section 5.2. An L^1 result is presented in Section 5.3. The chapter closes with a method that can be applied to a wider class of problems (not necessarily linear) following [20].

Finally, in Chapter 6, after a brief discussion of the formal derivation of the hydrodynamic limit, we give the proof of the two results mentioned earlier.

Chapter 2: Particle System

As a motivation of what follows, we begin with a presentation of the dynamics of the particle system. The system at hand consists of Stokes' equations of fluid motion for a medium and the kinematic equations of motion for a number of particles embedded in the fluid. The fluid interacts with the particles through the boundary conditions on the surface of the particles. Although the ensuing system cannot be solved exactly, there is an iterative procedure that in some cases may be shown to converge to the solution. Such an instance appears in the so called dilute regime which we discuss briefly.

Due to linearity of Stokes' problem, hydrodynamic forces acting on the particles depend linearly to the particle velocities. This dependence is encapsulated in the form of the hydrodynamic friction $G_{ij}(x_1, \ldots, x_N)$ tensor. The friction $G_{ij}(x_1, \ldots, x_N)$ measures the interaction (through the fluid) between two particles with centers at points x_i and x_j. Naturally, we have a freedom of choice in constructing approximations of G_{ij} or its inverse mobility tensor and we give the two most important examples in literature. Of course, we should always keep in mind the very interesting case of identity friction, which physically corresponds to particles that move freely without interacting with each other.

2.1 Particle Dynamics with Noise

A physical motivation for the occurrence of the Fokker-Planck equation that is being studied here is provided in the realm of particle dynamics. The particle system under study can be found e.g. in [36, 37] etc. The inclusion of "white noise" is one of the ways to maintain dissipative structure.

Consider N identical spherical particles, of uniform density, mass m and radius R, located in physical space $\mathbb{R}^3$. Let x_i be the center of the i'th particle and B_i the open ball with center x_i and radius R. S_i is the surface of the ball. The particles are immersed in a slow, incompressible, Stokes flow. All forces are assumed central so the particles cannot rotate. The particle system (fluid + particles) obeys the system of equations:

$$-\eta \triangle u + \nabla p = 0 \qquad x \in \mathbb{R}^3 \setminus \bigcup_i \bar{B}_i, \tag{2.1a}$$

$$\nabla \cdot u = 0 \qquad x \in \mathbb{R}^3 \setminus \bigcup_i \bar{B}_i, \tag{2.1b}$$

$$u(x) = v_i \quad \text{for} \quad x \in S_i \quad 1 \leq i \leq N, \tag{2.1c}$$

$$u(x) \to 0 \quad \text{as} \quad |x| \to \infty, \tag{2.1d}$$

$$\sigma(x) = -p\mathrm{I} + \eta(\nabla u + \nabla u^T), \tag{2.1e}$$

$$\frac{dx_i}{dt} = v_i, \tag{2.1f}$$

$$m\frac{dv_i}{dt} = -\int_{S_i} \sigma \cdot n dS - \nabla_{x_i} \mathcal{U}(\mathbf{x}). \tag{2.1g}$$

In the above system, η is the fluid viscosity, $u(x)$ a velocity vector in $\mathbb{R}^3$, and $p(x)$ the fluid pressure at a point x in space. We assume an external potential $\mathcal{U}(\mathbf{x})$

that is a function of the configurational vector $\mathbf{x} = (x_1, x_2, \ldots, x_N) \in \mathbb{R}_x^{3N}$. Later we will drop the boldface typeface and describe by x the configurational vector in $\mathbb{R}_x^{3N}$ (for now we use it to avoid confusion with the typical spatial variable in $\mathbb{R}^3$). In (2.1c) & (2.1d) we consider two boundary conditions. (2.1d) is the condition of a vanishing velocity field at infinity and (2.1c) assumes that the velocity on the surface of the i'th particle is the constant vector $v_i \in \mathbb{R}^3$.

Forces that act on a surface in the fluid (hydrodynamic forces) are described by the surface integral $-\int_{S_i} \sigma \cdot n \, dS$ of the viscous stress tensor $\sigma(x)$. Finally, (2.1g) is the equation of particle motion that states that the force on the particle has a part caused by hydrodynamic interactions and one caused by the external potential $\mathcal{U}(x)$.

Of central importance to our theory, is that linearity implies that the force the i'th particle exerts on the fluid (F_i), can be written in the form

$$F_i = \sum_{j=1}^{N} G_{ij} v_j = \int_{S_i} \sigma \cdot n \, dS.$$

The tensor $\{G_{ij}\}_{i,j=1}^{N}$ is symmetric and non-negative in $\mathbb{R}_x^{3N} \times \mathbb{R}_x^{3N}$. The non-negativity of the friction tensor follows from a study of the energy dissipation of the particle system, since

$$\sum_{i,j=1}^{N} v_i^T G_{ij} v_j = 2\eta \int_{\mathbb{R}^3 \setminus \bigcup_i B_i} e : e \, dx \tag{2.2}$$

with $e = \frac{1}{2}(\nabla u + \nabla u^T)$ (rate of strain tensor) is the energy dissipation of the particle system.

Introducing a more compact vector/tensor notation, we set $F = (F_1, \ldots, F_N)$, $v = (v_1, \ldots, v_N)$ for the $\mathbb{R}^{3N}$ hydrodynamic force and velocity vectors. The friction inverse is the mobility tensor $H = G(x)^{-1}$ that has been the object of study of polymer dynamics for dilute sedimentations.

Naturally, the N particle system does not have an exact solution because of the complexity of the boundary. The usual approach for treating hydrodynamical interactions is to introduce an approximation for the friction or mobility tensors that respects the following properties:

- It is valid for all configurations, admissible and non-admissible. By non-admissible, we mean cases where particles overlap. Since we are dealing with a kinetic formulation that doesn't account for collisions this is a very important property to have. Of course, this also allows for a freedom in the choice of the friction or mobility for non-admissible configurations.

- It is non-negative for all configurations. This property is consistent with the non-negativity of the "exact" friction and mobility that guarantees dissipation of energy for the system.

We are working under the assumption that the particles are located in a "thermal bath", modeled by a stochastic term. The dynamics of particles, after the inclusion of Brownian motion, is described by the system of equations

$$\frac{dx}{dt} = v, \tag{2.3a}$$

$$m\frac{dv}{dt} = -Gv - \nabla\mathcal{U}(x) + \sqrt{2}G^{1/2}(x)\frac{dW}{dt}. \tag{2.3b}$$

$W(t)$ is a Brownian vector in $\mathbb{R}^{3N}$, and $\frac{dW}{dt}$ is the "white noise". The derivative in the Brownian vector is used only as a notational instrument, since the Brownian measure is not differentiable in the classical sense. The statistics of the "white noise" is given by

$$\mathbb{E}\left(\frac{dW(t)}{dt}\right) = 0, \tag{2.4a}$$

$$\mathbb{E}\left(\frac{dW(t)}{dt}\frac{dW(t')}{dt'}\right) = \delta(t-t'), \tag{2.4b}$$

where $\mathbb{E}$ here stands for the expectation w.r.t. to Brownian measure.

2.2 Reflections Method

Although, it is impossible to have an exact solution for the general N particle Stokes problem for the domain $D = \mathbb{R}^3 \setminus \cup_k \overline{B}_k$, we can try an approximation process using reflections. The method is an iterative technique in which one solves the Stokes problem in a simpler domain and makes a successive correction at each step. In our case, we solve the problem for a single particle and correct the BCs on the surface of particles in the next step.

2.2.1 Approximation Scheme

We introduce the formulation of the reflections technique. Let $U_i(x)$ be the velocity on the surface of the i'th particle. Notice, that in general, we don't assume a homogeneous velocity on particle surface.

The ensuing steps of the method are:

- 1-step: Solve for $i = 1, \ldots, N$

$$\eta \triangle u_i^{(1)}(x) - \nabla p_i^{(1)}(x) = 0,$$

$$\nabla \cdot u_i^{(1)}(x) = 0 \quad \text{in} \quad x \in \mathbb{R}^3 \setminus \overline{B}_i,$$

with B.Cs

$$u_i^{(1)}(x) = U_i(x) \qquad x \in S_i.$$

We also define $u^{(1)}(x) = \sum_{i=1}^{N} u_i^{(1)}(x)$ and $p^{(1)}(x) = \sum_{i=1}^{N} p_i^{(1)}(x)$.

- n-step: Solve for $i = 1, \ldots, N$

$$\eta \triangle u_i^{(n)}(x) - \nabla p_i^{(n)}(x) = 0$$

$$\nabla \cdot u_i^{(n)}(x) = 0 \quad \text{in} \quad x \in \mathbb{R}^3 \setminus \overline{B}_i$$

with B.Cs

$$u_i^{(n)}(x) = -\sum_{j \neq i} u_j^{(n-1)}(x) \qquad x \in S_i.$$

We finally define $u^{(n)}(x) = u^{(n-1)}(x) + \sum_{i=1}^{N} u_i^{(n)}(x)$ and $p^{(n)}(x) = p^{(n-1)}(x) + \sum_{i=1}^{N} p_i^{(n)}(x)$.

We have reserved subscript notation to denote particle number ($i = 1, \ldots, N$), and superscripts to denote the step number ($n = 1, 2, \ldots$). We argue that (under certain conditions which we will present) the limit of $u^{(n)}(x)$ solves the N particle Stokes problem. The convergence of the reflections technique depends on the relative position of particles and holds in an appropriate space.

In order to prepare for the presentation of the convergence result, we borrow the notation from Jabin & Otto, [36]. This formulation borrows from the language of operators.

We describe the solution of the single-particle problem for the j'th particle with the help of an operator T_j. T_j acts on a vector U defined on the surface of the particle and maps it onto the solution of the Stokes' problem for the particle. The operator will be called Stokes operator for the j'th particle, i.e.

$$T_j : L_0^\infty(S_j) \to L^\infty(\mathbb{R}^3 \setminus B_j)$$

$$U \to T_j U$$

,where L_0^∞ is L^∞ with U picked with an average normal component that is zero, i.e

$$\int_{S_j} U \cdot n \, dS = 0.$$

We also consider the family of operators $\{A_{ij}\}_{i,j=1}^{N}$, from $L_0^\infty(\cup_k B_k)$ to $L_0^\infty(\cup_k B_k)$, with the property

$$A_{ij}U = \begin{cases} T_j(U|_{S_j})|_{S_i} & i \neq j \\ U & i = j. \end{cases}$$

Using the language of operators, the reflections technique is formalized by

- 1-step:

$$u^{(1)}(x) = \sum_{i=1}^{N} T_i(U_i)$$

- n-step:

$$u^{(n)}(x) = u^{(n-1)}(x) + \sum_{i=1}^{N} u_i^{(n)}(x),$$

with

$$u_i^{(n)}(x) = \sum_j T_i(\mathrm{I} - A)_{ij}^n U_j.$$

The solution, if the technique converges, should be formally written as

$$u(x) = \sum_{n=0}^{\infty} \sum_{i,j} T_i(\mathrm{I} - A)_{ij}^n U_j,$$

or using formal Neumann expansion

$$u(x) = \sum_{i,j} T_i A_{ij}^{-1} U_j.$$

2.2.2 Dilute Regime & Convergence of Reflections Technique

We are about to present a case where convergence of the reflections method holds. This case goes back to the work in [36] in which the dilute regime of particle sedimentation is described. This regime is the one in which particles approximately behave like free particles (with no interaction between them).

The **dilute regime** is characterized by $R \lesssim d_{min}$ and $\Lambda \lesssim 1$, where

$$d_{ij} = |x_i - x_j| \qquad \& \qquad d_{min} = \min_{i \neq j} d_{ij}.$$

The critical parameter Λ, is

$$\Lambda = \frac{R}{d_{min}} N^{\frac{2}{3}}$$

for a cloud of N identical spherical particles with radius R each.

We use the symbol $\lesssim$, for inequalities, instead of $\leq$, in order to suppress the use and having to re-evaluate constants that do not depend on $R, N \,\&\, d_{min}$. We may now proceed to

Theorem 3. *In the dilute regime the following estimates hold:*

(i)

$$\|T_j U\|_{S_i} \lesssim \frac{R}{d_{ij}} \|U\|_{S_j} \qquad \textit{for } j \neq i,$$

(ii)

$$\|A_{ij} U\|_{S_i} \lesssim \frac{R}{d_{ij}} \|U\|_{S_j} \qquad \textit{for } j \neq i,$$

(iii)

$$\|(\mathrm{I} - A)^n_{ij} U\|_{S_i} \leq \begin{cases} (C\Lambda)^{n-1} \frac{R}{d_{ij}} \|U\|_{S_j} & \textit{for } j \neq i \\ (C\Lambda)^{n-1} \frac{\Lambda}{N} \|U\|_{S_i} & \textit{for } i = j, \end{cases}$$

and some constant $C > 0$,

(iv)

$$\|A^{-1}_{ij} U\|_{S_i} \lesssim \frac{R}{d_{ij}} \|U\|_{S_j} \qquad \textit{for } j \neq i,$$

(v)

$$\|A_{ii}^{-1}U\|_{S_i} \lesssim \|U\|_{S_i}$$

for $\|\cdot\|_{S_i}$ *being the* L^∞ *norm on* S_i *.*

If we also define

$$G_{ij}U = \int_{S_i} \sigma(T_i A_{ij}^{-1}U) \cdot n \, dS$$

for a constant vector U*, then the following hold*

(vi)

$$|G_{ii} - 6\pi\eta R\mathrm{I}| \lesssim \eta \frac{R^3 N^{\frac{1}{3}}}{d_{min}^2},$$

(vii)

$$|G_{ij}| \lesssim \eta \frac{R^2}{d_{ij}} \qquad \textit{for } j \neq i,$$

where $|\cdot|$ *is now the matrix norm induced by* $\|\cdot\|$*.*

Proof. Care will be given in presenting the most important steps of the proof of the statement. The proof is presented in many of the results in [36].

The statements (i) & (ii) are a straightforward result of the regularity of the Stokes problem solution for a ball. As we show in appendix this is trivial for the Stokes problem with constant B.Cs. The same regularity estimate holds for the Stokes problem solution with inhomogeneous BCs as shown after more delicate analysis.

Statements (iv) & (v) are also an easy consequence of the estimate (iii). The formal Neumann expansion for A^{-1} is

$$A_{ij}^{-1} = \sum_{n=0}^{\infty} (\mathrm{I} - A)_{ij}^n$$

So, with the help of (iii), for $j \neq i$

$$\|A_{ij}^{-1} U\|_{S_i} \lesssim \sum_{n=1}^{\infty} \Lambda^{n-1} \frac{R}{d_{ij}} \|U\|_{S_i} \lesssim \frac{R}{d_{ij}} \|U\|_{S_i}.$$

For $j = i$, since

$$A_{ii}^{-1} U = U + \sum_{n=2}^{\infty} (\mathrm{I} - A)_{ii}^n U$$

we have the estimate

$$\|A_{ii}^{-1} U - U\|_{S_i} \lesssim \sum_{n=2}^{\infty} \Lambda^{n-1} \frac{\Lambda}{N} \|U\|_{S_i} \lesssim \frac{\Lambda^2}{N} \|U\|_{S_i}$$

which yields estimate (v).

Some extra effort is needed in establishing (iii). The argument is carried out using induction. The inductions step uses the relation

$$(\mathrm{I} - A)_{ij}^{n+1} U = \sum_k (\mathrm{I} - A)_{ik} (\mathrm{I} - A)_{kj}^n U = - \sum_{k \neq i} A_{ik} (\mathrm{I} - A)_{kj}^n U.$$

Using (ii), the above relation yields

$$\|(\mathrm{I} - A)_{ij}^{n+1} U\|_{S_i} \lesssim \sum_{k \neq i} \frac{R}{d_{ik}} \|(\mathrm{I} - A)_{kj}^n U\|_{S_k}.$$

The reason why the critical parameter Λ appears in the estimate (iii) has its foundations in the inequalities

$$\sum_{j \neq i} \frac{R}{d_{ij}} \lesssim \Lambda \qquad \& \qquad \sum_{j \neq i} \left(\frac{R}{d_{ij}} \right)^2 \lesssim \frac{\Lambda^2}{N}$$

which can be proven with an application of a covering lemma (for the first one) and some extra algebra for the second. The two for-mentioned inequalities are used in the induction step. We omit the details of the induction step.

Using single particle solution for a constant vector U we have,

$$\int_{S_i} \sigma(T_j U) \cdot n \, dS = \delta_{ij} \, 6\pi\eta R U.$$

For non-constant vectors it can still be proved using the regularity of Stokes' solution that

$$\int_{S_i} \sigma(T_i U) \cdot n \, dS \lesssim \eta R \|U\|.$$

According to the definition of tensor G_{ij} we also have

$$G_{ii} U - 6\pi\eta R U = \int_{S_i} \sigma(T_i(A_{ii}^{-1} - U)) \cdot n \, dS.$$

The above yields,

$$\|G_{ii} U - 6\pi\eta R U\| \lesssim \eta R \|A_{ii}^{-1} U - U\| \lesssim \eta \frac{R^3 N^{1/3}}{d_{min}^2} \|U\|.$$

With similar arguments (vii) is established.

$\square$

2.3 Approximations of Friction & Mobility Tensors

2.3.1 Stokeslet Approximation & Oseen Tensor

We have asked the question of constructing a mobility or friction tensor with the properties mentioned earlier. The first related construction will be based on a simplified particle model.

Assume a very viscous flow in $\mathbb{R}^3$ and point particles localized at position x_i for $i = 1, \ldots, N$. Let F_i be the hydrodynamic force acting on the i'th particle (no other force is assumed to act on it). The flow is once again a Stokes, incompressible, with equation of motion adjusted to be

$$\nabla p(x) - \eta \triangle u(x) = -\sum_{i=1}^{N} F_i \delta(x - x_i) \qquad x \in \mathbb{R}^3. \tag{2.5}$$

The boundary condition at infinity is the same as in the original problem, namely $u \to 0$ as $|x| \to \infty$. See, for instance, [19]. The role that the Dirac delta function plays is to localize forces, since we are dealing with point particles. For the sake of mathematical rigor, we may view the delta function as a distribution, or might as well consider the weak formulation of the problem.

The above problem is solvable, as one can see in [19] and with the help of the superposition principle we find

$$u(x) = \sum_{i=1}^{N} H(x - x_i) F_i, \tag{2.6}$$

where the Green's function is the 3×3 mobility tensor, given by

$$H(x) = \frac{1}{8\pi\eta_s |x|} \left(\mathrm{I} + \frac{x \otimes x}{|x|^2} \right).$$

The tensor is divergence free

$$\nabla \cdot H(x) = 0,$$

as one would expect due to the fluid incompressibility. It is positive definite, in the sense that $\xi^T H \xi > 0$ for $\xi \in \mathbb{R}^3 \setminus \{0\}$. This property is extremely important in the theory for the mean field limit, when the number of particles $N \to \infty$. Finally, it exhibits asymptotic behavior

$$H(x) \sim \frac{C}{|x|}\mathrm{I} \quad \text{for} \quad |x| \gg 1.$$

The first mathematical difficulty appears in the fact that H is singular at 0. This prohibits us from assigning a value to the velocity of particles simply by $V_i = u(x_i)$. Engineers go around this difficulty by assigning the value $H(0) = \frac{1}{\zeta}\mathrm{I}$. Here $\zeta = 6\pi\eta$ is the friction that a single particle of radius 1 experiences in a viscous environment.

We have solved one problem by defining H at 0 and we can now consider the $3N \times 3N$ tensor $\{H_{ij}\}_{i,j=1}^{N} = H(x_i - x_j)$ for a given particle configuration. It turns out this tensor is not non-negative for all configurations. In fact, problems usually arise for small distances between particles (by small here we mean compared to 1). The name that this tensor carries is Oseen tensor. The flow studied above is known as Stokeslet flow. We will see how we can get around the inconvenience of a non positive hydrodynamic tensor.

2.3.2 Rotne-Prager-Yamakawa Tensor

In this section, we follow the idea found in [65] for the approximation of the friction tensor based on the variational formulation of the Stokes problem ((2.1a)-(2.1d)). The same approximation has been reproduced in [73] with the use of a different method. The variational technique states that the energy dissipation, seen as the integral in (2.2), is minimized for the solution of the problem.

By picking a trial velocity field u^{ap}, we calculate an approximation G_{ij}^{ap} with the property that $G_{ij}^{ap} \geq G_{ij}$. This not only guarantees the positivity of G_{ij}^{ap}, but also that this approximation is bounded below by G_{ij}. Briefly, the choice of the approximate u^{ap} will be the superposition of the solution to the Stokes problem for each individual sphere. This might appear a crude approximation but for particles that are well separated it is satisfactory. Thus, $u^{ap}(x) = \sum_{i=1}^{N} u_i(x)$, where

$$\begin{aligned} u_i(x) = &\left(\frac{3}{4}\frac{R}{|x-x_i|} + \frac{1}{4}\frac{R^3}{|x-x_i|^3}\right) v_i \\ &+ \left(\frac{3}{4}\frac{R}{|x-x_i|} - \frac{3}{4}\frac{R^3}{|x-x_i|^3}\right) \frac{(x-x_i)\otimes(x-x_i)}{|x-x_i|^2} v_i \end{aligned}$$

is the outer solution ($|x-x_i| \geq R$) of the one particle Stokes problem.

To simplify the computation of the integral that appears in (2.2), we extend the integration from $\mathbb{R}^3 \setminus B_i$ to $\mathbb{R}^3$. Doing this we can only add to the value of the integral (since the integrand term is non negative), so $G_{ij}^{ap} \geq G_{ij}$ is reinforced. For this we make the extra assumption that the inner solution for one particle is 0.

The actual computation will give the following expressions for the approximation of the friction tensor:

The approximate friction has diagonal elements

$$G_{ii} = 6\pi\eta R\mathrm{I}.$$

The non-diagonal elements are

$$G_{ij} = \frac{9}{2}\pi\eta\frac{R^2}{|x_{ij}|}\Big[\left(\mathrm{I} + \frac{x_{ij}\otimes x_{ij}}{|x_{ij}|^2}\right) + \frac{2R^2}{|x_{ij}|^2}\left(\frac{1}{3}\mathrm{I} - \frac{x_{ij}\otimes x_{ij}}{|x_{ij}|^2}\right)\Big]$$

for $|x_{ij}| \geq 2R$, and

$$G_{ij} = 6\pi\eta R\Big[\left(1 - \frac{9}{32}\frac{|x_{ij}|}{R}\right)\mathrm{I} + \frac{3}{32}\frac{x_{ij}\otimes x_{ij}}{R|x_{ij}|}\Big]$$

for $|x_{ij}| \leq 2R$.

The task of finding eigenvalues for each configuration of the RPY friction in the N particle system requires the solution of an algebraic equation of order $3N$. It is therefore more natural to obtain bounds for the eigenvalues. For the 2-particle system, when the radius is R and the centers of the particles are distanced by $|d| = |x_{12}|$, the exact eigenvalues are:

For $|d| \geq 2R$

$$\lambda_{1,2} = 6\pi\eta R \pm \frac{36\pi\eta R^2}{|d|}\left(\frac{1}{4} - \frac{1}{6}\frac{R^2}{|d|^2}\right)$$
$$\lambda_{3,4} = 6\pi\eta R \pm \frac{36\pi\eta R^2}{|d|}\left(\frac{1}{8} + \frac{1}{12}\frac{R^2}{|d|^2}\right),$$

and for $|d| \leq 2R$

$$\lambda_{1,2} = 6\pi\eta R \pm 36\pi\eta R\left(\frac{1}{6} - \frac{1}{32}\frac{|d|}{R}\right)$$
$$\lambda_{3,4} = 6\pi\eta R \pm 36\pi\eta R\left(\frac{1}{6} - \frac{3}{64}\frac{|d|}{R}\right).$$

Each of $\lambda_{1,2}$ is simple, whereas $\lambda_{3,4}$ are both double. The minimum eigenvalue for overlapping spheres ($|d| \leq 2R$) is $\lambda_{min} \geq \frac{9}{8}\pi\eta|d|$ and for non- overlapping ($|d| \geq 2R$), $\lambda_{min} \geq \frac{9}{4}\pi\eta R$.

Introducing the friction tensor $G(x)$ for the particle configuration $x = (x_1, x_2, \ldots, x_N)$, it is evident that the RPY approximation satisfies

$$\lambda_{min}(x)\mathrm{I} \leq G(x) \leq \lambda_{max}\mathrm{I},$$

in the sense of non-negative forms. For two particles we have just shown that

$$\lambda_{min}(x_1, x_2) \sim |x_1 - x_2|.$$

Finally, we shall mention the interesting case of a particle system in the absence of hydrodynamic interactions. In such a regime, particles move like "free particles" in the medium and a coarse approximation can be made by assuming $G(x) = 6\pi\eta R\mathrm{I}$.

Chapter 3: Well-Posedness of the Fokker-Planck Equation

In this chapter we present the Cauchy problem $\partial_t f + Lf = 0, \quad f|_{t=0} = f_0$ (FP equation + initial data) describing the statistical evolution of the particle system discussed in the previous section. For this and the next two chapters the focus is directed in giving results on existence (uniqueness), regularity and large time asymptotics for the non scaled operator L that corresponds to particles with constant mass.

In particular, we discuss how the FP equation arises in a simple application of Itô's formula and give formulations of the equation that are equivalent and are appropriate for a study in different functional settings. The writing of operator L as the sum of a transport term $\mathcal{T}$, and a collision term $\mathcal{C}$ is followed by a brief study of the properties of these operators as well as of the semi-group $\left(e^{-tL}\right)_{t\geq 0}$. The main question answered in this section is the existence of a **unique**! distributional solution under the assumption of initial data in L^2 or even better in L^1. The theory for this is well established for quite irregular friction $G(x)$, and potential $\mathcal{U}(x)$ (see [17, 49, 50]) and is given for completeness.

3.1 Cauchy Problem

The objective from this point forward is to pass from the particle dynamics system to a kinetic level of description. Consider the random vector $X = (x, v) \in \mathbb{R}^{3N,3N}_{x,v}$, that follows the motion described by the stochastic differential system (2.3a)-(2.3b). The statistics of motion at time t is characterized by the probability density function (p.d.f.) $f(t, x, v)$. This function describes the probability of finding the random vector $\tilde{X} = (\tilde{x}, \tilde{v})$ between the states X and $X + dX$, for an infinitesimal phase space vector dX, i.e.

$$f(t, x, v)dx = \mathbb{P}(\{\omega \in \Omega |\, X \leq \tilde{X}(\omega) \leq X + dX\})$$

in an appropriate probability space $(\Omega, \mathfrak{F}, \mathbb{P})$.

Consider the semigroup P_t, defined by

$$P_t\psi(x) = \mathbb{E}(\psi(X_t)|X_0 = X)$$

and acting on bounded measurable functions $\psi : \mathbb{R}^{6N} \to \mathbb{R}$. Let $\mathcal{L}$ be the generator of the semi-group ($\mathcal{L} = \frac{d}{dt}\Big|_{t=0} P_t$) and $\mathcal{L}^*$ its adjoint in L^2. $\mathcal{L}$ can be found for our system with the help of Itô's formula (or Feynmann-Kac formula for the matter), so that we arrive to the Backward-Kolmogorov equation of "observables" $\partial_t P_t\psi = \mathcal{L}P_t\psi$. Then, the density $f(t, x, v)$ satisfies the forward Fokker-Planck (Kolmogorov) equation $\partial_t f = \mathcal{L}^* f$, i.e.

$$\partial_t f + \nabla_x \cdot (vf) - \frac{1}{m}\nabla_v \cdot ((G(x)v + \nabla\mathcal{U}(x))f) = \frac{1}{m^2}\nabla_v \cdot (G(x)\nabla_v f)\,.$$

In order to study the diffusion limit of the above, one needs to introduce the appropriate scaling to separate conservative and dissipative terms. The scaling procedure is described with more detail in [14] and follows the change of variables,

$$m = \epsilon^2, \qquad v' = \epsilon v, \qquad x' = x.$$

The above change of variables leads to the following equation (after we re-introduce the notation for x, v in the place of x', v')

$$\partial_t f_\epsilon + \frac{1}{\epsilon}\mathcal{T}(f_\epsilon) = \frac{1}{\epsilon^2}\mathcal{C}(f_\epsilon), \tag{3.1}$$

where

$$\mathcal{T}(f) = v \cdot \nabla_x f - \nabla\mathcal{U}(x) \cdot \nabla_v f$$

is the sum of an advective/transport term $v \cdot \nabla_x f$, and a confinement $F(x) \cdot \nabla_v f$ for $F(x) = -\nabla_x \mathcal{U}(x)$. For simplicity, we call operator $\mathcal{T}$ the transport operator.

The term

$$\mathcal{C}(f) = \nabla_v \cdot \left(G(x)\left(\nabla_v f + vf\right)\right),$$

on the other hand, is the dissipative (or collision) part of the equation.

The scaling we have chosen is the only one that provides separation of scales for the collision and the transport terms. The limit $\epsilon \to 0$ corresponds to small mass limit. In the first part of our study (in what relates to questions of well-posedness, regularity and asymptotics) we will focus in the non parametrized version of the equation, with $m = \epsilon = 1$. This equation can be presented in the form of the Cauchy problem

$$\partial_t f + Lf = 0, \qquad f\big|_{t=0} = f_0(x, v), \tag{3.2}$$

with $L = \mathcal{T} - \mathcal{C}$ an operator which from now on will be called Fokker-Planck operator.

Remark 1. *As a side remark on the history on the equation* (3.2)*, it should be noted that Adriaan Fokker [22] and Max Planck [60], were the first to derive a PDE for a stochastic equation with noise in other than spatial variables (with the exception of Lord Rayleigh in* 1891*). Fokker obtained a stationary equation for a probability density* $W(q,t)$*, with* q *being the angular momentum of a dipole in an environment with fluctuations. Planck derived the non-stationary equation on his own, few years later. The first instance of an equation with exactly the same structure (density* $f(t,x,v)$*, diffusion only in velocities), appeared originally in [42] in the work of Oskar Klein . Later, Hans Kramers derived the same equation in [46]. Thus, a more accurate name for* (3.2) *could be Klein-Kramers (or even Klein-Kramers-Chandrasekhar according to others). For the above information and much more see [21]. For an earlier account see [10].*

To give some perspective, the fundamental solution to the 1-d (for x, v) case with constant friction $\gamma > 0$ and initial conditions $f(0,x,v) = \delta(x_0, v_0)$ for the equation

$$\partial_t f + v\partial_x f = \gamma \partial_v^2 f \qquad (n = 1),$$

is known for quite some time (see e.g [44, 45]) and is given by

$$f(t,x,v) = \frac{2\sqrt{3}}{\pi\gamma^2 t^2} exp\left(-\frac{|v - v_0|^2}{4\gamma t} - 3\frac{|x - x_0 - \frac{1}{2}(v + v_0)t|^2}{\gamma t^3} \right)$$

(see appendix for more details and other solvable cases).

We can use the fundamental solution to extract various regularity estimates, but we would rather focus on regularity results for cases where an explicit solution cannot be obtained. An explicit solution can in general be obtained for a quadratic potential $\mathcal{U}(x) = x^2$ with constant diffusion. The regularity question will be studied in detail later.

The general structure of a FP equation in phase space is

$$\partial_t f + \mathrm{div}_{x,v}(bf) = \mathrm{div}_{x,v}(\sigma\sigma^T \nabla_{x,v} f),$$

with $b(x,v) \in \mathbb{R}^{6N}$ being a vector field, and $\sigma(x,v) \geq 0$ a possibly degenerate matrix in $\mathbb{R}^{6N\times 6N}$. In our case, diffusion acts only upon the velocity variable and the ellipticity condition

$$\xi^T \sigma(x,v)\xi \geq \lambda|\xi|^2 \qquad \text{for some} \quad \lambda > 0, \, \forall \xi \in \mathbb{R}^{6N}$$

fails trivially even for uniformly positive $G(x)$. More specifically, the failure of ellipticity won't be a problem for the existence of a unique weak solution. A weak solution can be constructed, for quite irregular coefficients b and σ, in the realm of the theory of renormalized solutions first presented in [17].

On the other hand the failure of ellipticity poses some technical issues in the study of regularity and relaxation to a unique global equilibrium state. The interplay between the transport and collision terms will in fact be responsible for regularization in the missing x direction. This phenomenon has been studied thoroughly by tools of hypoellipticity theory.

A similar problem will arise in the study of convergence to a global equilibrium state. In terms of the long time (asymptotic) behavior, the collision operator acts only in the velocity variable and tends to draw the system in the so called local equilibria states which are Maxwellians. The transport term on the other hand is the reason that a solution is drawn away from each local equilibrium and driven towards a unique, global equilibrium state $\mathcal{M}_{eq}(x,v)$, where

$$\mathcal{M}_{eq}(x,v) = e^{-\mathcal{U}(x)} \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{3N/2}}$$

for our problem. The potential $\mathcal{U}(x)$ is normalized so that $\int e^{-\mathcal{U}(x)}\,dx = 1$. This problem will be solved by constructing a norm for which L becomes coercive.

A different formalism for the (3.2) problem is presented by considering the equation for $h = \frac{f}{\mathcal{M}_{eq}}$. The equation for h can now be written as $\partial_t h + Lh = 0$, for the operator

$$Lh = v \cdot \nabla_x h - \nabla \mathcal{U}(x) \cdot \nabla_v h - \nabla_v \cdot (G(x)\nabla_v h) + v \cdot G(x)\nabla_v h,$$

and h is now normalized by $\iint h_0 \mathcal{M}_{eq}\,dv\,dx = 1$.

A second commonly used conjugated formulation of the equation is by considering the equation for $h = \frac{f}{\mathcal{M}_{eq}^{1/2}}$. The FP operator L now becomes

$$Lh = v \cdot \nabla_x h - \nabla \mathcal{U}(x) \cdot \nabla_v h + G^{1/2}\left(-\nabla_v + \frac{v}{2}\right) \cdot G^{1/2}\left(\nabla_v + \frac{v}{2}\right) h,$$

and the normalization for h now becomes $\iint h_0 \mathcal{M}_{eq}^{1/2}\,dv\,dx = 1$.

The significance of the two conjugated versions of the Cauchy problem becomes apparent when each one is attached to an appropriate functional setting. This setting is the "weighted" L^2, i.e. $L^2(\mu)$ (with μ being the stationary measure of L) for the first one, and the flat L^2 space for the second. L can now, in both cases, be written as $L = A^*A + B$, with B being the anti-symmetric transport term $\mathcal{T}$ and A^*A the self-adjoint diffusion operator $\mathcal{C}$. The adjoint of an operator will be understood in the corresponding functional space.

3.2 Properties of the Operators $\mathcal{C}$, L and the Semigroup $(e^{-tL})_{t\geq 0}$

We shall review some of the important properties of operators $\mathcal{C}$, L and $(e^{-tL})_{t\geq 0}$. To simplify things, especially for the reason of having an operator L with a unique stationary measure, we assume a non degenerate matrix $G(x)$.

For a full matrix $G(x)$, the null space of $\mathcal{C}$ is

$$\mathcal{N}(\mathcal{C}) = \{f(x,v) | \exists \phi(x) \quad \text{s.t} \quad f(x,v) = \phi(x)\mathcal{M}(v)\},$$

which consists of the local equilibria states spanned by the standard Maxwellian in velocity space $\mathcal{M}(v) = \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{\frac{3N}{2}}}$.

The null space of L is contained in the intersection of the null spaces of $\mathcal{T}$, $\mathcal{C}$. The following proposition sheds light on the nature of stationary states for L.

Proposition 1. *Assume a potential $\mathcal{U}(x) \in C^1(\mathbb{R}^{3N}_x)$, with $e^{-\mathcal{U}(x)} \in L^1(\mathbb{R}^{3N}_x)$. Then, there exists a unique stationary state for equation $\partial_t f + Lf = 0$, characterized by the Maxwell-Boltzmann distribution*

$$\mathcal{M}_{eq}(x,v) = \frac{e^{-E(x,v)}}{Z},$$

with $E(x,v) = \frac{1}{2}|v|^2 + \mathcal{U}(x)$ the Hamiltonian of the system. Z is the partition function $Z = (2\pi)^{3N/2} \int e^{-\mathcal{U}(x)}\,dx$.

This observation motivates for a treatment that is customized for the Hilbert space $\mathcal{H} = L^2(\mu)$ setting, with $d\mu = \mathcal{M}_{eq}(x,v)dvdx$. Operators can now be considered as acting from $\mathcal{H}$ to $\mathcal{H}$. In this setting, the collision operator $\mathcal{C}$ is symmetric with $\mathcal{C} = A^*A$ for $A = G^{1/2}(x)\nabla_v$. With a bit of more work it can be shown that

$\mathcal{C}$ is in fact self-adjoint, see e.g. [24, 63]. The transport term $\mathcal{T}$ is antisymmetric ($\mathcal{T}^* = -\mathcal{T}$).

The domain of $\mathcal{C}$ is defined by

$$\mathcal{D}(\mathcal{C}) = \{h \in L^2(\mu) | (-\nabla_v + v) \cdot G(x)\nabla_v h \in L^2(\mu)\}.$$

At the same time, the range of $\mathcal{C}$ is characterized by

$$\mathcal{R}(\mathcal{C}) = \{h \in L^2(\mu) | \int h\, d\mu = 0\}.$$

Operator L generates the continuous semigroup e^{-tL}. Since L can be written in the form $A^*A + B$ with B being antisymmetric, the semigroup e^{-tL} is non expansive in $\mathcal{H}$ as a result. Indeed, let $h = e^{-tL}h_0$

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\langle h, h\rangle &= \langle Lh, h\rangle = \langle (A^*A + B)h, h\rangle \\ &= -\langle Ah, Ah\rangle = -\|Ah\| \le 0\end{aligned}$$

Thus, $\|h\| \le \|h_0\|$ implying $\|e^{-tL}\|_{L^2(\mu)} \le 1$. This property, for instance, allows for short time estimates with $L^2(\mu)$ data to be essentially global in time estimates.

3.3 Well-Posedness

3.3.1 A priori Energy & Weak Formulation

In this section we give a weak formulation for the FP equation and an existence theory that is based on the notion of renormalized solutions (for kinetic equations with irregular coefficients) as first presented in [17].

The original work was initiated for proving existence of renormalized solutions to transport equations i.e.

$$\partial_t u + b(x) \cdot \nabla_x u = 0 \tag{3.3}$$

for a vector field $b(x) \in \mathbb{R}^n$. For this equation it is a trivial task to formally obtain an energy functional in L^p (for $p < \infty$) provided that

$$\nabla \cdot b(x) \in L^1(0, T; L^\infty(\mathbb{R}^n)) \quad \text{for} \quad p \geq 1.$$

Consider a function $\beta \in C^1(\mathbb{R})$ with $\beta' \in L^\infty$ and multiply the (3.3) by $\beta'(u)$. This yields the equation

$$\partial_t \beta(u) + b \cdot \nabla_x \beta(u) = 0$$

which if it admits a weak solution for every choice of β then this solution constitutes a so called renormalized solution. This is a step forward the already classical approach to distributional solutions because the existence of a renormalized solution in fact implies uniqueness of a distributional solution by a simple energy argument.

To study the L^p theory all one needs to do is multiply (3.3) by $\beta'(u) = u^{p-1}$.

After integration by parts one gets

$$\|u(t)\|_p \leq e^{CT}\|u(0)\|_p \quad \text{a.e. in } (0,T] \quad \text{for} \quad C > 0.$$

The energy inequality can be proven not just formally but also rigorously here. The inequality is enough to give weak compactness and assert the existence of a weak solution. Uniqueness is the difficult part of our theory and will be given in detail for our problem. We need a regularization type of argument to address uniqueness and this will require some extra conditions on the coefficients.

Going back to our example, in [49] & [50] there is an extension of the DiPerna-Lions theory to the FP equation of the type

$$\partial_t f + \nabla_x \cdot (bf) + \frac{1}{2}\nabla_x \cdot (\sigma\sigma^T \nabla_x f) = 0$$

for a general, possibly degenerate, matrix $\sigma(x) \in \mathbb{R}^{n\times n}$ and $b(x) \in \mathbb{R}^n$. We will prove the theorem for our equation but first give the details of the formulation.

There exists a variety of a priori energies we can pick for (3.1). As noted already, we will work on the $L^2_{\mathcal{M}_{eq}}$ framework. After multiplying the equation by $h\mathcal{M}_{eq}$ and integrating first in phase space and then in time we get,

$$\frac{1}{2}\int h^2(t,x,v)\,d\mu + \int_0^t \int |G^{1/2}(x)\nabla_v h(s,x,v)|^2\,d\mu\,ds = \frac{1}{2}\int h^2(0,x,v)\,d\mu.$$

Under the additional assumption $G(x) \geq \lambda \mathrm{I}, \quad \lambda > 0$, the energy estimate is

$$\frac{1}{2}\int h^2(t,x,v)\,d\mu + \lambda\int_0^t \int |\nabla_v h(s,x,v)|^2\,d\mu\,ds \leq \frac{1}{2}\int h^2(0,x,v)\,d\mu$$

which gives at least some hint of extra regularity.

So, if $f_0 \in L^2_{\mathcal{M}_{eq}}$ then the solution $f(t,x,v)$ will remain in $L^2_{\mathcal{M}_{eq}}$. At the same time the maximum principle implies that bounded initial data will remain bounded. Therefore, we seek to construct solutions for initial data $f_0 \in L^2_{\mathcal{M}_{eq}} \cap L^\infty(\mathbb{R}^{6N}_{x,v})$.

We are ready to proceed in the weak formulation to (3.1) with initial data $h(0,x,v)$ in $L^2_{\mathcal{M}_{eq}} \cap L^\infty(\mathbb{R}^{6N}_{x,v})$. We characterize h as a weak solution to (3.1), if

$$-\int_0^T \int h\partial_t \phi \, d\mu \, dt - \int h(0,\cdot)\phi(0,\cdot) \, d\mu + \int_0^T \int h \nabla \mathcal{U}(x) \cdot \nabla_v \phi \, d\mu \, dt$$
$$-\int_0^T \int hv \cdot \nabla_x \phi \, d\mu \, dt + \int_0^T \int \nabla_v h \cdot G(x) \nabla_v \phi \, d\mu \, dt = 0$$

for any smooth ϕ, compactly supported in $[0,T) \times \mathbb{R}^{6N}_{x,v}$. The weak formulation for (3.2) is similar.

The following theorem can be proven.

Theorem 4. *Assume that the potential $\mathcal{U}(x)$ and diffusion matrix $G^{1/2}(x)$ satisfy the following assumptions:*

$$(i) \quad G(x)v + \nabla \mathcal{U}(x) \in (W^{1,1}_{loc})^{3N} \qquad (ii) \quad tr(G) \in L^\infty$$
$$(iii) \quad \frac{G(x)v + \nabla \mathcal{U}(x)}{1 + |x| + |v|} \in (L^\infty)^{3N}$$
$$(iv) \quad G^{1/2}(x) \in (W^{1,2}_{loc})^{3N\times 3N} \qquad (v) \quad \frac{G^{1/2}(x)}{1+|x|} \in (L^\infty)^{3N\times 3N}.$$

Given initial data $f_0 \in L^2_{\mathcal{M}_{eq}} \cap L^\infty$ there exists a unique weak solution f to (3.1) *s.t.*

$$f \in L^\infty([0,T], L^2_{\mathcal{M}_{eq}} \cap L^\infty)$$

satisfying the additional condition $G(x)^{1/2} \nabla_v f \in (L^2_{\mathcal{M}_{eq}}([0,T], L^2_{\mathcal{M}_{eq}}))^{3N}$.

Remark 2. *$tr(G)$ is the trace of tensor $G(x)$. The above assumptions are pretty general in their nature and in many cases they become obsolete, e.g. for bounded*

$G(x)$ (ii) and (v) become trivial. Furthermore, for smooth $\mathcal{U}(x)$, $G(x)$ with the appropriate growth at infinity conditions they are all superficial. Assumption (ii) is needed for a general L^p theory (including L^2) but not for $L^2_{\mathcal{M}_{eq}}$. It is also worth mentioning that all the conditions are actually used for the uniqueness result since showing existence is the "easy" part and requires only bounded initial data.

For instance, let us for a moment consider the special case of *constant diffusion* ($G(x) = \mathrm{I}$). One may typically assume a C^1 potential $\mathcal{U}(x)$ that grows sufficiently fast at infinity. The initial data will be L^1, although the following result also applies for measure initial data. It can be shown,

Theorem 5. *Assume a potential $\mathcal{U}(x) \in C^1(\mathbb{R}^{3N}_x)$ with a uniformly bounded Hessian $|\nabla^2 \mathcal{U}(x)| \leq C$ for $C > 0$. Then the Cauchy problem*

$$\partial_t f + Lf = 0 \qquad f(0, \cdot, \cdot) = f_0 \in L^1(\mathbb{R}^{3N}_x \times \mathbb{R}^{3N}_v)$$

admits a unique solution $f(t, x, v) \in C(\mathbb{R}_+, \mathcal{D}'(\mathbb{R}^{3N}_x \times \mathbb{R}^{3N}_v))$ that satisfies the additional

$$f \in L^\infty_{loc}(\mathbb{R}_+, L^1(\mathbb{R}^{3N}_x \times \mathbb{R}^{3N}_v)) \cap L^2_{loc}(\mathbb{R}_+, H^1_v(\mathbb{R}^{3N}_x \times \mathbb{R}^{3N}_v)).$$

Proof. The proof is very similar in philosophy to the one that will be presented shortly for the general case. □

We take a note for later that the regularity in this case is similar to the one achieved for full rank (uniformly positive) diffusion matrices as we will show.

For now let us present the proof.

Proof. For the sake of a complete presentation we give the steps of the proof in the spirit of [50]. Since regularization plays a central role, we begin by defining a mollification kernel

$$\text{i.e.} \qquad \rho_\delta(x,v) = \frac{1}{\delta^{6N}}\rho\left(\frac{x}{\delta}\right)\rho\left(\frac{v}{\delta}\right)$$

where ρ is a smooth compactly supported ($\rho \in \mathcal{D}(\mathbb{R}^n)$), normalized ($\int \rho\, dx = 1$), nonnegative function ($\rho \geq 0$).

The **existence** part is the easiest. One assumes existence for smooth $\mathcal{U}(x)$, $G(x)$ holds and regularizes potential and diffusion matrix by $\mathcal{U}_\delta = \rho_\delta \star \mathcal{U}$, $G_\delta^{1/2} = \rho_\delta \star G^{1/2}$. Then, equation

$$\partial_t f + v \cdot \nabla_x f - \nabla \mathcal{U}_\delta(x) \cdot \nabla_v f = \nabla_v \cdot (G_\delta(x)(\nabla_v f + vf))$$

has a solution f_δ that depends on $\delta > 0$. Using the energy estimate, provided that $tr(G) \in L^\infty$, we can extract a subsequence of f_δ and pass to the limit (since we have convergence in $L^\infty(0,T;L^2(\mathbb{R}^{6N}_{x,v}))$) to obtain a solution in the weak sense.

The **uniqueness** is proven by showing short- in-time stability of the solution. Stability implies that a zero initial condition remains zero. Thus we need to establish an estimate of the form,

$$\|f\|_{L^2} \leq Ke^{Ct}\|f_0\|_{L^2}.$$

There are two main technical difficulties in proving the stability estimate. The first to be overcome is the need to do computations beyond the formal level. To do this, we regularize the equation by convoluting with ρ_δ, so that we obtain an equation for $f_\delta = \rho_\delta \star f$. This yields the regularized Fokker-Planck equation

$$\partial_t f_\delta + v \cdot \nabla_x f_\delta - \nabla \mathcal{U}(x) \cdot \nabla_v f_\delta - \nabla_v \cdot (G(x)(\nabla_v f_\delta + vf_\delta)) = U_\delta + \nabla_v \cdot (G^{1/2} R_\delta), \quad (3.4)$$

for

$$U_\delta = -[\rho_\delta, v\cdot\nabla_x - (\nabla\mathcal{U}(x) + Gv)\cdot\nabla_v](f) + [\rho_\delta, tr(G)](f)$$
$$+ [\rho_\delta, G^{1/2}\nabla_v](G^{1/2}\nabla_v f)$$

$$\text{and} \qquad R_\delta = [\rho_\delta, G^{1/2}\nabla_v](f).$$

The commutator $[\rho_\delta, c(x)\cdot\nabla]$ between a mollification function ρ_δ and a derivation is defined by

$$[\rho_\delta, c(x)\cdot\nabla] = \rho_\delta \star (c\cdot\nabla f) - c\cdot\nabla(\rho_\delta\star f)$$

where $c(x)$ is a vector field. It can actually be proven that in the limit $\delta\to 0$,

$$U_\delta \to 0 \quad \text{in} \quad L^\infty + L^2([0,T], L^1_{loc})$$
$$R_\delta \to 0 \quad \text{in} \quad L^\infty([0,T], L^2_{loc})$$

if the following conditions

$$Gv + \nabla\mathcal{U}(x) \in (W^{1,1}_{loc})^{3N}, \quad tr(G)\in L^\infty, \quad G^{1/2}\in (W^{1,2}_{loc})^{3N\times 3N}$$

hold.

The second fix in the proof is based on the fact that it is much simplified in bounded regions. The trick here is to consider a cut-off function $\phi_R(x,v) = \phi\left(\frac{x}{R}\right)\phi\left(\frac{v}{R}\right)$, where ϕ is a smooth function s.t. $\phi = 1$ in the ball of radius 1 and vanishes outside the ball of radius 2. If we multiply (3.4) by $f_\delta\phi_R$ and integrate in $\mathbb{R}^{6N}_{x,v}$, we get

$$
\begin{aligned}
&\frac{d}{dt}\int \frac{f_\delta^2}{2}\phi_R - \frac{1}{2}\int f_\delta^2 tr(G)\phi_R + \int |G^{1/2}\nabla_v f_\delta|^2\phi_R \\
&= \frac{1}{2}\int f_\delta^2(v\cdot\nabla_x\phi_R - (Gv + \nabla\mathfrak{U}(x))\cdot\nabla_v\phi_R) - \int f_\delta G^{1/2}\nabla_v f_\delta \cdot G^{1/2}\nabla_v\phi_R \\
&\qquad + \int U_\delta f_\delta\phi_R - \int (G^{1/2}\nabla_v f_\delta)\cdot R_\delta\phi_R - \int f_\delta(G^{1/2}\nabla_v\phi_R)\cdot R_\delta.
\end{aligned}
$$

The idea is to bound all the terms in r.h.s and send $R\to\infty$ uniformly in δ and see what terms vanish. The first two terms are

$$
\Big|\int f_\delta^2(v\cdot\nabla_x\phi_R - (Gv+\nabla\mathfrak{U})\cdot\nabla_v\phi_R)\Big| \le C\Big|\Big|\frac{G(x)v+\nabla\mathfrak{U}(x)}{1+|x|+|v|}\Big|\Big|_{L^\infty}\|\nabla\phi\|_{L^\infty}\int_{|x|,|v|\ge R} f_\delta^2
$$

and

$$
\Big|\int f_\delta G^{1/2}\nabla_v f_\delta\cdot G^{1/2}\nabla_v\phi_R\Big| \le C\Big|\Big|\frac{G(x)^{1/2}}{1+|x|}\Big|\Big|_{L^\infty}\|\nabla\phi\|_{L^\infty}\int_{|x|,|v|\ge R}|f_\delta||G^{1/2}\nabla_v f_\delta|.
$$

If assumptions (iii) & (v) are satisfied, then it can be shown that when $R\to\infty$ the two integrals above go to zero $\forall\delta\le 1$ (as well as L^1 in time). The other integrals can be bounded in similar manner. So, for any $\eta>0$, we can find large enough radius $R>0$ s.t. uniformly in $\delta>0$ the following holds

$$
\int \frac{f^2}{2}\phi_R \le \eta + C\int_0^t\int f^2\phi_R.
$$

Taking $R\to\infty$ proves the stability estimate.

□

3.3.2 Propagation of L^1 Initial Data

A more natural assumption for the choice of initial data is to be L^1 (for convenience we consider $L^1\cap L^\infty$). The study of a general L^p theory is possible. First, we consider a regular function β of one variable and we multiply (3.2) by $\beta'(f)$. After integration in the x, v variables we obtain:

$$\frac{d}{dt}\int \beta(f) - \int (f\beta'(f) - \beta(f))tr(G) + \int \beta''(f)|G(x)^{1/2}\nabla_v f|^2 = 0.$$

Consider a sequence of convex, regular functions β_n that converge to the absolute value. Given the fact that $G(x)$ is bounded it is implied that

$$\frac{d}{dt}\int |f| \leq C\int |f| \quad \Longrightarrow \quad \|f(t)\|_{L^1} \leq e^{CT}\|f_0\|_{L^1}$$

Thus, a unique weak solution for initial data $L^1 \cap L^\infty$ exists under the same conditions of previous theorem. The solution is now modified to belong in the space

$$f \in L^\infty([0,T], L^1 \cap L^\infty), \qquad G(x)^{1/2}\nabla_v f \in (L^1([0,T], L^1))^{3N}.$$

Chapter 4: Regularity

The question of Sobolev regularity is undertaken in this section. Due to the nice structure of the Fokker-Planck operator, one can anticipate regularization properties under certain plausible assumptions on the coefficients of L, despite the fact that L is a hypoelliptic operator (diffusive only in velocity). There are two types of regularity results treated here. Local (instantaneous) results, and short time regularity estimates. Estimates of the latter type are stronger and being coupled with the non expansivity of the semigroup e^{-tL} imply global regularization.

Starting with local results, we discuss only the basics of the language of pseudo-differential operators, stating Hörmander's theorem for operators of the type $\sum_{i=0}^{p} X_i^2 + X_0$ and Kohn's method of proof for this result. We shall also present an explicit local estimate for the problem with constant friction $\gamma > 0$.

We begin with exact estimates for the solution to the equation with a quadratic potential, based on the representation of the solution found in [15]. Next, we introduce certain entropies $\mathcal{E}(t, f)$ that imply immediate short time estimates, an idea first presented in [31]. We can apply this method for initial data in $L^2(\mu)$ and $LlogL(\mu)$, where μ is the unique stationary measure. C Villani in [71] gives a method of regularization which builds upon a system of differential inequalities.

This method gives regularity for unbounded initial data as one expects from problems of the kind. Here we apply the method for the slightly more general case of a diffusion $G(x) = \gamma(x)\mathrm{I}$. Finally, we present a result of regularization from L^1 data.

4.1 Local Regularity

4.1.1 Pseudodifferential Operators

The following is the statement of the Hörmander theorem valid for operators of type

$$L = \sum_{j=1}^{p} X_j^2 + X_0. \tag{4.1}$$

This form has been traditionally called Hörmander form of a second order differential operator.

Remark 3. *The Hörmander theorem is also valid for operators of the type* $L = \sum_{j=1}^{p} X_j^2$ *and* $L = \sum_{j=1}^{p} X_j^* X_j + X_0$*. The latter is treated by view of the observation that* $X_j^* = -X_j + c_j$ *for* $c_j \in C^\infty$*, with the adjoint being understood in the Hilbert setting* $\mathcal{H} = L^2(\mathbb{R}^n)$*.*

Here, the operators $X_0, X_1, \ldots, X_p$ are derivations (vector fields). Derivations are first order differential operators with C^∞ coefficients, i.e. $X_i = a_i(x) \cdot \nabla$ etc. Before we present Hörmander's theorem, we shall define commutators of the derivations using induction and Lie bracket notation,

$$X_{ij} = [X_i, X_j] = X_i X_j - X_j X_i$$

$$\ldots$$

$$X_{i_1 i_2 \ldots i_\alpha} = [X_{i_1}, [X_{i_2}, \ldots, [x_{i_{\alpha-1}}, X_{i_\alpha}] \ldots]].$$

Theorem 6. *For the operator* L*, mentioned in* (4.1)*, we say that* $X_0, X_1, \ldots, X_p$ *satisfy the Hörmander condition at a point* $x_0 \in \Omega$ *iff there exists some* $r(x_0) \in \mathbb{N}$

s.t the vector space generated by $X_{i_1 i_2 \ldots i_\alpha}$ at x_0 for $|\alpha| \leq r(x_0) - 1$ spans the whole tangent space. If the Hörmander condition is satisfied for all points in Ω (open set), it can be proven that

$$\|u\|_{1/r} \leq C(\|Lu\|_0^2 + \|u\|_0^2),$$

for some $r > 0$ and $\|u\|_s$ the typical H^s norm.

In [29] there is a proof of the above inequality based in an approach by J.J Kohn [43]. The proof by Kohn does not give the optimal exponent for $1/r$ in Hörmander's estimate, which is $1/3$, but rather the exponent $1/4$. For this proof, one needs to introduce the basics of the language of pseudo-differential operators.

The starting point is the introduction of the notion of a symbol. A symbol $p(x, \xi)$ of order m (real), is a function $p : \mathbb{R}^{2n} \ni (x, \xi) \to \mathbb{R}$ that can be expanded in homogeneous terms w.r.t. the ξ variable

$$p(x, \xi) \sim \sum_j p_{m-j}(x, \xi),$$

with $p_n(x, \xi)$ satisfying

$$p_n(x, \lambda\xi) = \lambda^n p_n(x, \xi) \qquad \text{for} \quad |\xi| \geq 1.$$

The operator P that corresponds to the above symbol

$$P(p)u(x) = \frac{1}{(2\pi)^n} \int e^{ix\cdot\xi} p(x, \xi) \hat{u}(\xi) d\xi,$$

is called a pseudo-differential operator of order m.

A more convenient characterisation of the symbol class $\mathbf{S}^m$ for symbols of order m is the following. A symbol $p(x, \xi) \in \mathbf{S}^m$ with $m \in \mathbb{R}$ if

$$|\partial_x^\alpha \partial_\xi^\beta p(x, \xi)| \leq C_{\alpha,\beta} \langle \xi \rangle^{m-|\beta|} \qquad \forall \alpha, \beta \in \mathbb{N} \cap \{0\},$$

where $\langle \xi \rangle$ stands for the "japanese bracket" symbol $\langle \xi \rangle = (1 + |\xi|^2)^{1/2}$.

Pseudo-differential operators are generalizations of differential operators. They form an algebra in the sense that a composition of two pseudo-differential operators of order m_1 and m_2 is a pseudo-differential operator of order $m_1 + m_2$. Notice that the symbol of the composition is NOT the product of the two symbols of the operators! More detailed, if $p_1 \in \mathbf{S}^{m_1}$ and $p_2 \in \mathbf{S}^{m_2}$, then

$$P(p_1)P(p_2) = P(p_1 p_2) + P(p_3)$$

where $p_3 \in \mathbf{S}^{m_1+m_2-1}$. This implies that the commutation $[P(p_1), P(p_2)]$ is an operator with symbol in $\mathbf{S}^{m_1+m_2-1}$. This observation lies in the heart of any proof of Hörmander's theorem that uses pseudo-differential calculus.

The adjoint (in L^2) of a pseudo-differential operator of order m, is a pseudo-differential operator of the same order, i.e. if $p \in \mathbf{S}^m$

$$P(p)^* = P(p) + P(q) \qquad \text{with} \quad q \in \mathbf{S}^{m-1}.$$

Pseudo-differential operators of order 0 form an algebra of bounded operators in $L^2(\mathbb{R}^n)$, i.e.

$$\|P(p)\| \leq C \qquad \text{if} \quad p \in \mathbf{S}^0.$$

In this family of operators of special importance are operators Λ^s that correspond to the symbol $p(x, \xi) = \langle \xi \rangle^s = (1 + |\xi|^2)^{s/2}$. Notice that the classical Sobolev space H^s can now be related to Λ^s by

$$H^s(\mathbb{R}^n) = \Lambda^{-s} L^2(\mathbb{R}^n).$$

Kohn proved the following as a step for showing hypoellipticity of the operator (4.1).

Theorem 7. *Consider the pseudo-differential operators of order* 0 *for which the following inequality holds*

$$\|Pu\|_\epsilon^2 \le C(\|Lu\|_0^2 + \|u\|_0^2), \tag{4.2}$$

for some $\epsilon > 0$ *and* $C > 0$*. The operators that satisfy the above inequality belong to a class which we denote by* $\mathcal{P}$*. If* $P \in \mathcal{P}$ *the following are satisfied:*

(i) $P^* \in \mathcal{P}$

(ii) $X_j\Lambda^{-1} \in \mathcal{P} \qquad j = 0, \ldots, p$

(iii)$[X_j, P] \in \mathcal{P} \qquad j = 0, \ldots, p$

(iv) $TP, PT \in \mathcal{P}$ *for any pseudo-differential operator* T *of order* 0.

In order to understand how this theorem combined with the Hörmander condition leads to hypoellipticity, we should start with the following simple observation. For proving regularity it suffices to prove (4.2) for $P = \mathrm{I}$ or P being the operator $F\Lambda^{-1}$ for all directions of derivations F. The Hörmander condition is satisfied if $X_0, X_1, \ldots, X_p$ create all directions which in turn with the help of Kohn's theorem proves that $F\Lambda^{-1} \in \mathcal{P}$ for all directions F. The rest is an iterative use of (4.2) that proves $u \in H^s$ for all $s \geq 0$.

To provide a common framework with results that will follow, we consider the Hilbert setting $\mathcal{H} = L^2(\mu)$, and operators of the form $L = A^*A + B$. We shall also consider the finite-dimensional Hilbert space $\mathcal{V}$ which will be the space of all variables on which an operator acts (typically we can think of it as $\mathbb{R}^n$). In the above abstract setting A is an operator $A : \mathcal{H} \longrightarrow \mathcal{H} \otimes \mathcal{V}$ and $B : \mathcal{H} \longrightarrow \mathcal{H}$.

This form has more structure and lies in the heart of many results shown later. To make a notational clarification, one can view A as an array of derivations, namely $A = (A_1, \ldots, A_n)$. We can now define commutations involving an array of derivations and a derivation, or two arrays of derivations in the following way:

$[A, B]$ will be viewed as the array $([A_1, B], [A_2, B], \ldots, [A_n, B])$, or with the help of tensorization as

$$[A, B] := AB - (B \otimes \mathrm{I})A.$$

On the other hand, $[A, A]$ will be the matrix of operators defined by $[A, A]_{i,j} = [A_i, A_j]_{i,j}$.

We have already stated the Fokker-Planck operator can be written in form $L = A^*A + B$, with

$$A = G^{1/2}(x)\nabla_v,\ A^* = -G^{1/2}(x)(\nabla_v - v) \quad \text{and} \quad B = v \cdot \nabla_x - \nabla \mathcal{U}(x) \cdot \nabla_v.$$

For $G(x) = \mathrm{I}$, $[A, B] = \nabla_x$, and the Hörmander condition is satisfied. The commutator algebra gets significantly more complicated for non constant diffusion $G(x)$. At the same time, it is worth noticing that local estimates tell us nothing about regularization in time. Therefore, the biggest part of this section will be devoted to answering the second question.

4.1.2 Algebraic Core

For simplicity let us consider the equation

$$\partial_t f + v \cdot \nabla_x f - \nabla \mathcal{U}(x) \cdot \nabla_v f = \gamma \nabla_v \cdot (\nabla_v f + v f)$$

with constant friction $\gamma > 0$ and initial data $f|_{t=0} = f_0$. The equation allows for a closed and more symmetric structure of the FP operator L. The potential $\mathcal{U}(x)$ satisfies

$$e^{-\mathcal{U}(x)} \in S(\mathbb{R}^n_x) \qquad \text{for} \qquad |\nabla^2 \mathcal{U}(x)| \leq C.$$

Remark 4. *The estimates that will follow are formal and can be proven for slightly weaker assumptions on the potential $\mathcal{U}(x)$. In [32] the potential assumed is the so called "high degree" potential that behaves like $\mathcal{U}(x) = |x|^{2m}$ (for $m \geq 1$) at infinity .*

After conjugating f with $\mathcal{M}_{eq}^{1/2}$ we have the new functions $h = \frac{f}{\mathcal{M}_{eq}^{1/2}}$ and $h_0 = \frac{f_0}{\mathcal{M}_{eq}^{1/2}}$. Operator L now takes the form $L = X_0 + \sum_j b_j^* b_j$ where X_0 is the field $X_0 = v \cdot \nabla_x - \nabla \mathcal{U}(x) \cdot \nabla_v$ described as in $X_0 = \gamma^{-1} \sum_j (a_j b_j^* - a_j^* b_j)$ and b_j^*, b_j the annihilation-creation pair.

The operators a_j, b_j are given by

$$a_j = \gamma^{1/2}(\partial_{x_j} - \frac{1}{2}\partial_{x_j}\mathcal{U}(x)), \qquad a_j^* = \gamma^{1/2}(-\partial_{x_j} + \frac{1}{2}\partial_{x_j}\mathcal{U}(x))$$
$$b_j = \gamma^{1/2}(\partial_v j + \frac{1}{2}v_j), \qquad b_j^* = \gamma^{1/2}(-\partial_v j + \frac{1}{2}v_j).$$

A more compact, vectorial notation, can be used with the introduction of operators

$$a = (a_1, \ldots, a_n)^T \qquad \text{and} \qquad b = (b_1, \ldots, b_n)^T$$

with adjoints $a^* = (a_1^*, \ldots, a_n^*)$, $b^* = (b_1^*, \ldots, b_n^*)$.

Operators a, b satisfy the canonical commutator relations

$$[b, b] = [b^*, b^*] = 0$$

$$[b, b^*] = \gamma Id \quad (\text{element-wise}[b, b^*]_{i,j} = [b_i, b_j^*] \quad \text{e.t.c})$$

as well as

$$[a, a] = [a^*, a^*] = 0$$

$$[a, a^*] = \gamma \nabla^2 \mathcal{U}(x).$$

Also a, b commute in the sense

$$[a^\sharp, b^\natural] = 0$$

for $\sharp$ and $\natural$ corresponding to either $*$ or nothing.

The Lie algebra structure is summarized by the following commutation relations between a, b and the field X_0, i.e.

$$[b, X_0] = a, \qquad [b^*, X_0] = a^* \qquad \text{and}$$

$$[a, X_0] = -\nabla^2 \mathcal{U}(x) b, \qquad [a^*, X_0] = -b^* \nabla^2 \mathcal{U}(x).$$

The natural Sobolev scaling is introduced with the help of the operator

$$\Lambda^2 = 1 + a^* a + b^* b.$$

This satisfies the relations,

$$[\Lambda^2, X_0] = -b^* (\nabla^2 \mathcal{U}(x) - \mathrm{I}) a - a^* (\nabla^2 \mathcal{U}(x) - \mathrm{I}) b$$

as can be easily checked, and

$$[\Lambda^m, (1+b^*b)^l] = 0 \qquad \text{for} \quad m, l \in \mathbb{R}.$$

It is proven, e.g. see [32] that

Theorem 8. *Under the assumptions that we made on $\mathcal{U}(x)$ there exists $C > 0$ s.t.*

$$\forall h \in \mathcal{S}(\mathbb{R}^{n,n}_{x,v}), \qquad \|\Lambda^\epsilon h\| \leq C_{\gamma,\mathcal{U}}(\|Lh\|^2 + \|h\|^2) \qquad 0 \leq \epsilon \leq \frac{1}{4}.$$

A proof of the above theorem can be given with the help of pseudo differential calculus and key hypoelliptic estimates. Here we are only giving a brief sketch of the proof based on the following estimate

$$\begin{aligned} \forall h \in \mathcal{S}(\mathbb{R}^{n,n}_{x,v}), \quad \|\Lambda^\epsilon h\|^2 &\leq \langle Lh, (M+M^*)h\rangle + \langle MLh, Mh\rangle \\ &\quad + C\left(\langle Lh, h\rangle + \|h\|^2\right), \qquad 0 \leq \epsilon \leq \frac{1}{4}. \end{aligned} \tag{4.3}$$

where $M = \Lambda^{2\epsilon-2}a^*b$, for some $C > 0$ that depends on γ and $\mathcal{U}(x)$.

Although we are not presenting the proof of the estimate which can in fact be found in [32], we will show how it implies the estimate in the above theorem. All terms in the r.h.s of the hypoelliptic estimate will be bounded by $\|Lh\|^2 + \|h\|^2$. We use the fact that $\Lambda^{2\epsilon-2}a^*$ and $a\Lambda^{2\epsilon-2}$ are bounded operators and that

$$\|\Lambda^{2\epsilon-2}a^*\| \quad \& \quad \|a\Lambda^{2\epsilon-2}\| \leq 1 \qquad \text{for} \quad 0 \leq \epsilon \leq \frac{1}{4}.$$

This implies immediately that

$$\|Mh\|^2 = \|\Lambda^{2\epsilon-2}a^*bh\|^2 \leq \|bh\|^2 = \langle Lh, h\rangle.$$

At the same time for the adjoint of M we have

$$M^* = b^* a\Lambda^{2\epsilon-2} = b^*(1+b^*b)^{-1/2}(1+b^*b)^{1/2}a\Lambda^{2\epsilon-2}$$

$$= b^*(1+b^*b)^{-1/2}a(1+b^*b)^{1/2}\Lambda^{2\epsilon-2} = b^*(1+b^*b)^{-1/2}a\Lambda^{2\epsilon-2}(1+b^*b)^{1/2}.$$

Since both operators $b^*(1+b^*b)^{-1/2}$ and $a\Lambda^{2\epsilon-2}$ are bounded, we have that

$$\|M^*h\|^2 \leq \|h\|^2 + \|bh\|^2.$$

With these estimates at hand, we get

$$2\langle Lh, (M+M^*)h\rangle = 2\langle Lh, Mh\rangle + 2\langle Lh, M^*h\rangle \leq 2\|Lh\|\,\|Mh\|$$

$$+2\|Lh\|\,\|M^*h\| \leq \|Lh\|^2 + \|Mh\|^2 + 2\|Lh\|\,(\|bh\|^2 + \|h\|^2)^{1/2}$$

$$\leq 2(\|Lh\|^2 + \|h\|^2) + \|Lh\|\,(\langle Lh, h\rangle + \|h\|^2)^{1/2}$$

$$\leq C(\|Lh\|^2 + \|h\|^2),$$

and

$$\langle MLh, Mh\rangle = \langle \Lambda^{2\epsilon-2}a^*bLh, \Lambda^{2\epsilon-2}a^*bh\rangle = \langle a\Lambda^{4\epsilon-4}a^*bLh, bh\rangle$$

$$\leq \|a\Lambda^{4\epsilon-4}a^*bLh\|\,\|bh\| \leq \|Lh\|\,\|bh\| \leq 2(\|Lh\|^2 + \|h\|^2).$$

The above two estimates and (4.3) prove the estimate in the theorem.

4.2 Short time & Global Regularity

4.2.1 Exact Regularity Estimates for the Quadratic Potential

Before we start with the functional related techniques for obtaining short time regularity estimates, we deviate a bit by extracting exact estimates for the solution in the case of a quadratic potential. We assume a quadratic potential plus smooth perturbations of it (typically $\omega_0 \frac{|x|^2}{2} + \Phi(x)$ for $\Phi(x) \in H^\infty(\mathbb{R}^{3N}_x)$, $\omega_0 > 0$), in the spirit of [15]. The friction matrix is assumed identity $G(x) = \mathrm{I}$.

As a part of a procedure of showing algebraic rates of convergence to the unique global equilibrium state for equation $\partial_t f + Lf = 0$ with initial data $f|_{t=0} = f_0$, the authors in [15] show that the solution propagates Sobolev regularity under the above assumptions for the potential $\mathcal{U}(x)$. This was done by constructing an exact solution taking the Fourier transform, and proving all the estimates with the help of the transformed solution. Here, we use the same solution to derive the short time regularity estimates. Unfortunately, the exact solution offers no insight on the long time behavior of the solution.

Consider the equation

$$\partial_t f + v \cdot \nabla_x f - x \cdot \nabla_v f = \nabla_v \cdot (\nabla_v f + vf). \tag{4.4}$$

We initially avoid using the perturbative part to simplify the computations a bit. Also, to make things slightly more presentable we let $n = 3N$. The Fourier transform

$\hat{f}(t,\xi,\eta)$ of $f(t,x,v)$ is defined by

$$\hat{f}(t,\xi,\eta) = \iint f(t,x,v)e^{-i(x\cdot\xi+v\cdot\eta)}\,dv\,dx,$$

with (ξ,η) being the conjugate variables of (x,v). The equation for the Fourier transform $\hat{f}$ is

$$\partial_t\hat{f} + \eta\cdot\nabla_\xi\hat{f} + (\eta-\xi)\cdot\nabla_\eta\hat{f} + |\eta|^2\hat{f} = 0. \tag{4.5}$$

The characteristic lines for the equation for $\hat{f}$ satisfy

$$\begin{pmatrix} \frac{d\xi}{dt} \\ \frac{d\eta}{dt} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix},$$

with assigned initial data $\xi(0)=\xi_0$, $\eta(0)=\eta_0$. The eigenvalues of the above system are $\frac{1}{2}\pm i\frac{\sqrt{3}}{2}$, with eigenvectors $\begin{pmatrix} 2 \\ 1 \end{pmatrix} \pm i \begin{pmatrix} 0 \\ \sqrt{3} \end{pmatrix}$. Subsequently, the characteristic system has the solution $(\xi(t) \quad \eta(t))^T = e^{t/2}X(t)(\xi_0 \quad \eta_0)^T$, for $X(t)$ the 2×2 matrix

$$X(t) = \begin{pmatrix} \cos\left(\frac{\sqrt{3}}{2}t\right) - \frac{1}{\sqrt{3}}\sin\left(\frac{\sqrt{3}}{2}t\right) & \frac{2}{\sqrt{3}}\sin\left(\frac{\sqrt{3}}{2}t\right) \\ -\frac{2}{\sqrt{3}}\sin\left(\frac{\sqrt{3}}{2}t\right) & \cos\left(\frac{\sqrt{3}}{2}t\right) + \frac{1}{\sqrt{3}}\sin\left(\frac{\sqrt{3}}{2}t\right) \end{pmatrix}.$$

This induces the characteristic flow $\chi_t(\xi,\eta) = (\chi_t^1(\xi,\eta), \chi_t^2(\xi,\eta))$ s.t.

$$\chi_t^1(\xi_0,\eta_0) = \xi(t) \qquad \text{and} \qquad \chi_t^2(\xi_0,\eta_0) = \eta(t).$$

The solution for $\hat{f}$ is given by

$$\hat{f}(t,\xi,\eta) = \hat{f}(0,\chi_{-t}^1(\xi,\eta),\chi_{-t}^2(\xi,\eta))e^{-\int_0^t |\chi_{-s}^2(\xi,\eta)|^2\,ds}.$$

Remark 5. *If we allow a perturbation $\Phi(x)$ on the quadratic potential, then the r.h.s of* (4.5) *now contains the extra term $i\eta\cdot\widehat{\nabla\Phi f}$, and the full solution to the* *transformed equation now reads*

$$
\begin{aligned}
\hat{f}(t,\xi,\eta) &= \hat{f}(0,\chi^1_{-t}(\xi,\eta),\chi^2_{-t}(\xi,\eta))e^{-\int_0^t |\chi^2_{-s}(\xi,\eta)|^2\,ds} \\
&\quad + i\int_0^t \chi^2_{-s}(\xi,\eta)\widehat{\nabla\Phi f}(t-s,\chi^1_{-s}(\xi,\eta),\chi^2_{-s}(\xi,\eta))e^{-\int_0^s |\chi^2_{-\sigma}(\xi,\eta)|^2\,d\sigma}\,ds.
\end{aligned}
$$

The following estimate proven in [15] gives control of $\hat{f}$. We present it here.

Lemma 1. *There exists some $K > 0$, s.t. for any $0 \leq t \leq 1$ and $(\xi,\eta) \in \mathbb{R}^n \times \mathbb{R}^n$, we have*

$$\int_0^t |\chi^2_{-s}(\xi,\eta)|^2\,ds \geq K(t^3|\xi|^2 + t|\eta|^2).$$

Proof. See paragraph 5 in [15]. □

Let us begin, for instance, with an estimate for $\|f\|_{L^2}$, that can be established with the help of the lemma above. For this estimate, we want to have some control of $\hat{f}_0(\cdot,\cdot) = \hat{f}(0,\cdot,\cdot)$ uniformly in ξ,η. This can be easily obtained, since by conservation of mass one has

$$\sup_{t\geq 0}\sup_{\xi,\eta\in\mathbb{R}^n} |\hat{f}(t,\xi,\eta)| \leq \|f_0\|_{L^1}.$$

Indeed, for a quadratic potential, the L^2 norm for a solution can be controlled explicitly by

$$
\begin{aligned}
\|f\|_{L^2}^2 &= \iint |f|^2\,dv\,dx = \iint |\hat{f}|^2\,d\eta\,d\xi \\
&= \iint |\hat{f}_0|^2 e^{-2\int_0^t |\chi_{-s}^2(\xi,\eta)|^2\,ds}\,d\eta\,d\xi \le \|f_0\|_{L^1}^2 \iint e^{-2K(t^3|\xi|^2+t|\eta|^2)}\,d\eta\,d\xi \\
&= \|f_0\|_{L^1}^2 \left(\int e^{-2Kt^3|\xi|^2}\,d\xi\right)\left(\int e^{-2Kt|\eta|^2}\,d\eta\right) \\
&= \|f_0\|_{L^1}^2 \left(\int_0^\infty \int_{\partial B(0,\rho)} e^{-2Kt^3\rho^2}\,dS\,d\rho\right)\left(\int_0^\infty \int_{\partial B(0,\rho)} e^{-2Kt\rho^2}\,dS\,d\rho\right) \\
&= \|f_0\|_{L^1}^2 |\partial B(0,1)|^2 \left(\int_0^\infty \rho^{n-1} e^{-2Kt^3\rho^2}\,d\rho\right)\left(\int_0^\infty \rho^{n-1} e^{-2Kt\rho^2}\,d\rho\right) \\
&= \|f_0\|_{L^1}^2 |\partial B(0,1)|^2 \frac{1}{2(2Kt^3)^{n/2}}\Gamma\left(\frac{n}{2}\right)\frac{1}{2(2Kt)^{n/2}}\Gamma\left(\frac{n}{2}\right) \\
&= \frac{\|f_0\|_{L^1}^2 |\partial B(0,1)|^2}{4(2K)^n t^{2n}}\left(\Gamma\left(\frac{n}{2}\right)\right)^2,
\end{aligned}
$$

where $\Gamma(\cdot)$ is the gamma function, and $\partial B(0,1)$ is the surface area of the n dimensional unit sphere.

The following theorem gives a precise short-time estimate, for the non-weighted Sobolev norm of the solution, for a potential of the type $\mathcal{U}(x) = \frac{x^2}{2} + \Phi(x)$.

Theorem 9. *Assume a smooth solution f to the equation*

$$
\partial_t f + v \cdot \nabla_x f - x \cdot \nabla_v f - \nabla\Phi(x) \cdot \nabla_v f = \nabla_v \cdot (\nabla_v f + vf),
$$

with initial data $f_0 \in L^1(\mathbb{R}^n \times \mathbb{R}^n)$, where the perturbative part $\Phi(x)$ satisfies $\Phi(x) \in H^k(\mathbb{R}^n)$ for all $k \ge 0$. Then, there exists $t_0 > 0$ s.t.

$$
\|f\|_{H^{m,l}_{x,v}} \le \frac{C}{t^{n+\frac{3}{2}m+\frac{1}{2}l}} \qquad \text{for} \quad 0 < t \le t_0 \quad \text{and} \quad C > 0.
$$

Proof. We begin by writing the solution for $\hat{f}$ in the form $\hat{f} = A(t,\xi,\eta) + iB(t,\xi,\eta)$, where the functions A and B have already been presented in Remark 5.

$$\|f\|^2_{H^{m,l}_{x,v}} = \iint |\nabla^m_x \nabla^l_v f|^2 \, dv \, dx = \iint |\xi|^{2m}|\eta|^{2l}|\hat{f}|^2 d\xi \, d\eta$$
$$\leq 2 \iint |\xi|^{2m}|\eta|^{2l}(|A(t,\xi,\eta)|^2 + |B(t,\xi,\eta)|^2) \, d\xi \, d\eta.$$

The computation of the first part in the above integral gives

$$\iint |\xi|^{2m}|\eta|^{2l}|A(t,\xi,\eta)|^2 \, d\eta \, d\xi = \iint |\xi|^{2m}|\eta|^{2l}|\hat{f}_0|^2 e^{-2\int_0^t |\chi^2_{-s}(\xi,\eta)|^2 \, ds} \, d\eta \, d\xi$$
$$\leq \|f_0\|^2_{L^1} \iint |\xi|^{2m}|\eta|^{2l} e^{-2K(t^3|\xi|^2 + t|\eta|^2)} \, d\eta \, d\xi$$
$$= \|f_0\|^2_{L^1} \left(\int |\xi|^{2m} e^{-2Kt^3|\xi|^2} \, d\xi \right) \left(\int |\eta|^{2l} e^{-2Kt|\eta|^2} \, d\eta \right)$$
$$= \|f_0\|^2_{L^1} \left(\int_0^\infty \int_{\partial B(0,\rho)} \rho^{2m} e^{-2Kt^3\rho^2} \, dS \, d\rho \right) \left(\int_0^\infty \int_{\partial B(0,\rho)} \rho^{2l} e^{-2Kt\rho^2} \, dS \, d\rho \right)$$
$$= \|f_0\|^2_{L^1} |\partial B(0,1)|^2 \left(\int_0^\infty \rho^{2m+n-1} e^{-2Kt^3\rho^2} \, d\rho \right) \left(\int_0^\infty \rho^{2l+n-1} e^{-2Kt\rho^2} \, d\rho \right)$$
$$= \frac{\|f_0\|^2_{L^1} |\partial B(0,1)|^2}{4(2K)^{n+m+l}} \frac{1}{t^{2n+3m+l}} \Gamma\left(m + \frac{n}{2}\right) \Gamma\left(l + \frac{n}{2}\right).$$

The second integral, $\int |\xi|^{2m}|\eta|^{2l}|B(t,\xi,\eta)|^2 \, d\xi \, d\eta$, will be shown to be bounded by a constant for sufficiently small values of $t > 0$. First, we begin with some estimates for $B(t,\xi,\eta)$ already pointed out in [15] par. 5.

Indeed, with the help of mass conservation and induction, it is proven that

$$\sup_{t \geq 0} |\widehat{\nabla \Phi f}(t,\xi,\eta)| \leq \frac{C_k}{(1 + |\xi|^2 + |\eta|^2)^k} \qquad \text{for} \quad C_k > 0, \quad \forall k \geq 0.$$

The above estimate is used in

$$|B(t,\xi,\eta)| \leq \int_0^t |\chi^2_s(\xi,\eta)| |\widehat{\nabla \Phi f}(t-s, \chi^1_{-s}(\xi,\eta), \chi^2_{-s}(\xi,\eta))| \, e^{-\int_0^s |\chi^2_{-\sigma}(\xi,\eta)| d\sigma} \, ds$$
$$\leq \int_0^t (s|\xi| + |\eta|) \frac{C_k}{(1 + |\chi^1_{-s}(\xi,\eta)|^2 + |\chi^2_{-s}(\xi,\eta)|^2)^k} \, e^{-K(s^3|\xi|^2 + s|\eta|^2)} \, ds.$$

By continuity of the characteristic lines, it follows that there exists $t_0 > 0$ s.t.

$$|\chi^1_{-s}(\xi,\eta)|^2 + |\chi^2_{-s}(\xi,\eta)|^2 \geq \frac{1}{2}(|\xi|^2 + |\eta|^2) \qquad \text{for} \quad s \leq t_0.$$

This implies

$$|B(t,\xi,\eta)| \leq \frac{C_k}{(1+|\xi|^2+|\eta|^2)^k} \int_0^t (s|\xi| + |\eta|) e^{-K(s^3|\xi|^2 + s|\eta|^2)}\, ds \qquad \text{for} \quad t \leq t_0.$$

The last integral is bounded, so this results to

$$|B(t,\xi,\eta)| \leq \frac{C'_k}{(1+|\xi|^2+|\eta|^2)^k} \qquad \text{for} \quad C'_k > 0, \quad \forall k \geq 0.$$

The above estimate holds for all $k \geq 0$. Given any choice for m and l, one can pick k large enough so that the integral $\int |\xi|^{2m}|\eta|^{2l}|B(t,\xi,\eta)|^2\, d\xi\, d\eta$ is bounded for $t \leq t_0$.

Combining the two estimates,

$$\int |\xi|^{2m}|\eta|^{2l}|A(t,\xi,\eta)|^2\, d\xi\, d\eta = O(1/t^{2n+3m+l}) \quad \text{and}$$

$$\int |\xi|^{2m}|\eta|^{2l}|B(t,\xi,\eta)|^2\, d\xi\, d\eta = O(1)$$

for $t \leq t_0$, finishes the proof.

$\square$

4.2.2 Héraou Method

The technique we are going to present in this section , with a slight modification, was first presented in [31]. It shows regularization from $L^2(\mu)$ to $H^1(\mu)$, where μ is the measure with density $\mathcal{M}_{eq}(x,v)$. It uses the energy functional

$$\mathcal{E}(t,h) := \int h^2\, d\mu + at \int |\nabla_v h|^2\, d\mu + 2bt^2 \int \nabla_v h \cdot \nabla_x h\, d\mu + ct^3 \int |\nabla_x h|^2\, d\mu, \quad (4.6)$$

which is shown to be dissipative for appropriate choice of values $a, b, c > 0$.

The Héraou technique has some advantages, but it also comes with a slight expense. Among the advantages is the fact that it makes use of a single functional that happens to be dissipative for certain values of parameters. It is also a very basic technique, since it relies solely on estimates that are based on the Cauchy-Schwartz and the Young inequalities. In the implementation of the method we have to assume a potential $\mathcal{U}(x) \in C^2(\mathbb{R}^{3N}_x)$ that has a bounded Hessian (e.g. quadratic potential etc). This assumption on the potential can be relaxed a bit, with a method that uses a system of inequalities to extract regularity estimates based on an approach by C. Villani.

Here we are about to generalize slightly on the method in [31], by assuming a smooth, diagonal diffusion matrix $G(x) = \gamma(x)\mathrm{I}$, with $\gamma(x), \nabla_x\gamma(x)$ bounded by

$$\lambda_0 \le \gamma(x) \le \Lambda_0 \qquad \text{for} \quad \lambda_0, \Lambda_0 > 0,$$

$$\text{and} \qquad |\nabla_x \gamma(x)| \le \Lambda_1 \qquad \text{for} \quad \Lambda_1 > 0.$$

The norm used above is the Hilbert-Schmidt norm i.e. $|\nabla_x\gamma(x)| = \sqrt{\sum_i |\partial_{x_i}\gamma(x)|^2}$. With all the above assumptions, we prove the following.

Theorem 10. *Let $\mathcal{U}(x) \in C^2(\mathbb{R}^{3N}_x)$ with $\inf \mathcal{U}(x) > -\infty$, having a bounded Hessian*

$$|\nabla^2 \mathcal{U}(x)| \leq C.$$

Assume a solution to the Fokker-Planck equation, with initial data $h_0 \in L^2(\mu)$. It can be proven that there exist parameters $a, b, c > 0$ (generally aligned as in $1 \gg a \gg b \gg c$) s.t.

$$\frac{d}{dt}\mathcal{E}(t,h) \leq 0 \qquad \text{for} \quad 0 \leq t \leq 1,$$

for the functional (4.6). *More specifically, it is shown that*

$$\int |\nabla_v h|^2 \, d\mu = O(t^{-1}), \qquad \int |\nabla_x h|^2 \, d\mu = O(t^{-3}) \qquad \text{for} \quad 0 < t \leq 1.$$

Proof. As stated already, we are using the energy functional (4.6) which we show it is dissipative for carefully selected parameters $a, b, c > 0$. More precisely we will show that there is a constant $K > 0$ s.t.

$$\begin{aligned}
&\frac{d}{dt}\mathcal{E}(t,h) \leq \\
&- K\left(\int |\nabla_v h|^2 \, d\mu + at \int |\nabla_v^2 h|^2 \, d\mu + bt^2 \int |\nabla_x h|^2 \, d\mu + ct^3 \int |\nabla_{vx}^2 h|^2 \, d\mu\right).
\end{aligned}$$

The result then follows by the form of the energy functional.

The derivative of the energy $\mathcal{E}(t,h)$ is

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}(t,h) = {} & \frac{d}{dt}\int h^2 \, d\mu + at\frac{d}{dt}\int |\nabla_v h|^2 \, d\mu \\
& + 2bt^2 \frac{d}{dt}\int \nabla_v h \cdot \nabla_x h \, d\mu + ct^3 \frac{d}{dt}\int |\nabla_x h|^2 \, d\mu \\
& + a\int |\nabla_v h|^2 \, d\mu + 4bt \int \nabla_v h \cdot \nabla_x h \, d\mu + 3ct^2 \int |\nabla_x h|^2 \, d\mu. \qquad (4.7)
\end{aligned}$$

Before we compute the estimates of the time derivatives of expression (4.7), we simply give the bound (based on Cauchy-Schwartz and Young inequalities) for the last line. Indeed,

$$
\begin{aligned}
& a \int |\nabla_v h|^2 \, d\mu + 4bt \int \nabla_v h \cdot \nabla_x h \, d\mu + 3ct^2 \int |\nabla_x h|^2 \, d\mu \\
& \leq a \int |\nabla_v h|^2 \, d\mu + 4b \int |\nabla_v h|^2 \, d\mu + bt^2 \int |\nabla_x h|^2 \, d\mu + 3ct^2 \int |\nabla_x h|^2 \, d\mu \\
& \leq (4b + a) \int |\nabla_v h|^2 \, d\mu + (b + 3c)t^2 \int |\nabla_x h|^2 \, d\mu.
\end{aligned}
$$

In general, we assume the ordering $a \gg b \gg c$ for parameters a, b, c, which somewhat simplifies the analysis. With this assumption, the above expression is bounded by $C \left(a \int |\nabla_v h|^2 \, d\mu + bt^2 \int |\nabla_x h|^2 \, d\mu \right)$ for $C > 1$.

The easy term in (4.7), is the time derivative of $\|h\|^2_{L^2(\mu)}$ which is controlled by

$$
\frac{1}{2} \frac{d}{dt} \int h^2 \, d\mu = - \int |G^{1/2}(x) \nabla_v h|^2 \, d\mu \leq -\lambda_0 \int |\nabla_v h|^2 \, d\mu. \tag{4.8}
$$

In order to compute the time evolution of the remaining norms, we use the following trick. Assume that we want to find the derivative of the norm $\|Ch\|_{L^2(\mu)}$ for a given derivation operator C. After we compute the action of the operator $\partial_t + L$ on the first order differential operator C i.e. $(\partial_t + L)Ch = [L, C]h$, we multiply by Ch and integrate in μ to get

$$
\frac{1}{2} \frac{d}{dt} \int |Ch|^2 \, d\mu + \int |ACh|^2 \, d\mu = \int Ch \cdot [L, C]h \, d\mu.
$$

The last integral term, $\int Ch \cdot [L, C]h \, d\mu$, should be controlled with further analysis.

The above computation has been performed in the appendix for $C = \nabla_v$ and $C = \nabla_x$, resulting in equations (4.9) & (4.11). For equation (4.10), we have performed the computation in a straightforward manner.

$$\frac{1}{2}\frac{d}{dt}\int |\nabla_v h|^2\, d\mu + \int |G^{1/2}(x)\nabla_v^2 h|^2\, d\mu = -\int \nabla_v h \cdot \nabla_x h\, d\mu$$
$$-\int \nabla_v h \cdot G(x)\nabla_v h\, d\mu, \tag{4.9}$$
$$\frac{d}{dt}\int \nabla_x h \cdot \nabla_v h\, d\mu = \int \nabla_v h \cdot \nabla^2 \mathfrak{U}(x)\nabla_v h\, d\mu - \int |\nabla_x h|^2\, d\mu$$
$$-\int \nabla_v h \cdot G(x)\nabla_x h\, d\mu - 2\int (G(x)\nabla_{vx}^2 h) : \nabla_v^2 h\, d\mu - \int \nabla_v^2 h : (\nabla_x G(x)\cdot \nabla_v h)\, d\mu, \tag{4.10}$$
$$\frac{1}{2}\frac{d}{dt}\int |\nabla_x h|^2\, d\mu + \int |G^{1/2}(x)\nabla_{vx}^2 h|^2\, d\mu = \int \nabla_x h \cdot \nabla^2 \mathfrak{U}(x)\nabla_v h\, d\mu$$
$$-\int \nabla_{vx}^2 h : (\nabla_x G(x)\cdot \nabla_v h)\, d\mu. \tag{4.11}$$

Each of the above derivatives will be treated separately. Let us only note here that notation-wise we have chosen the use of operators rather than present calculations componentwise. That way we avoid a heavy notation use. To be more specific, I have tried to keep the following conventions. The $\cdot$ symbol, as usual, stands for the dot product between two vectors, but it is also used when a third order tensor is multiplied by a vector to give a second order tensor. The symbol $:$ is the usual tensor product between second order tensors. The product of a second order tensor and a vector and that between two second order tensors is the usual one (matrix multiplication). Norms for a tensor of any order are the usual Hilbert-Schmidt norms. For the reader who wants to be meticulous about the computations with components, these have been thoroughly performed in the appendix.

The rate of change of the Hérau energy i.e (4.7) with the help of (4.8)-(4.11) is shown to be controlled by

$$
\begin{aligned}
\frac{d}{dt}\mathcal{E}(t,h) \leq -2\lambda_0 \int |\nabla_v h|^2 \, d\mu - 2\lambda_0 a t \int |\nabla_v^2 h|^2 \, d\mu \\
- 2bt^2 \int |\nabla_x h|^2 \, d\mu - 2\lambda_0 c t^3 \int |\nabla_{vx}^2 h|^2 \, d\mu \qquad (4.12) \\
+ 2bt^2 \int \nabla_v h \cdot \nabla^2 \mathcal{U}(x) \nabla_v h \, d\mu - 2\lambda_0 a t \int |\nabla_v h|^2 \, d\mu \qquad (4.13) \\
+ 2ct^3 \int \nabla_x h \cdot \nabla^2 \mathcal{U}(x) \nabla_v h \, d\mu - 2at \int \nabla_v h \cdot \nabla_x h \, d\mu \\
- 2bt^2 \int \gamma(x) \nabla_v h \cdot \nabla_x h \, d\mu \qquad (4.14) \\
- 4bt^2 \int \gamma(x) \nabla_{vx}^2 h : \nabla_v^2 h \, d\mu \qquad (4.15) \\
- 2bt^2 \int \nabla_x \gamma(x) \cdot \nabla_v^2 h \nabla_v h \, d\mu - 2ct^3 \int \nabla_x \gamma(x) \cdot \nabla_{vx}^2 h \nabla_v h \, d\mu \qquad (4.16) \\
+ (4b + a) \int |\nabla_v h|^2 \, d\mu + (b + 3c) t^2 \int |\nabla_x h|^2 \, d\mu. \qquad (4.17)
\end{aligned}
$$

We have grouped the terms in $\frac{d}{dt}\mathcal{E}(t,h)$ in a way that all the terms in first line i.e (4.12) have a sign, and will be used to dominate the terms in the lines (4.13)-(4.17) that follow. The conditions that suffice for such a control will be unveiled in this study. We begin for instance with (4.17). The estimate that appears in (4.17), which is the first estimate of terms contained in $\frac{d}{dt}\mathcal{E}(t,h)$ that we obtained, can be trivially controlled by terms in (4.12) for $\lambda_0 > a \, (H1)$. Notice that since we look for values of a, b with the ordering $a \gg b$, no extra condition for b is necessary. Another remark to be made here is that since $t \in [0,1]$, we are going to use in many occasions the fact that $t^3 \leq t^2 \leq t \leq 1$ without further notice.

Next is (4.13). The second integral term in (4.13) has a sign, so rightfully we focus on the first term. This term is

$$\leq 2bC \int |\nabla_v h|^2 \, d\mu \leq \lambda_0 \int |\nabla_v h|^2 \, d\mu,$$

if $b \leq \frac{\lambda_0}{2C}$ $(H2)$.

For the terms in (4.14), one has

$$\begin{aligned}(4.14) &\leq (2cC + 2a + 2\Lambda_0 b)t \int |\nabla_v h||\nabla_x h| \, d\mu \\ &\leq \frac{b}{2}t^2 \int |\nabla_x h|^2 \, d\mu + \frac{(2cC + 2a + 2\Lambda_0 b)^2}{2b} \int |\nabla_v h|^2 \, d\mu,\end{aligned}$$

which can be controlled by terms in (4.12) if $\frac{(2cC+2a+2\Lambda_0 b)^2}{2b} \leq \lambda_0$ $(H3)$. This practically implies that a^2/b should be sufficiently small.

As for (4.15),

$$\begin{aligned}(4.15) &\leq 4b\Lambda_0 t^2 \int |\nabla^2_{vx} h||\nabla^2_v h| \, d\mu \\ &\leq 2b\Lambda_0 \left(\frac{\lambda_0 c}{2\Lambda_0 b} t^3 \int |\nabla^2_{vx} h|^2 \, d\mu + \frac{2\Lambda_0 b}{\lambda_0 c} t \int |\nabla^2_v h|^2 \, d\mu \right),\end{aligned}$$

which is dominated by (4.12) if $\frac{b^2}{ac} \leq \frac{\lambda_0^2}{4\Lambda_0^2}$ $(H4)$.

Finally, let's treat (4.16). This line is comprised of 2 integral terms. The first one is controlled by

$$\text{1st integral in (4.16)} \leq b\Lambda_1 \left(\int |\nabla_v h|^2 \, d\mu + t \int |\nabla^2_v h|^2 \, d\mu \right),$$

which is trivially dominated by the terms in (4.12) if $\frac{b}{a} \leq \frac{\lambda_0}{\Lambda_1}$ $(H5)$.

Concluding with the second integral term in (4.16), one has

$$\begin{aligned}\text{2nd integral in (4.16)} &\leq 2c\Lambda_1 t^3 \int |\nabla^2_{vx} h||\nabla_v h| \, d\mu \\ &\leq c\Lambda_1 \left(\frac{\lambda_0 c}{2c\Lambda_1} t^3 \int |\nabla^2_{vx} h|^2 \, d\mu + \frac{2c\Lambda_1}{\lambda_0 c} t \int |\nabla_v h|^2 \, d\mu \right),\end{aligned}$$

which is also controlled by terms in (4.12) if $c \leq \frac{\lambda_0^2}{\Lambda_1^2}$ $(H6)$.

The last step in the proof, is to put down all the assumptions $(H1) - (H6)$ that a, b, c need to satisfy so that $\mathcal{E}(t, h)$ is dissipative. All these assumptions boil down to our ability to pick $a, b, c > 0$ so that the quantities a, $\frac{b}{a}$, $\frac{c}{b}$, $\frac{b^2}{ac}$, and $\frac{a^2}{b}$ can be made sufficiently small, for a given choice of the constants $\lambda_0, \Lambda_0, \Lambda_1, C > 0$.

This becomes a trivial task, once we can prove that we are able to send all these quantities to 0 at the same times. Take for instance the choice of sequences $a_j = 2^{-j}$, $b_j = 2^{-\frac{5}{3}j}$, and $c_j = 2^{-2j}$. Then $b_j/a_j = 2^{-\frac{2}{3}j}$, $c_j/b_j = 2^{-\frac{1}{3}j}$, $b_j^2/(a_j c_j) = 2^{-\frac{1}{3}j}$, and $a_j^2/b_j = 2^{-\frac{1}{3}j}$, which finishes the proof (pick j large enough).

$\square$

Remark 6. *An important point to mention is the necessity to include a mixed derivative term* $\int \nabla_x h \cdot \nabla_v h \, d\mu$ *in the energy functional. The main reason behind this choice is the fact that the dissipation of the mixed derivative contains the term* $\int |\nabla_x h|^2 \, d\mu$, *which isn't seen in* $\frac{d}{dt} \int |\nabla_x h|^2 \, d\mu$, *nor in* $\frac{d}{dt} \int |\nabla_v h|^2 \, d\mu$. *At the same time, the mixed term can be trivially controlled by a combination of* $\int |\nabla_x h|^2 \, d\mu$ *and* $\int |\nabla_v h|^2 \, d\mu$ *(with a simple Cauchy-Schwartz). This means that with proper manipulations, like the ones presented here, the inclusion of this term in* $\mathcal{E}(t, h)$ *plays a central role in proving that* $\mathcal{E}(t, h)$ *dissipates. As we are going to see, this is a technical trick which is not just important for regularization, but also important in convergence results. In fact, it will be shown that hypoelliptic regularization and hypoelliptic coercivity (hypocoercivity) are phenomena very similar in nature that can be treated with similar techniques.*

4.2.3 Regularization in the Entropic Sense

The technique that we described can lend itself to the treatment of solutions that belong in the $L \log L(\mu)$ space. Indeed, in many physically relevant cases the initial data does not belong in the weighted $L^2(\mu)$ space but has finite entropy. Unfortunately, the Hérau method cannot work in the entropic case, unless we assume $G(x) = \mathrm{I}$. This part will be explained in more detail later. For now, let us present what will serve as the Lyapunov functional in this case, which is

$$\begin{aligned}\mathcal{E}(t,h) = \int h \log h \, d\mu + at \int h |\nabla_v \log h|^2 \, d\mu + 2bt^2 \int h (\nabla_v \log h \cdot \nabla_x \log h) \, d\mu \\ + ct^3 \int h |\nabla_x \log h|^2 \, d\mu. \qquad (4.18)\end{aligned}$$

One cannot help but notice the striking resemblance in the form of this functional with the one presented for the $L^2(\mu)$ theory. Indeed, the relevant Sobolev norms have the form of integrals $\int \frac{|Ch|^2}{h} \, d\mu$, for a first order differential operator C. Of particular importance is the mixed term $\int \frac{\nabla_x h \cdot \nabla_v h}{h} \, d\mu$ which plays the same role as in the $L^2(\mu)$ theory, i.e. when differentiated along the semi-group e^{-tL} it provides the important $-\int \frac{|\nabla_x h|^2}{h} \, d\mu$ term for the dissipation of $\mathcal{E}(t,h)$.

The resemblance is even more obvious when it comes to the proof of dissipation for $\mathcal{E}(t,h)$. One is occupied with the task of bounding each term that appears in $\frac{d}{dt}\mathcal{E}(t,h)$ with terms of the type

$$\int h |\nabla_v \log h|^2 \, d\mu, \quad t \int h |\nabla_v^2 \log h|^2 \, d\mu, \quad t^2 \int h |\nabla_x \log h|^2 \, d\mu, \text{ and}$$
$$t^3 \int h |\nabla_{vx}^2 \log h|^2 \, d\mu.$$

Next, we make sure that there exists a choice of $a, b, c > 0$, so that when we sum all the terms in $\frac{d}{dt}\mathcal{E}(t,h)$, a combination of the four terms above with a negative sign appears making $\mathcal{E}(t,h)$ non increasing in $0 \leq t \leq 1$.

The precise statement of this result is :

Theorem 11. *Consider the potential $\mathcal{U}(x) \in C^2(\mathbb{R}^{3N}_x)$ with $\inf \mathcal{U}(x) > -\infty$, having a bounded Hessian*

$$|\nabla^2 \mathcal{U}(x)| \leq C.$$

Assume also a solution h (with $\int h_0 \, d\mu = 1$) to the F-P equation

$$\partial_t h + v \cdot \nabla_x h - \nabla \mathcal{U}(x) \cdot \nabla_v h = \triangle_v h - v \cdot \nabla_v h,$$

having finite initial entropy $\int h_0 \log h_0 \, d\mu < \infty$. It can be shown, that for the energy functional (4.18) *and for a certain choice of the parameters $a, b, c > 0$, we have*

$$\frac{d}{dt}\mathcal{E}(t,h) \leq 0 \qquad \text{for} \quad 0 \leq t \leq 1.$$

More specifically, it is shown that

$$\int h|\nabla_v \log h|^2 \, d\mu \leq O(t^{-1}), \qquad \int h|\nabla_x \log h|^2 \, d\mu \leq O(t^{-3}) \qquad \text{for} \quad 0 < t \leq 1.$$

Proof. As explained already, the main goal will be to find $a, b, c, K > 0$, s.t. for $0 \leq t \leq 1$, the following holds

$$\frac{d}{dt}\mathcal{E}(t,h) \leq -K \left(\int h|\nabla_v \log h|^2 \, d\mu + at \int h|\nabla_v^2 \log h|^2 \, d\mu \right.$$
$$\left. + bt^2 \int h|\nabla_x \log h|^2 \, d\mu + ct^3 \int h|\nabla_{vx}^2 \log h|^2 \, d\mu \right).$$

Starting with the time derivative of $\mathcal{E}(t,h)$, one has

$$\begin{aligned}\frac{d}{dt}\mathcal{E}(t,h) &= \frac{d}{dt}\int h\log h\,d\mu + at\frac{d}{dt}\int h|\nabla_v \log h|^2\,d\mu\\ &+ 2bt^2\frac{d}{dt}\int h(\nabla_v\log h\cdot\nabla_x\log h)\,d\mu + ct^3\frac{d}{dt}\int h|\nabla_x\log h|^2\,d\mu\\ &+ a\int h|\nabla_v\log h|^2\,d\mu + 4bt\int h(\nabla_v\log h\cdot\nabla_x\log h)\,d\mu + 3ct^2\int h|\nabla_x\log h|^2\,d\mu.\end{aligned} \tag{4.19}$$

In similar manner like before, the last line in the equation above is bounded above by

$$(4b+a)\int h|\nabla_v\log h|^2\,d\mu + (b+3c)t^2\int h|\nabla_x\log h|^2\,d\mu.$$

It remains to treat the r.h.s terms in the first two lines of (4.19).

We start with the easy term, which is the evolution of the entropic term (in math literature, this entropic term is often addressed with the name *Kullback* entropy functional, and its rate of change with a "$-$" sign as the *Fisher* information) i.e.

$$\frac{d}{dt}\int h\log h\,d\mu = -\int\frac{|\nabla_v h|^2}{h}\,d\mu = -\int h|\nabla_v\log h|^2\,d\mu. \tag{4.20}$$

The evolution of the remaining integral terms is more complex in nature but can be studied in an abstract setting. Here we follow the computations performed in [71] where the differentiation is performed along the semi-group e^{-tL}. In short, it is shown that

$$
\begin{aligned}
-\frac{d}{dt}\int h\left(C\log h\cdot C'\log h\right)\,d\mu = \\
&\int h\left(C\log h\cdot [C',B]\log h\right)\,d\mu + \int h\left([C,B]\log h\cdot C'\log h\right)\,d\mu \\
&+2\int h\left(CA\log h\cdot C'A\log h\right)\,d\mu \\
&+\int h\left([C,A^*]A\log h\cdot C'\log h\right)\,d\mu + \int h\left(C\log h\cdot [C',A^*]A\log h\right)\,d\mu \\
&+\int h\left(CA\log h\cdot [A,C']\log h\right)\,d\mu + \int h\left([A,C]\log h\cdot C'A\log h\right)\,d\mu \\
&+\int h\left([A,C]^*(A\log h\otimes C'\log h)\right)\,d\mu + \int h\left([A,C']^*(A\log h\otimes C\log h)\right)\,d\mu.
\end{aligned}
$$

where C, C' are first order differential operators of the type $a(x)\cdot\nabla$, with $a(x)$ being a smooth vector field with derivatives that grow at most polynomially. The convention in the notation used above is that $\otimes$ stands for the dyadic product, and the action of an operator on another described by

$$
[C,A^*]A = \sum_j [C,A_j^*]A_j, \qquad CAu\cdot[A,C']u = \sum_{i,j} C_iA_ju[A_i,C_j']u,
$$

$$
CAu\cdot C'Au = \sum_{i,j} C_iA_juC_i'A_ju \quad \text{and} \quad [A,C]^*(Au\otimes C'u) = \sum_{i,j}[A_i,C_j]^*(A_iuC_j'u).
$$

More precisely, it is shown with the help of the above, that the evolution of the log Sobolev norms is governed by the equations

$$
\begin{aligned}
-\frac{1}{2}\frac{d}{dt}\int h|\nabla_v \log h|^2\, d\mu &= \int h(\nabla_v \log h \cdot \nabla_x \log h)\, d\mu \\
&\quad + \int h|\nabla_v^2 \log h|^2\, d\mu + \int h|\nabla_v \log h|^2\, d\mu, \qquad (4.21)\\
-\frac{1}{2}\frac{d}{dt}\int h|\nabla_x \log h|^2\, d\mu &= -\int h(\nabla_x \log h \cdot \nabla^2 \mathcal{U}(x) \nabla_v \log h)\, d\mu \\
&\quad + \int h|\nabla_{vx}^2 \log h|^2\, d\mu, \qquad (4.22)\\
-\frac{d}{dt}\int h\left(\nabla_x \log h \cdot \nabla_v \log h\right)\, d\mu &= \int h|\nabla_x \log h|^2\, d\mu \\
-\int h(\nabla_v \log h \cdot \nabla^2 \mathcal{U}(x) \nabla_v \log h)\, d\mu &+ 2\int h(\nabla_{vx}^2 \log h : \nabla_v^2 \log h)\, d\mu \\
&\quad + \int h(\nabla_v \log h \cdot \nabla_x \log h)\, d\mu. \qquad (4.23)
\end{aligned}
$$

It is now evident that the idea that was implemented in the proof in the previous section actually works in this case without any real change. Equations (4.20)-(4.23) above, when inserted in the equation for $\frac{d}{dt}\mathcal{E}(t,h)$, will give a number of integrals dominated by just 4 of these integral terms which have a (negative) sign. For instance, from (4.20) comes the contribution $-\int h|\nabla_v \log h|^2\, d\mu$. From (4.21) comes the contribution $-\int h|\nabla_v^2 \log h|^2\, d\mu$. From (4.22) comes the contribution $-\int h|\nabla_{vx}^2 \log h|^2\, d\mu$ (contained in the second line of (4.22)), and finally from (4.23) comes the main contribution $-\int h|\nabla_x \log h|^2\, d\mu$. The remaining of the integral terms in the above equations are dominated by the main contributions stated above (if one uses the C-S and Young inequalities with appropriate choice of $a, b, c > 0$). In detail,

$$
\begin{aligned}
\frac{d}{dt}\mathcal{E}(t,h) \leq &-\int h|\nabla_v \log h|^2\, d\mu - 2at\int h|\nabla_v^2 \log h|^2\, d\mu \\
&\qquad\qquad - 2bt^2\int h|\nabla_x \log h|^2\, d\mu - 2ct^3\int h|\nabla_{vx}^2 \log h|^2\, d\mu \\
&+ 2bt^2\int h(\nabla_v \log h\cdot \nabla^2\mathcal{U}(x)\nabla_v \log h)\, d\mu - 2at\int h|\nabla_v \log h|^2\, d\mu \\
&- (2at+2bt^2)\int h(\nabla_v \log h\cdot \nabla_x \log h)\, d\mu + 2ct^3\int h(\nabla_x \log h\cdot \nabla^2\mathcal{U}(x)\nabla_v \log h)\, d\mu \\
&- 4bt^2\int h(\nabla_{vx}^2 \log h : \nabla_v^2 \log h)\, d\mu \\
&+ (4b+a)\int h|\nabla_v \log h|^2\, d\mu + (b+3c)t^2\int h|\nabla_x \log h|^2\, d\mu.
\end{aligned}
$$

It turns out that the choice of parameters $a, b, c > 0$ that we made in the $H^1(\mu)$ regularization, i.e., choosing the values of a, $\frac{b}{a}$, $\frac{c}{b}$, $\frac{a^2}{b}$, and $\frac{b^2}{ac}$ to be sufficiently small, is good enough to prove dissipation for $\mathcal{E}(t,h)$.

$\square$

Remark 7. *A brief explanation of why the above method fails for a general diffusivity $G(x)$ is the following. The short answer is that the more complicated structure of the commutator algebra does not allow for good log Sobolev bounds. If we assume A, B to be first order derivation operators, then bounds for $\int \frac{Ah\cdot Bh}{h}\, d\mu$ are achieved by the simple inequality*

$$
\int \frac{Ah\cdot Bh}{h}\, d\mu \leq \left(\int \frac{|Ah|^2}{h}\, d\mu\right)^{1/2}\left(\int \frac{|Bh|^2}{h}\, d\mu\right)^{1/2}.
$$

Now, if we want to continue taking advantage of such expressions and continue e.g. with an inequality of the type

$$
\left(\int \frac{|Ah|^2}{h}\, d\mu\right)^{1/2} \leq C_1\left(\int \frac{|R_1 h|^2}{h}\, d\mu\right)^{1/2} + \ldots + C_k\left(\int \frac{|R_k h|^2}{h}\, d\mu\right)^{1/2},
$$

for some integer $k \geq 1$ and $C_i > 0$, then A as an operator should be bounded pointwise by $R_1, \ldots, R_k$ and not just bounded in the operator sense like

$$\|Ah\|_{L_2(\mu)} \leq C_1 \|R_1 h\|_{L_2(\mu)} + \ldots + C_k \|R_k h\|_{L_2(\mu)}.$$

This makes it impossible, for instance, for exact estimates for an integral like $\int \frac{|v \cdot \nabla_x h|^2}{h} \, d\mu$, whereas it is true on the other hand that

$$\int |v \cdot \nabla_x h|^2 \, d\mu \leq C \left(\int |\nabla_x h|^2 \, d\mu + \int |\nabla_{vx}^2|^2 \, d\mu \right).$$

The last comment that we make here is to explain exactly what we mean when saying that a derivation operator A is bounded pointwise by $R_1, \ldots, R_k$. In brief, by writing A as $A = a(x) \cdot \nabla = \sum_j a_j(x) \partial_j$, and $R_i = r_i(x) \cdot \nabla = \sum_j r_{ij}(x) \partial_j$, we then say that A is bounded pointwise by $R_1, \ldots, R_k$ iff

$$a_j(x) \leq \sum_{i=1}^{k} C_i |r_{ij}(x)|, \qquad \text{for} \quad C_i > 0, \quad \forall x \in \mathbb{R}^n \text{ and } \quad \forall j.$$

4.2.4 Hypoellipticity à la Villani

We now present a different approach based on a method developed by C.Villani and discussed in [71]. In this approach, instead of dealing directly with an energy functional, one tries to construct a system of differential inequalities to derive the short time (hypoelliptic) estimates. This method offers the advantage that it relaxes the assumption on the potential $\mathcal{U}(x)$, and most importantly that it offers the optimal estimate in H_v^3 instead of H_v^1 as was the case using the Hérau technique. As shown here, estimates of H_v^3 imply estimates for H_v^1 by a standard interpolation argument. The assumptions for the control of $\gamma(x)$ are the same as in the previous section.

Theorem 12. *Assume a potential $\mathcal{U}(x) \in C^2(\mathbb{R}_x^{3N})$ with $\inf \mathcal{U}(x) > -\infty$, satisfying the condition*

$$|\nabla^2 \mathcal{U}(x)| \leq C(1 + |\nabla \mathcal{U}(x)|) \qquad \text{for} \quad C > 0.$$

It can be shown that the solution h of the Fokker-Planck equation with bounded in $L^2(\mu)$ initial data satisfies the following regularity estimates when $0 \leq t \leq 1$,

$$\int |\nabla_x h|^2 \, d\mu = O(t^{-3}) \qquad \text{and} \qquad \int |\nabla_v^3 h|^2 \, d\mu = O(t^{-3}).$$

Let us make a couple of observations before we start with the proof of the theorem.

Remark 8. *The first one is related to the fact that $|\nabla \mathcal{U}(x)|^2$ defines a bounded operator from $H^1(\mu) \to L^2(\mu)$ (as long as $|\nabla^2 \mathcal{U}(x)|$ is dominated by $|\nabla \mathcal{U}(x)|$), which* *implies*

$$\int |\nabla \mathcal{U}(x)|^2 g^2 \, d\mu \le C' \left(\int g^2 \, d\mu + \int |\nabla_x g|^2 \, d\mu \right) \tag{4.24}$$

$$\text{and} \qquad \int |\nabla^2 \mathcal{U}(x)|^2 g^2 \, d\mu \le C' \left(\int g^2 \, d\mu + \int |\nabla_x g|^2 \, d\mu \right). \tag{4.25}$$

A second estimate we will use in many instances is

$$\int |v|^2 g^2 \, d\mu \le C' \left(\int g^2 \, d\mu + \int |\nabla_v g|^2 \, d\mu \right) \qquad \text{for} \quad C' > 0. \tag{4.26}$$

The constant C' depends on C, N only (and only on N for (4.26)).

Proof. A straightforward computation gives

$$\begin{aligned}
&\int |\nabla \mathcal{U}(x)|^2 g^2 e^{-\mathcal{U}(x)} \, dx = - \int \nabla_x (e^{-\mathcal{U}(x)}) \cdot \nabla \mathcal{U}(x) g^2 \, dx \\
&= \int \nabla_x \cdot (\nabla \mathcal{U}(x) g^2) e^{-\mathcal{U}(x)} \, dx = \int \triangle \mathcal{U}(x) g^2 e^{-\mathcal{U}(x)} \, dx + 2 \int g \nabla \mathcal{U}(x) \cdot \nabla_x g \, e^{-\mathcal{U}(x)} \, dx \\
&\le \frac{1}{12NC^2} \int |\triangle \mathcal{U}(x)|^2 g^2 e^{-\mathcal{U}(x)} \, dx + \frac{12NC^2}{4} \int g^2 e^{-\mathcal{U}(x)} \, dx \\
&\qquad + \frac{1}{4} \int |\nabla \mathcal{U}(x)|^2 g^2 e^{-\mathcal{U}(x)} \, dx + 4 \int |\nabla_x g|^2 e^{-\mathcal{U}(x)} \, dx \\
&\le \frac{6NC^2}{12NC^2} \int (1 + |\nabla \mathcal{U}(x)|^2) g^2 e^{-\mathcal{U}(x)} \, dx + 3NC^2 \int g^2 e^{-\mathcal{U}(x)} \, dx \\
&\qquad + \frac{1}{4} \int |\nabla \mathcal{U}(x)|^2 g^2 e^{-\mathcal{U}(x)} \, dx + 4 \int |\nabla_x g|^2 e^{-\mathcal{U}(x)} \, dx \\
&= \frac{3}{4} \int |\nabla \mathcal{U}(x)|^2 g^2 e^{-\mathcal{U}(x)} \, dx + \left(\frac{1}{2} + 3NC^2 \right) \int g^2 e^{-\mathcal{U}(x)} \, dx + 4 \int |\nabla_x g|^2 e^{-\mathcal{U}(x)} \, dx.
\end{aligned}$$

This sums up to

$$\int |\nabla \mathcal{U}(x)|^2 g^2 e^{-\mathcal{U}(x)} \, dx \le (2 + 12NC^2) \int g^2 e^{-\mathcal{U}(x)} \, dx + 16 \int |\nabla_x g|^2 e^{-\mathcal{U}(x)} \, dx,$$

which concludes the estimate after we multiply by $\mathcal{M}(v)$ and integrate in v.

The only point to be made about the above inequalities is that we have used the trivial inequality

$$|\triangle \mathcal{U}(x)|^2 \le 3N |\nabla^2 \mathcal{U}(x)|^2 \le 6NC^2 (1 + |\nabla \mathcal{U}(x)|^2).$$

The estimate (4.25) follows from

$$\int |\nabla^2 \mathcal{U}(x)|^2 g^2 e^{-\mathcal{U}(x)}\, dx \leq 2C^2 \int (1 + |\nabla \mathcal{U}(x)|^2) g^2 e^{-\mathcal{U}(x)}\, dx.$$

Similarly, for (4.26)

$$\begin{aligned}\int |v|^2 g^2 \mathcal{M}(v)\, dv &= -\int \nabla_v(\mathcal{M}(v)) \cdot v g^2\, dv = \int \nabla_v \cdot (v g^2) \mathcal{M}(v)\, dv \\ &= 3N \int g^2 \mathcal{M}(v)\, dv + 2\int g v \cdot \nabla_v g \mathcal{M}(v)\, dv \\ &\leq 3N \int g^2 \mathcal{M}(v)\, dv + \frac{1}{2}\int |v|^2 g^2 \mathcal{M}(v)\, dv + 2\int |\nabla_v g|^2 \mathcal{M}(v)\, dv,\end{aligned}$$

which yields

$$\int |v|^2 g^2 \mathcal{M}(v)\, dv \leq 6N \int g^2 \mathcal{M}(v)\, dv + 4\int |\nabla_v g|^2 \mathcal{M}(v)\, dv.$$

After multiplication by $e^{-\mathcal{U}(x)}$ and integration in x the desired result follows. $\square$

Remark 9. *(Interpolation) The second main observation has to do with justification of why this result is slightly stronger than the one encountered in the last paragraph. Here, we prove that* $\int |\nabla_v^3 h|^2\, d\mu = O(t^{-3})$ *coupled with* $\int h^2\, d\mu < \infty$ *implies* $\int |\nabla_v h|^2\, d\mu = O(t^{-1})$ *and* $\int |\nabla_v^2 h|^2\, d\mu = O(t^{-2})$*. This is in fact a straightforward interpolation result. As mentioned in [71] one can easily prove the interpolation inequality*

$$\int |\nabla_v^j h|^2\, d\mu \leq C \left(\int h^2\, d\mu\right)^{1-j/3} \left(\int |\nabla_v^3 h|^2\, d\mu\right)^{j/3} \qquad \text{for} \quad 0 \leq j \leq 3,$$

using Hermite polynomials. The n'th Hermite polynomial is defined by $H_n(v) = (-1)^n e^{\frac{v^2}{2}} \frac{d^n}{dv^n} e^{-\frac{v^2}{2}}$ *with* $n \geq 0$*. Hermite polynomials form an orthonormal basis in* $L^2(\mathcal{M}(v)dv)$ *i.e.*

$$\int H_n(v) H_m(v) \mathcal{M}(v)\, dv = \delta_{nm}.$$

The polynomials satisfy, among others, the property $H_n'(v) = 2nH_{n-1}(v)$. Consider the function $h(x,v)$ and its expansion in Hermite polynomials $h(x,v) = \sum_n c_n H_n(v)$, with coefficients c_n that depend on x in general. It follows that

$$\int |\nabla_v h|^2 \mathcal{M}(v)\, dv = \int \Big| \sum_n c_n H_n'(v) \Big|^2 \mathcal{M}(v)\, dv = \int \Big| \sum_n 2nc_n H_{n-1}(v) \Big|^2 \mathcal{M}(v)\, dv$$
$$= 4 \int \sum_n n^2 c_n^2 |H_{n-1}(v)|^2 \mathcal{M}(v)\, dv = 4 \sum_n n^2 c_n^2.$$

In similar fashion,

$$\int |\nabla_v^2 h|^2 \mathcal{M}(v)\, dv = 16 \sum_n n^4 c_n^2.$$

Integrating in $e^{-\mathcal{U}(x)}dx$ and applying the Hölder inequality yields the desired interpolation inequality.

Remark 10. *(Evolution of H_v^3 seminorm). The evolution of the mixed derivative, as well as the H_x^1 and H_v^3 semi-norms is used in the proof. The first two have already been presented in equations (II) & (IV). The evolution of H_v^3 is described by the equation*

$$\frac{1}{2}\frac{d}{dt} \int |\nabla_v^3 h|^2 \, d\mu + \int |G^{1/2}(x) \nabla_v^4 h|^2 \, d\mu = -3 \int |G^{1/2}(x) \nabla_v^3 h|^2 \, d\mu$$
$$- 3 \int \nabla_v^3 h \vdots \nabla_{xvv}^3 h \, d\mu. \tag{4.27}$$

We are now ready to proceed with the proof of the main result.

Proof. We begin with the first r.h.s integral of (4.11),

$$\begin{aligned}
&\int \nabla_x h \cdot \nabla^2 \mathcal{U}(x) \nabla_v h \, d\mu = -\int \nabla^2_{vx} h : \nabla^2 \mathcal{U}(x) h \, d\mu + \int \nabla_x h \cdot \nabla^2 \mathcal{U}(x) v h \, d\mu \\
&\leq \frac{5}{2\lambda_0} \int |\nabla^2 \mathcal{U}(x)|^2 h^2 \, d\mu + \frac{2\lambda_0}{20} \int |\nabla^2_{vx} h|^2 \, d\mu \\
&\qquad + \frac{5}{2\lambda_0} \int |\nabla^2 \mathcal{U}(x)|^2 h^2 \, d\mu + \frac{2\lambda_0}{20} \int |v|^2 |\nabla_x h|^2 \, d\mu \\
&\leq \frac{5}{\lambda_0} \left((2 + 12NC^2) \int h^2 \, d\mu + 16 \int |\nabla_x h|^2 \, d\mu \right) + \frac{\lambda_0}{10} \int |\nabla^2_{vx} h|^2 \, d\mu \\
&\qquad + \frac{\lambda_0}{10} \left(6N \int |\nabla_x h|^2 \, d\mu + 4 \int |\nabla^2_{vx} h|^2 \, d\mu \right) \\
&\leq \frac{5(2 + 12NC^2)}{\lambda_0} \int h^2 \, d\mu + \left(\frac{80}{\lambda_0} + \frac{3}{5} \lambda_0 N \right) \int |\nabla_x h|^2 \, d\mu + \frac{\lambda_0}{2} \int |\nabla^2_{vx} h|^2 \, d\mu.
\end{aligned}$$

It is not necessary to keep track of the optimal constants in computations like the above. One just has to make sure that the last integral term $\int |\nabla^2_{vx} h|^2 \, d\mu$ is multiplied by a constant strictly less than λ_0 (here $\frac{\lambda_0}{2}$).

The second integral in r.h.s of (4.11) is bounded by

$$-\int \nabla_x \gamma(x) \cdot \nabla^2_{vx} h \nabla_v h \, d\mu \leq \Lambda_1 \left(\frac{\Lambda_1}{\lambda_0} \int |\nabla_v h|^2 \, d\mu + \frac{\lambda_0}{4\Lambda_1} \int |\nabla^2_{vx} h|^2 \, d\mu \right).$$

The above two estimates, together with (4.11), can be combined to give

$$\frac{1}{2} \frac{d}{dt} \int |\nabla_x h|^2 \, d\mu \leq -\frac{\lambda_0}{4} \int |\nabla^2_{vx} h|^2 \, d\mu + C \left(\int (h^2 + |\nabla_x h|^2 + |\nabla_v h|^2) \, d\mu \right).$$

The integral term $\int |\nabla_v h|^2 \, d\mu$ in the estimate above is an extra term that does not appear in the estimate obtained by Villani when he treats $G(x) = \mathrm{I}$. On the other hand, the estimates for the mixed derivative $\int \nabla_x h \cdot \nabla_v h \, d\mu$ and the H^3_v semi-norm $\int |\nabla^3_v h|^2 \, d\mu$ are the same as in the Villani theory.

Indeed, the derivative for the H^3_v semi-norm satisfies

$$\frac{1}{2} \frac{d}{dt} \int |\nabla^3_v h|^2 \, d\mu \leq -\frac{\lambda_0}{4} \int |\nabla^4_v h|^2 \, d\mu + C \left(\int |\nabla^2_{vx} h|^2 \, d\mu + \int |\nabla^3_v h|^2 \, d\mu \right).$$

The estimates for H^1_x and H^3_v combined for sufficiently small $c > 0$ yield

$$\frac{1}{2}\frac{d}{dt}\left(\int |\nabla_x h|^2\,d\mu + c\int |\nabla_v^3 h|^2\,d\mu\right) \leq -K\left(\int |\nabla_v^4 h|^2\,d\mu + \int |\nabla_{vx}^2 h|^2\,d\mu\right)$$
$$+C\left(\int h^2\,d\mu + \int |\nabla_x h|^2\,d\mu + \int |\nabla_v h|^2\,d\mu + \int |\nabla_v^3 h|^2\,d\mu\right).$$

Some of the estimates of the terms involved in $\frac{d}{dt}\int \nabla_x h\cdot\nabla_v h\,d\mu$ are e.g.

$$\int \nabla_v h\cdot\nabla^2\mathcal{U}(x)\nabla_v h\,d\mu = -\int h\nabla^2\mathcal{U}(x):\nabla_v^2 h\,d\mu + \int h\nabla^2\mathcal{U}(x):(v\nabla_v h)\,d\mu$$
$$\leq \frac{1}{4}\int |\nabla_x h|^2\,d\mu + C\left(\int \left(h^2 + |\nabla_v h|^2 + |\nabla_v^2 h|^2\right)\,d\mu\right).$$

Also, we have

$$\int \gamma(x)\nabla_{vx}^2 h:\nabla_v^2 h\,d\mu = -\int \gamma(x)\nabla_x h\cdot\nabla_v\triangle_v h\,d\mu + \int \gamma(x)\nabla_x h\cdot(v\nabla_v^2 h)\,d\mu$$
$$\leq \frac{1}{4}\int |\nabla_x h|^2\,d\mu + C\left(\int |\nabla_v^2 h|^2\,d\mu + \int |\nabla_v^3 h|^2\,d\mu\right)$$

etc. These sum up to

$$\frac{d}{dt}\int \nabla_v h\cdot\nabla_x h\,d\mu \leq -\frac{1}{2}\int |\nabla_x h|^2\,d\mu$$
$$+C\left(\int h^2\,d\mu + \int |\nabla_v h|^2\,d\mu + \int |\nabla_v^2 h|^2\,d\mu + \int |\nabla_v^3 h|^2\,d\mu\right).$$

System of inequalities:

Borrowing once again notation from [71], we sum up the estimates to a system of inequalities. Set

$$X = \int |\nabla_x h|^2\,d\mu, \quad \mathcal{M} = \int \nabla_x h\cdot\nabla_v h\,d\mu, \quad Y_j = \int |\nabla_v^j h|^2\,d\mu \quad 1 \leq j \leq 4.$$

Also, for the sake of brevity, we consider

$$Z = Y_4, \quad \mathcal{E} = X + cY_3 \quad \text{for some}\, c > 0.$$

Then, we end up with the system of inequalities

$$\begin{aligned}
&\frac{d}{dt}\mathcal{E} \le -KZ + C(1 + Y_1 + \mathcal{E}),\\
&\frac{d}{dt}\mathcal{M} \le -KX + C(1 + Y_1 + Y_2 + Y_3),\\
&|\mathcal{M}| \le \sqrt{XY_1},\\
&\quad Y_1 \le CY_2^{1/2} \le C'Y_3^{1/3} \le C''Y_4^{1/4} \qquad \text{(interpolation inequality).}
\end{aligned}$$

For the above system of inequalities, it can be proven that when $0 < t \le 1$,

$$\mathcal{E}(t) \le \frac{C}{t^3} \qquad \text{for} \quad C > 0.$$

This concludes the proof of the hypoellipticity estimate.

□

In the preceding proof we made use of the following lemma for providing the regularity estimate.

Lemma 2. *(Villani) Assume the nonnegative continuous functions* $\mathcal{E}, X, Y, Z$ *on* $[0, 1]$*, and the continuous* $\mathcal{M}$ *on* $[0, 1]$*, that satisfy the system of inequalities*

$$K(X + Y) \le \mathcal{E} \le C(X + Y), \tag{4.28a}$$

$$|\mathcal{M}| \le C\mathcal{E}^{1-\delta}, \tag{4.28b}$$

$$\frac{d\mathcal{E}}{dt} \le -KZ + C\mathcal{E}, \tag{4.28c}$$

$$Y \le C(X + Z)^{1-\theta}, \tag{4.28d}$$

$$\frac{d\mathcal{M}}{dt} \le -KX + CY, \tag{4.28e}$$

for $K, C > 0$ *and* $\delta, \theta \in (0, 1)$*. It can be proven that there exists a constant* $\bar{C} > 0$

(depending on C, K, δ, θ) s.t

$$\mathcal{E}(t) \leq \frac{\bar{C}_{C,K,\theta,\delta}}{t^{1/k}} \qquad \text{for} \quad 0 < t < 1,$$

and $k = \min\left(\delta, \frac{\theta}{1-\theta}\right)$.

Remark 11. *In the example of regularization implied by the system presented above (when we proved regularity estimates for $H^3_v(\mu)$ and $H^1_x(\mu)$), the values of $\theta = \frac{1}{4}$ and $\delta = \frac{1}{3}$ give the desired optimal exponent $1/k = 3$.*

One could start with the observation that the constant $\bar{C}$ that appears in the estimate is not dependent on the regularity of f_0. This allows for the a priori estimate being true for the most general initial data that ensure a unique solution (say $f_0 \in L^1$). At the same time, there is nothing that ensures an optimal exponent (here $1/k$) for the estimate proven. Care must be taken in the choice of the interpolation inequalities so that one gets the best possible values for θ, δ.

Although the system of inequalities looks quite specific in nature, it has been correctly conjectured that it can be applied to more general cases of hypoelliptic regularization. A brief logistic explanation could be the following. Assume the evolution of a second order differential operator that is diffusive in a certain number of variables (in the problem we consider it is the velocity variables) and the remaining variables (which we might call missing variables) appear in derivatives of order at most one. If the quantity $\mathcal{E}$ corresponds to a general Sobolev norm (or rather seminorm) in all possible variables i.e. $\mathcal{E} = \int |\nabla^k_x \nabla^l_v f|^2$, $X = \int |\nabla^k_x f|^2$, and $Y = \int |\nabla^l_v f|^2 + \int f^2$. The inequality (4.28a) can be proven trivially with the help of Hölder's inequality in Fourier space. Inequalities (4.28c) and (4.28e) describe the evolution of $\mathcal{E}$ and the mixed derivative term $\mathcal{M}$. The propagation of $\mathcal{E}$ is controlled by a norm similar to $-Y$ but of order higher by one. That of $\mathcal{M}$ is controlled by $-X$. Finally, inequalities (ii) and (iv) are usually proven via interpolation techniques and are necessary for the system closure.

We present a proof to the above lemma, based on a proposition that appears to be more in the core of the proof. The idea lies in proving that a certain integro-differential inequality satisfies algebraic short time estimates with exponent that can be explicitly specified.

Lemma 3. *Assume the non increasing, continuous $\mathcal{E}(t)$, defined on the interval $[0,1]$ which satisfies the integro-differential inequality*

$$\int_I \mathcal{E}(s)\,ds \le C_1 \sup_I \mathcal{E}(t)^\alpha + C_2 \int_I (\mathcal{E}'(s))^\beta\,ds + C_3 \int_I \mathcal{E}(s)^\gamma\,ds,$$

for any interval I in $[0,1]$ with the constants $0 \le \alpha < 1$, $0 < \beta, \gamma < 1$, and $C_1, C_2, C_3 > 0$. Then, $\mathcal{E}(t)$ is shown to satisfy

$$\mathcal{E}(t) \le \frac{C}{t^{1/k}} \qquad 0 < t \le 1,$$

for $k = \min(1-\alpha, \frac{1-\beta}{\beta})$ and a constant C that depends only on the constants C_1, C_2, α, β.

Proof. We consider $E > 0$ (w.l.o.g. we can assume $E > 1$) and try to find the first time t_0 s.t $\mathcal{E}(t_0) < E$. We will show that $t_0 \le \frac{C'}{E^k}$. Let I_n be the time interval for which $2^n E \le \mathcal{E}(t) \le 2^{n+1}E$, for $n \ge 0$. Then, the length of this interval is

$$|I_n| = m\left(t \in [0,1] : 2^n E \le \mathcal{E}(t) \le 2^{n+1}E, \quad n \ge 0\right),$$

where m is the Lebesgue measure on the real line. Using the fact that $\mathcal{E}(t)$ is non increasing it is evident that $t_0 = \sum_{n\geq 0} |I_n|$. We will find a bound on $|I_n|$ related to E using the given integro-differential inequality and treating each term separately i.e.

$$\sup_{I_n} \mathcal{E}^\alpha \leq 2^{(n+1)\alpha} E^\alpha, \qquad \int_{I_n} \mathcal{E}(s)^\gamma \, ds \leq 2^{(n+1)\gamma} E^\gamma |I_n|,$$

and

$$\int_{I_n} (\mathcal{E}'(s))^\beta \, ds \leq \left(\int_{I_n} \mathcal{E}'(s) \, ds \right)^\beta |I_n|^{1-\beta} \leq 2^{n\beta} E^\beta |I_n|^{1-\beta}.$$

In the second line above, we have employed Hölder inequality.

Putting all terms together, one gets

$$2^n E |I_n| \leq \int_{I_n} \mathcal{E}(s) \, ds \leq C_1 2^{(n+1)\alpha} E^\alpha + C_2 2^{n\beta} E^\beta |I_n|^{1-\beta} + C_3 2^{(n+1)\gamma} E^\gamma |I_n|.$$

It is only the first two terms in the r.h.s that can control the l.h.s (for large enough E). If the first term in r.h.s dominates, then

$$|I_n| \leq 3 C_1 2^\alpha \frac{1}{2^{n(1-\alpha)}} \frac{1}{E^{1-\alpha}}.$$

In the case of the second term being the dominant

$$|I_n| \leq 3 C_2^{1/\beta} \frac{1}{2^{n(1-\beta)/\beta}} \frac{1}{E^{(1-\beta)/\beta}}.$$

Summing on $|I_n|$ terms gives $t_0 \leq C'/E^k$, for the value of k as presented in the lemma. $\square$

In this final step in the proof of Villani's lemma, we show that $\mathcal{E}(t)$ satisfies an inequality similar to the one in the lemma above. To make $\mathcal{E}(t)$ non increasing, one considers $e^{-Ct}\mathcal{E}(t)$ which should satisfy the same short time estimates. With this transformation, equation (4.28c) now becomes $\mathcal{E}' \leq -KZ$. Employing inequalities (4.28a) - (4.28e), we get

$$
\begin{aligned}
\mathcal{E}(t) &\leq C(X(t)+Y(t)) \\
&\leq C\left(X(t)+C(X(t)+Z(t))^{1-\theta}\right) \leq CX(t)+C^2(X(t)+Z(t))^{1-\theta} \\
&\leq C\left(-\frac{1}{K}\mathcal{M}'(t)+\frac{C}{K}Y(t)\right)+C^2(X(t)+Z(t))^{1-\theta} \\
&\leq -\frac{C}{K}\mathcal{M}'(t)+\frac{C^2}{K}Y(t)+C^2X(t)^{1-\theta}+C^2Z(t)^{1-\theta} \\
&\leq -\frac{C}{K}\mathcal{M}'(t)+C^2\left(\frac{C}{K}+1\right)(X(t)^{1-\theta}+Z(t)^{1-\theta}) \\
&\leq -\frac{C}{K}\mathcal{M}'(t)+C^2\left(\frac{C}{K}+1\right)\left(\left(\frac{1}{K}\mathcal{E}(t)\right)^{1-\theta}+Z(t)^{1-\theta}\right).
\end{aligned}
$$

This is short of the inequality

$$\mathcal{E}(t) \leq C_1\mathcal{M}'(t)+C_2\mathcal{E}(t)^{1-\theta}+C_3(-\mathcal{E}'(t))^{1-\theta},$$

for constants $C_1, C_2, C_3 > 0$. Integrating over a time interval I, we have

$$\int_I \mathcal{E}(s)\,ds \leq 2C_1\sup_I \mathcal{E}(t)^{1-\delta}+C_2\int_I \mathcal{E}(s)^{1-\theta}\,ds+C_3\int_I(-\mathcal{E}'(s))^{1-\theta}\,ds.$$

We can now conclude with the second lemma, with the choice of exponents $\alpha = 1-\delta$ and $\beta = \gamma = 1-\theta$.

4.2.5 Higher Order Sobolev Regularity

We can use the differential system presented previously to obtain short time estimates for higher Sobolev norms both in weighted and in "flat" space (with no weights). The analysis is not really that much different to the one followed in the previous paragraph, with the exception that in this case it is the evolution of higher Sobolev norms that is considered, and so one must also find appropriate expressions for higher order mixed derivatives. We begin with a statement for identity ($G(x) = \mathrm{I}$) diffusivity, in which a stronger result for solutions with L^1 initial data can be proven. The exact statement is:

Theorem 13. *(Villani) Assume the unique solution to the F-P equation*

$$\partial_t f + v \cdot \nabla_x f - \nabla \mathcal{U}(x) \cdot \nabla_v f = \nabla_v \cdot (\nabla_v f + v f),$$

with initial data $f_0 \in L^1(\mathbb{R}^{n,n}_{x,v})$, a smooth potential $\mathcal{U}(x)$ satisfying $\inf \mathcal{U}(x) > -\infty$ *and having all its derivatives uniformly bounded, i.e. for $j \geq 2$*

$$|\nabla^j \mathcal{U}(x)| \leq C_j \qquad \textit{for some} \quad C_j > 0.$$

Under the above assumptions, it can be proven that for any positive integers $m > 0$, we have

$$\sum_{3k+l \leq 3m} \|f\|_{H^{k,l}_{x,v}(\mathbb{R}^n_x \times \mathbb{R}^n_v)} \leq \frac{C}{t^\beta} \qquad \textit{for} \quad 0 < t \leq 1,$$

where the constant $C(k, l, C_j, \lambda_0, \Lambda_j)$ does not depend on the regularity of the initial data, and a constant $\beta = \beta(n, m) > 0$. It can be proven with a bit of further analysis that the estimate above is true for any $\beta > n + 3m/2$.

Since, in the proof of this result initial data with L^1 regularity is used, one has to be especially careful with the interpolation technique that will be required. For this reason the following Nash type inequality is used instead of the more elementary interpolation with Hermite polynomials used in previous paragraph.

Lemma 4. *Assume a smooth function $f(x,v)$ with $(x,v) \in \mathbb{R}^n_x \times \mathbb{R}^n_v$. Assume also integers $\lambda, \mu \geq 0$ and $\lambda', \mu' > 0$. If*

$$\frac{\lambda}{\lambda'} + \frac{\mu}{\mu'} < 1,$$

there exists a constant C depending only on these integers s.t.

$$\iint |\nabla_x^{\lambda} \nabla_v^{\mu} f|^2 \leq C \left(\iint |\nabla_x^{\lambda'} f|^2 + \iint |\nabla_v^{\mu'} f|^2 \right)^{1-\theta} \left(\iint f \right)^{2\theta},$$

where

$$\theta = \frac{1 - \left(\frac{\lambda}{\lambda'} + \frac{\mu}{\mu'}\right)}{1 + \frac{n}{2}\left(\frac{1}{\lambda'} + \frac{1}{\mu'}\right)}.$$

The proof we are about to present is a slightly modified version of the one given by Villani, in the sense that we treat the more general case of non-constant diffusivity. This case involves slightly more elaborate treatment of the estimates for $\mathcal{E}(h)$ and $\mathcal{M}(h)$, which turn out being essentially the same as in the proof of the previous theorem. Unfortunately, we were unable to reproduce the result in the flat space $H^{k,l}_{x,v}(\mathbb{R}^n_x \times \mathbb{R}^n_v)$ but rather in $H^{k,l}_{x,v}(\mu)$. The reason for this is the slightly different structure in the evolution of Sobolev norms. Moreover, the interpolation estimates are not necessarily satisfied for every measure μ. Here we were able to prove them for the quadratic potential $\mathcal{U}(x) = \frac{x^2}{2}$ ($+L^\infty$ perturbations of it), using a simple extension of the interpolation result with Hermite polynomials (where interpolation is now performed both in x, v variables) presented in the previous paragraph. On the plus side, we were able to catch the optimal exponent $\beta = 3m$. Indeed, it is proven that:

Theorem 14. *Let $h(t,x,v)$ be a solution to the F-P equation*

$$\partial_t h + v\cdot\nabla_x h - \nabla\mathcal{U}(x)\cdot\nabla_v h = \gamma(x)\triangle_v h - \gamma(x)v\cdot\nabla_v h,$$

with $h_0 \in L_2(\mu)$. Assume a smooth $\mathcal{U}(x)\in C^\infty(\mathbb{R}^n_x)$ satisfying $\inf\mathcal{U}(x) > -\infty$ and

$$|\nabla^j\mathcal{U}(x)| \le C_j \qquad \text{for} \quad C_j > 0, \quad j\ge 2.$$

Assume also, $\gamma(x)\in C^\infty(\mathbb{R}^n_x)$ satisfies the extra conditions

$$\lambda_0 \le \gamma(x) \le \Lambda_0 \qquad \text{for} \quad \lambda_0, \Lambda_0 > 0,$$

and

$$|\nabla^j_x\gamma(x)| \le \Lambda_j \qquad \Lambda_j > 0 \quad j \ge 1.$$

It is proven that for any integer $m\ge 0$, the following estimates holds

$$\frac{d}{dt}\mathcal{E}_m(h) \le -K\int |\nabla_v^{3m+1}h|^2\,d\mu + C\mathcal{E}_m(h), \qquad \text{and}$$

$$\frac{d}{dt}\mathcal{M}_m(h) \le -K\int |\nabla_x^m h|^2\,d\mu + C\sum_{k<m,\,3k+l\le 3m}\int |\nabla_x^k\nabla_v^l h|^2\,d\mu \qquad \text{for} \quad K, C > 0,$$

where $\mathcal{E}_m(h) = \sum\limits_{3k+l\le 3m} a_{k,l}\int|\nabla_x^k\nabla_v^l h|^2\,d\mu$ is the generalized Sobolev norm that captures up to m derivatives in x and up to $3m$ in v (with coefficients $a_{k,l}>0$ that are going to be determined in detail), and $\mathcal{M}_m(h) = \int\nabla_x^m h\cdot\nabla_x^{m-1}\nabla_v h\,d\mu$ is the generalized mixed derivative functional. If in addition, the following interpolation inequalities hold for the measure μ (like in the case of a quadratic potential $\mathcal{U}(x) = \frac{x^2}{2}$)

$$\mathcal{E}_m(h) \le C\left(\int|\nabla_x^m h|^2 d\mu + \int|\nabla_v^{3m}h|^2 d\mu + \int h^2 d\mu\right),$$

$$\mathcal{M}_m(h) \leq C\mathcal{E}_m(h)^{1-\delta}$$

for some $\theta, \delta \in (0,1)$, then the Sobolev regularity estimate

$$\sum_{3k+l\leq 3m} \|h\|_{H^{k,l}_{x,v}(\mu)} \leq \frac{C}{t^\beta} \qquad \textit{for} \quad 0 < t \leq 1$$

holds, with $\beta > 0$ and $C > 0$.

As mentioned already, we make use of the following interpolation estimate.

Lemma 5. *(Interpolation in Gaussian measure) Let $h \in H^m_x(\mu) \cap H^{3m}_v(\mu)$ for the measure $\mu(dx, dv) = e^{-x^2/2}e^{-v^2/2}dxdv$. Then, $h \in H^{k,l}_{x,v}(\mu)$ for $3k + l \leq 3m$. Specifically,*

$$\int |\nabla^k_x \nabla^l_v h|^2 \, d\mu \leq C \left(\int |\nabla^m_x h|^2 \, d\mu \right)^{\frac{k}{m}} \left(\int |\nabla^{3m}_v h|^2 \, d\mu \right)^{\frac{l}{3m}} \left(\int h^2 \, d\mu \right)^{1-\frac{k}{m}-\frac{l}{3m}}$$

for some $C > 0$ independent of h.

Proof. Since one deals with Gaussian measure, the proof is based as usual on interpolation using Hermite polynomials. Hermite polynomials form a natural basis because they are orthonormal in the gaussian measure. The only slight deviation here is that interpolation is done w.r.t. both variables x, v. Indeed, if we write $h(x,v) = \sum_{n,r} c_{n,r} H_n(x) H_r(v)$, then it follows upon differentiation of h, subsequent integration using the orthonormality of Hermite functions, and finally by a simple

C-S inequality for series that

$$
\begin{aligned}
\int |\nabla_x^k \nabla_v^l h|^2 \, d\mu &= \int \Big| \sum_{n,r} c_{n,r} \nabla_x^k \nabla_v^l H_n(x) H_r(v) \Big|^2 \, d\mu \\
&= \int \Big| \sum_{n,r} c_{n,r} \frac{n!}{(n-k)!} \frac{r!}{(r-l)!} H_{n-k}(x) H_{r-l}(v) \Big|^2 e^{-\frac{x^2}{2}} e^{-\frac{v^2}{2}} \, dv \, dx \\
&= \sum_{n,r} c_{n,r}^2 \left(\frac{n!}{(n-k)!} \frac{r!}{(r-l)!} \right)^2 \le C \sum_{n,r} c_{n,r}^2 n^{2k} r^{2l} \\
&\le C \left(\sum_{n,r} c_{n,r}^2 n^{2m} \right)^{\frac{k}{m}} \left(\sum_{n,r} c_{n,r}^2 r^{6m} \right)^{\frac{l}{3m}} \left(\sum_{n,r} c_{n,r}^2 \right)^{1-\frac{k}{m}-\frac{l}{3m}} \\
&\le C \left(\int |\nabla_x^m h|^2 \, d\mu \right)^{\frac{k}{m}} \left(\int |\nabla_v^{3m} h|^2 \, d\mu \right)^{\frac{l}{3m}} \left(\int h^2 \, d\mu \right)^{1-\frac{k}{m}-\frac{l}{3m}} .
\end{aligned}
$$

□

Notice that the above lemma implies directly the interpolation inequality $\mathcal{E}_m(h) \le C \left(\int |\nabla_x^m h|^2 d\mu + \int |\nabla_v^{3m} h|^2 d\mu + \int h^2 d\mu \right)$ for the Gauss measure μ. The proof of the main result follows.

Proof. The proof begins with an estimate on higher order derivatives in both x, v variables, i.e we look after an estimate for

$$
\mathcal{E}_m(h) = \sum_{3k+l \le 3m} a_{k,l} \int |\nabla_x^k \nabla_v^l h|^2 \, d\mu,
$$

with coefficients $a_{k,l}$ that will be chosen appropriately.

Here, $|\nabla_x^k \nabla_v^l h|^2$ is short for the sum $\sum_{k_1+\ldots+k_n=k} \sum_{l_1+\ldots+l_n=l} |\partial_{x_1,\ldots,x_n}^{k_1,\ldots,k_n} \partial_{v_1,\ldots,v_n}^{l_1,\ldots,l_n} h|^2$. In general, the notation $\nabla_x^k \nabla_v^l$ will be used short of any operator of the type $\partial_{x_1,\ldots,x_n}^{k_1,\ldots,k_n} \partial_{v_1,\ldots,v_n}^{l_1,\ldots,l_n}$, where $k_1 + \ldots + k_n = k$ and $l_1 + \ldots + l_n = l$ for $k_i, l_i \ge 0$. This compact notation is used to avoid the extremely cumbersome alternative of carrying indices.

We differentiate the equation k times in the x variables and l times in the v variables (k, l are both non negative integers), to get

$$\partial_t \nabla_x^k \nabla_v^l h + l \nabla_x^{k+1} \nabla_v^{l-1} h + v \cdot \nabla_x^{k+1} \nabla_v^l h - \sum_{r=0}^{k} C_r \nabla^{r+1} \mathcal{U}(x) \cdot \nabla_x^{k-r} \nabla_v^{l+1} h$$
$$= \sum_{r=0}^{k} C_r \nabla_x^r \gamma(x) \Big(\nabla_x^{k-r} \nabla_v^l \triangle_v h - l \nabla_x^{k-r} \nabla_v^l h - v \cdot \nabla_x^{k-r} \nabla_v^{l+1} h \Big)$$

for specific constants C_r, ($C_0 = 1, \ldots$).

Then after we multiply by $\nabla_x^k \nabla_v^l h$ and integrate (in measure μ) by parts, we have

$$\frac{1}{2} \frac{d}{dt} \int |\nabla_x^k \nabla_v^l h|^2 \, d\mu - l \int \Big(\nabla_x^k \nabla_v^{l-1} \triangle_v h - v \cdot \nabla_x^k \nabla_v^l h \Big) \cdot \nabla_x^{k+1} \nabla_v^{l-2} h \, d\mu$$
$$- \sum_{r=1}^{k} C_r \int \nabla_x^k \nabla_v^l h \cdot \left(\nabla^{r+1} \mathcal{U}(x) \cdot \nabla_x^{k-r} \nabla_v^{l+1} h \right) \, d\mu$$
$$= - \sum_{r=0}^{k} C_r \int \nabla_x^k \nabla_v^{l+1} h \cdot \nabla_x^r \gamma(x) \nabla_x^{k-r} \nabla_v^{l+1} h \, d\mu$$
$$- \sum_{r=0}^{k} C_r l \int \nabla_x^k \nabla_v^l h \cdot \nabla_x^r \gamma(x) \nabla_x^{k-r} \nabla_v^l h \, d\mu.$$

This yields the following inequality

$$\frac{1}{2} \frac{d}{dt} \int |\nabla_x^k \nabla_v^l h|^2 \, d\mu \leq -\lambda_0 \int |\nabla_x^k \nabla_v^{l+1} h|^2 \, d\mu + C \int |\nabla_x^k \nabla_v^{l+1} h| |\nabla_x^{k+1} \nabla_v^{l-2} h| \, d\mu$$
$$+ C \sum_{r=1}^{k} \int |\nabla_x^k \nabla_v^{l+1} h| |\nabla^{r+1} \mathcal{U}(x)| |\nabla_x^{k-r} \nabla_v^l h| \, d\mu$$
$$+ C \sum_{r=1}^{k} \int |\nabla_x^k \nabla_v^{l+1} h| |\nabla_x^r \gamma(x)| |\nabla_x^{k-r} \nabla_v^{l+1} h| \, d\mu$$
$$+ C \sum_{r=0}^{k} \int |\nabla_x^k \nabla_v^l h| |\nabla_x^r \gamma(x)| |\nabla_x^{k-r} \nabla_v^l h| \, d\mu,$$

for some $C > 0$. This inequality, together with the fact that $|\nabla^j \mathcal{U}(x)|$, $|\nabla_x^j \gamma(x)|$ are bounded by assumption and the use of Young and C-S inequalities gives

$$
\begin{aligned}
\frac{1}{2}\frac{d}{dt}\int |\nabla_x^k \nabla_v^l h|^2 \, d\mu \le -\frac{\lambda_0}{2}\int |\nabla_x^k \nabla_v^{l+1} h|^2 \, d\mu + C\int |\nabla_x^{k+1}\nabla_v^{l-2} h|^2 \, d\mu \\
+ C\sum_{r=0}^{k}\int |\nabla_x^{k-r}\nabla_v^{l} h|^2 \, d\mu + C\sum_{r=1}^{k}\int |\nabla_x^{k-r}\nabla_v^{l+1} h|^2 \, d\mu.
\end{aligned}
$$

Now, we fix some integer $m > 0$ and choose coefficients $a_{k,l}$ with the only condition to be satisfied that $Ca_{k,3(m-k)} < \frac{\lambda_0}{2} a_{k+1,3(m-k-1)}$ for $0 \le k \le m-1$. From this choice of coefficients $a_{k,l}$, it is trivial to reduce the last inequality to

$$
\frac{d}{dt}\mathcal{E}_m(h) \le -K\int |\nabla_v^{3m+1} h|^2 \, d\mu + C\mathcal{E}_m(h),
$$

for the new constants $K, C > 0$.

In similar manner, differentiating $\mathcal{M}_m(h)$ gives

$$
\begin{aligned}
&\frac{d}{dt}\mathcal{M}_m(h)\\
&= \int \partial_t(\nabla_x^m h)\cdot \nabla_x^{m-1}\nabla_v h\, d\mu + \int \nabla_x^m h \cdot \partial_t(\nabla_x^{m-1}\nabla_v h)\, d\mu\\
&= \sum_{r=1}^{m} C_r \int \Big(\nabla^{r+1}\mathcal{U}(x)\cdot \nabla_x^{m-r}\nabla_v h\Big)\cdot \nabla_x^{m-1}\nabla_v h\, d\mu\\
&+ \sum_{r=0}^{m} C_r \int \nabla_x^r \gamma(x) \nabla_x^{m-r} h \cdot \Big(\nabla_x^{m-1}\nabla_v \triangle_v h - v\nabla_x^{m-1}\triangle_v h\Big)\, d\mu\\
&- \int |\nabla_x^m h|^2\, d\mu + \sum_{r=0}^{m-1} C_r \int \Big(\nabla^{r+1}\mathcal{U}(x)\cdot \nabla_x^{m-1-r}\nabla_v^2 h\Big)\cdot \nabla_x^m h\, d\mu\\
&+ \sum_{r=0}^{m-1} C_r \int \nabla_x^r\gamma(x)\Big(\nabla_x^{m-1-r}\nabla_v\triangle_v h - \nabla_x^{m-1-r}\nabla_v h - v\cdot \nabla_x^{m-1-r}\nabla_v^2 h\Big)\cdot \nabla_x^m h\, d\mu\\
&\le -\int |\nabla_x^m h|^2\, d\mu + C\sum_{r=1}^{m}\int |\nabla^{r+1}\mathcal{U}(x)||\nabla_x^{m-r}\nabla_v h||\nabla_x^{m-1}\nabla_v h|\, d\mu\\
&+ C\sum_{r=0}^{m}\int |\nabla_x^r\gamma(x)||\nabla_x^{m-r} h|\Big(|\nabla_x^{m-1}\nabla_v\triangle_v h| + |v\nabla_x^{m-1}\triangle_v h|\Big)\, d\mu\\
&+ C\sum_{r=0}^{m-1}\int |\nabla^{r+1}\mathcal{U}(x)||\nabla_x^{m-1-r}\nabla_v^2 h||\nabla_x^m h|\, d\mu\\
&+ C\sum_{r=0}^{m-1}\int |\nabla_x^r\gamma(x)|\Big(|\nabla_x^{m-1-r}\nabla_v\triangle_v h|\\
&\qquad\qquad + |\nabla_x^{m-1-r}\nabla_v h| + |v\cdot \nabla_x^{m-1-r}\nabla_v^2 h|\Big)|\nabla_x^m h|\, d\mu.
\end{aligned}
$$

Now, using the fact that $|\nabla^j\mathcal{U}(x)|$ and $|\nabla_x^j\gamma(x)|$ are bounded, and Young's inequality

$2ab \leq \epsilon a^2 + \frac{1}{\epsilon} b^2$ with $\epsilon > 0$, one gets

$$
\begin{aligned}
\frac{d}{dt}\mathcal{M}_m(h) &\leq -\frac{1}{2}\int |\nabla_x^m h|^2 \, d\mu + C\sum_{r=1}^{m}\int \left(|\nabla_x^{m-r}\nabla_v h|^2 + |\nabla_x^{m-1}\nabla_v h|^2\right) d\mu \\
&\quad + C\sum_{r=1}^{m}\int |\nabla_x^{m-r} h|^2 \, d\mu + C\int \left(|\nabla_x^{m-1}\nabla_v^2 h|^2 + |\nabla_x^{m-1}\nabla_v^3 h|^2\right) d\mu \\
&\quad + C\sum_{r=0}^{m-1}\int |\nabla_x^{m-1-r}\nabla_v^2 h|^2 \, d\mu \\
&\quad + C\sum_{r=0}^{m-1}\int \left(|\nabla_x^{m-1-r}\nabla_v h|^2 + |\nabla_x^{m-1-r}\nabla_v^2 h|^2 + |\nabla_x^{m-1-r}\nabla_v^3 h|^2\right) d\mu \\
&\leq -\frac{1}{2}\int |\nabla_x^m h|^2 \, d\mu + C\sum_{k<m,\, 3k+l\leq 3m}\int |\nabla_x^k \nabla_v^l h|^2 \, d\mu.
\end{aligned}
$$

Now that the two major estimates for $\mathcal{E}_m$ and $\mathcal{M}_m$ have been shown, consider the choice $X = \int |\nabla_x^m h|^2 \, d\mu$, $Y = \int (|\nabla_v^{3m} h|^2 + h^2) \, d\mu$, and $Z = \int |\nabla_v^{3m+1} h|^2 \, d\mu$. Coupled with the extra conditions in the assumptions of the theorem and the use of the differential inequality system we have already mentioned implies the desired Sobolev estimate.

The last step in the proof is to show that the Gaussian measure satisfies these extra interpolation assumptions made in the statement of this theorem. It turns out that this is the case with values of θ, δ that can be computed explicitly.

Indeed, one has the estimate the standard interpolation inequality

$$
\int |\nabla_v^{3m} h|^2 \, d\mu \leq C\left(\int |\nabla_v^{3m+1} h|^2 \, d\mu\right)^{\frac{3m}{3m+1}} \left(\int h^2 \, d\mu\right)^{\frac{1}{3m+1}}
$$

which implies that $1 - \theta = \frac{3m}{3m+1}$ and subsequently $\frac{\theta}{1-\theta} = \frac{1}{3m}$.

At the same time, the estimate for the mixed derivative is

$$
\begin{aligned}
|\mathcal{M}_m(h)| &\le \left(\int |\nabla_x^m h|^2\,d\mu\right)^{\frac{1}{2}} \left(\int |\nabla_x^{m-1}\nabla_v h|^2\,d\mu\right)^{\frac{1}{2}} \\
&\le \left(\int |\nabla_x^m h|^2\,d\mu\right)^{\frac{1}{2}} \left(\int |\nabla_x^m h|^2\,d\mu\right)^{\frac{m-1}{2m}} \left(\int |\nabla_v^{3m} h|^2\,d\mu\right)^{\frac{1}{6m}} \left(\int h^2\,d\mu\right)^{\frac{1}{2}-\frac{m-1}{2m}-\frac{1}{6m}} \\
&\le \left(\int |\nabla_x^m h|^2\,d\mu\right)^{1-\frac{1}{2m}} \left(\int |\nabla_v^{3m} h|^2\,d\mu\right)^{\frac{1}{6m}} \left(\int h^2\,d\mu\right)^{\frac{1}{3m}}.
\end{aligned}
$$

This implies that in our case $\delta = \frac{1}{2m}$ (for $m \ge 1$). This means that for the Gauss measure $d\mu = e^{-x^2/2}e^{-v^2/2}dxdv$, the exponent that one gets is $\beta = 3m$ which happens to be optimal.

□

We shall make a couple of remarks on the connection between the estimates for weighted and flat Sobolev spaces and the optimal exponents. The first one is that given initial data in $L^2(\mu)$, the weighted estimates actually hold for flat norms with the same exponent due to the inequality

$$
\begin{aligned}
&\int |\nabla_x^k \nabla_v^l f|^2\,dv\,dx = \int |\nabla_x^k \nabla_v^l (h\mathcal{M}_{eq})|^2\,dv\,dx = \int |\nabla_x^k \nabla_v^l (he^{-\mathcal{U}(x)}e^{-\frac{v^2}{2}})|^2\,dv\,dx \\
&\le C\int \Big|\sum_{\alpha=0}^{k}\sum_{\beta=0}^{l} \nabla_x^\alpha \nabla_v^\beta h \nabla_x^{k-\alpha}(e^{-\mathcal{U}(x)})\nabla_v^{l-\beta}(e^{-\frac{v^2}{2}})\Big|^2\,dv\,dx \\
&\le C\sup_{x,v}(1+|v|^2+\ldots+|v|^{2l})(1+|\mathcal{U}(x)|^2+\ldots+|\nabla^k\mathcal{U}(x)|^2)\mathcal{M}_{eq} \\
&\sum_{\alpha=0}^{k}\sum_{\beta=0}^{l}\int |\nabla_x^\alpha \nabla_v^\beta h|^2\,d\mu \le C\sum_{\alpha=0}^{k}\sum_{\beta=0}^{l}\int |\nabla_x^\alpha \nabla_v^\beta h|^2\,d\mu.
\end{aligned}
$$

Of course, this estimate says nothing about the case of L^1 initial data and as we have seen so far, one can expect that the optimal exponents (at least when $\gamma(x)$ is a constant) is given in $\|\nabla_x^m f\|_{L^2} + \|\nabla_v^{3m} f\|_{L^2} \le \frac{C}{t^{n+3m/2}}$ with C being independent of the regularity of f_0. In fact, Villani shows that when $\gamma(x)$ is constant the exponent that one gets by his technique is $3n + 3m/2$ and that with the help of the Nash inequality presented in the form of a lemma earlier, this exponent can decrease to any β with $\beta > n + 3m/2$ (but not shown in the critical case $\beta = n + 3m/2$).

4.2.6 Regularization from L^1 Data

We have already presented a result on the regularizing properties of the semigroup e^{-tL} for the F-P operator $Lh = v \cdot \nabla_x h - \nabla \mathcal{U}(x) \cdot \nabla_v h - (\nabla_v - v) \cdot \nabla_v h$, when one encounters entropic initial conditions (in $L \log L(\mu)$). In fact this result can be relaxed to L^1 data as it is evident in:

Theorem 15. *Consider the Cauchy problem $\partial_t f + Lf = 0$ with initial data $f_0 \in L^1(\mathbb{R}^n_x \times \mathbb{R}^n_v)$. The potential $\mathcal{U}(x)$ is assumed to be Lipschitz which guarantees a unique solution. Assume also bounded second moments for f_0 i.e. $\iint f_0(|x|^2+|v|^2) < \infty$. It is shown that the semigroup solution regularizes the relative entropy in the sense that*

$$\iint f \log \frac{f}{\mathcal{M}_{eq}} \, dvdx < \infty \qquad \textit{for} \quad t > 0.$$

Proof. We start with an estimate on the growth of the general moment $\iint f(|x|^2 + |v|^2)^{s/2}$ for $s > 0$. Computations are performed component-wise with the standard convention of summation over repeated indices.

$$\begin{aligned}
&\frac{d}{dt} \iint f(|x|^2 + |v|^2)^{s/2} = \iint \partial_t f(|x|^2 + |v|^2)^{s/2} \\
&= \iint \left(-v_j \partial_{x_j} f + \partial_{x_j} \mathcal{U}(x) \partial_{v_j} f - \partial^2_{v_j} f - \partial_{v_j}(v_j f)\right) (|x|^2 + |v|^2)^{s/2} \\
&= \iint v_j f \partial_{x_j} (|x|^2 + |v|^2)^{s/2} + \iint \left(-\partial_{x_j} \mathcal{U}(x) f + \partial_{v_j} f + v_j f\right) \partial_{v_j} (|x|^2 + |v|^2)^{s/2} \\
&= s \iint f(|x|^2 + |v|^2)^{\frac{s}{2}-1} \left(v_j x_j - v_j \partial_{x_j} \mathcal{U}(x) + v_j^2\right) - f \partial_{v_j} \left(v_j (|x|^2 + |v|^2)^{\frac{s}{2}-1}\right) \\
&\leq C \iint f(|x|^2 + |v|^2)^{s/2} - s \iint f \delta_{jj} (|x|^2 + |v|^2)^{\frac{s}{2}-1} \\
&\qquad\qquad - 2s \left(\frac{s}{2} - 1\right) \iint f v_j v_j (|x|^2 + |v|^2)^{\frac{s}{2}-2},
\end{aligned}$$

where we made use of the Lipschitz continuity of $\mathcal{U}(x)$. Taking $s = 2$ and using conservation of mass, one gets

$$\frac{d}{dt}\iint f(|x|^2 + |v|^2) \leq C \iint f(|x|^2 + |v|^2)$$

which implies $\iint f(|x|^2 + |v|^2) = O(e^{Ct})$, or $O(1+t)$ for $0 < t < 1$ (given $\int f_0(|x|^2 + |v|^2) < \infty$).

The regularization in $H^{k,l}_{x,v}(\mathbb{R}^n_x \times \mathbb{R}^n_v)$ coupled with the anisotropic Nash inequality used by Villani implies that $\iint f^2$ behaves like $O(t^{-\beta})$ for some $\beta > 0$ and $0 < t < 1$. One can now make use of the bounds for $\int f^2$ and $\int f(|x|^2 + |v|^2)$ to obtain

$$\begin{aligned}
\iint f \log f \, dv \, dx &\leq C_{3/2} \int f^{3/2} \, dv \, dx \\
&\leq C_{3/2} \left(\iint f(1 + |x|^2 + |v|^2) \, dv \, dx \right)^{\frac{1}{2}} \left(\iint \frac{f^2}{1 + |x|^2 + |v|^2} \, dv \, dx \right)^{\frac{1}{2}} \\
&\leq O\left(\left(\frac{1+t}{t^\beta} \right)^{1/2} \right).
\end{aligned}$$

While, at the same time

$$\iint f \log \frac{f}{\mathcal{M}_{eq}} \leq \iint f \log f + C \iint f(|x|^2 + |v|^2)$$

which completes the proof.

$\square$

Chapter 5: Convergence to Equilibrium

In this chapter, we present results of long time asymptotics for a regularized FP operator L. The results are strongly influenced by the theory of hypocoercivity systematically studied by C.Villani and outlined in [71]. Hypocoercivity deals with operators that are not coercive in a Hilbertian framework, but for which we can create an "appropriate" norm so that L is now coercive. The theory has some nice generalizations but results tailored for the F-P operator suffice here.

The method employed here bears a strong resemblance to the technique by Hérau used in the previous section. We make use of a functional $\mathcal{E}(h)$ that gives a measure of distance of a solution to the unique stationary state $\mathcal{M}_{eq}(x, v)$. In general, we try to prove under certain assumptions, that $-\mathcal{E}'(h) \geq K\mathcal{E}(h)^{1+\epsilon}$ for some $K > 0$, $\epsilon \geq 0$. If $\epsilon = 0$, exponential decay of $\mathcal{E}(h)$ is implied, otherwise for $\epsilon > 0$ the rate of decay is algebraic.

We start with a result for $L^2(\mu)$ initial data, exactly like we did when we studied short time regularity estimates. We then proceed to a result from [15], which gives an algebraic decay rate for the relative entropy functional. We can strengthen this result by adding extra terms to the entropy functional (terms including derivatives + a mixed derivative term), and show exponential convergence rate for $L \log L(\mu)$ initial data. Finally, a different approach based on a method by Dolbeault-Mouhot-Schmeiser is presented.

5.1 Hypocoercivity

The theorem proved in this paragraph can be generalized to any unbounded operator of the form $L = A^*A + B$ in some Hilbert space $\mathcal{H}$, with B being antisymmetric ($B^* = -B$). It is related to the study of convergence rates of a solution of the equation $\partial_t f + Lf = 0$ to the unique stationary state $\mathcal{M}_{eq}$. The Hilbert space $\mathcal{H}$ can be assumed being $L^2(\mu)$ for the main result in this paragraph. The norm in $L^2(\mu)$ will be denoted by $\|\cdot\|$ and it is generated by the inner product $\langle\cdot,\cdot\rangle$. The adjoint of an operator will be understood in this setting. Any norm and inner product in this paragraph that isn't specified otherwise will be assumed to be related to $L^2(\mu)$.

Before we begin with the main theory, we give two definitions to shed some light on the difference between coercivity and hypocoercivity. First, let us assume that the (unbounded) operator $-L$ generates the continuous, contraction, semigroup $(S_t)_{t\geq 0}$ on on the Hilbert space $\mathcal{H}$ (with inner product mentioned above), i.e. $S_t = e^{-tL}$. We will define the two notions of coercivity/hypocoercivity to hold on a Hilbert space $\widetilde{\mathcal{H}}$, which will in general be narrower than $\mathcal{H}$. $\widetilde{\mathcal{H}}$ is endowed with the inner product $\langle\cdot,\cdot\rangle_{\widetilde{\mathcal{H}}}$. For simplicity, we assume both Hilbert spaces to be real Hilbert spaces. We now give give definitions of coercivity and hypocoercivity in the spirit of [71], so that the importance of functional setting becomes apparent.

Definition 1. *An operator L is called λ-coercive (for $\lambda > 0$) on $\widetilde{\mathcal{H}}$, if*

$$\langle h, Lh\rangle_{\widetilde{\mathcal{H}}} \geq \lambda\|h\|^2_{\widetilde{\mathcal{H}}} \qquad \forall h \in D(L)\cap\widetilde{\mathcal{H}},$$

and coercive on $\widetilde{\mathcal{H}}$ if the above inequality holds for some $\lambda > 0$.

From this definition, it follows trivially that L is λ-coercive iff

$$\forall h \in \widetilde{\mathcal{H}}, \quad t \geq 0, \qquad \|e^{-tL}h\|_{\widetilde{\mathcal{H}}} \leq e^{-\lambda t}\|h\|_{\widetilde{\mathcal{H}}}.$$

The usual space $\widetilde{\mathcal{H}}$ on which we define coercivity is $\widetilde{\mathcal{H}} = \mathcal{H}/\mathcal{N}(L)$. The following definition provides a property weaker than that of coercivity, which we call hypocoercivity. It will hold on a generally different Hilbert space $\widetilde{\mathcal{H}}$.

Definition 2. *Assume an (unbounded) operator L on $\mathcal{H}$, generating a continuous semi-group $(e^{-tL})_{t\geq 0}$. We say that L is λ-hypocoercive on $\widetilde{\mathcal{H}}$ (for $\lambda > 0$) if there exist some $C > 0$ s.t.*

$$\|e^{-tL}h\|_{\widetilde{\mathcal{H}}} \leq Ce^{-\lambda t}\|h\|_{\widetilde{\mathcal{H}}} \qquad \forall h \in \widetilde{\mathcal{H}}, \quad \forall t \geq 0,$$

and hypocoercive on $\widetilde{\mathcal{H}}$ if the above inequality holds for some $\lambda > 0$.

The typical choice for $\widetilde{\mathcal{H}}$, when showing hypocoercivity, is $\mathcal{N}(L)^{\perp}$ endowed with a Sobolev norm e.g. $\widetilde{\mathcal{H}} = H^1/\mathcal{N}(L)$.

A couple of remarks should be made with respect to the above definitions.

Remark 12. *The first comment is that coercivity implies hypocoercivity with constant $C = 1$, as we already saw. The inverse statement is also true. If the constant C equals 1 in the definition of hypocoercivity, then for the inner product of $\widetilde{\mathcal{H}}$, λ-coercivity holds. This is just a consequence of Lummer-Phillips theorem.*

Remark 13. *The important feature of the constant C that appears in the definition of hypocoercivity is that it makes the property of hypocoercivity for a semi-group, invariant for* ***equivalent*** *norms. The importance of this property will be highlighted* *later. For now one should keep in mind that proving coercivity for an operator under a carefully designed norm automatically proves hypocoercivity for any equivalent norm.*

An example of how the coercivity condition can be relaxed to imply hypocoercivity is found in the following theorem presented in [31].

Theorem 16. *Let L be a generator of a contraction semi-group on a Hilbert space $\mathcal{H}$. Assume that there exists some $a > 0$ and a bounded operator M (on $\mathcal{H}$) s.t. the following condition holds,*

$$\forall h \in D(L) \cap \mathcal{H}, \quad t \geq 0, \qquad a\|h\|^2 \leq \langle Lh, h\rangle + \langle Lh, (M + M^*)h\rangle.$$

Then, it follows that L is hypocoercive on $\mathcal{H}$.

In the theorem that follows the key Sobolev space in which hypocoercivity is shown is the space $\mathcal{H}^1$ with norm

$$\|u\|_{\mathcal{H}^1} = \left(\|u\|^2 + \|Au\|^2 + \|Cu\|^2\right)^{1/2}.$$

We can now give a general theorem that proves hypocoercivity for our case:

Theorem 17. *Consider the operator $L = A^*A + B$ (with $B^* = -B$) and $C = [A, B]$. Assume that the following hold:*

(i) $[A, C] = [A^*, C] = 0$, $[A, A] = 0$,

(ii) $\|[A, A^*]h\| \leq \alpha \left(\|h\| + \|Ah\|\right)$ *for some* $\alpha > 0$,

(iii) $\|[B, C]h\| \leq \beta \left(\|Ah\| + \|A^2h\| + \|Ch\| + \|ACh\|\right)$ *for some* $\beta > 0$,

(iv) $A^*A + C^*C$ *is k-coercive.*

Then, there exists an inner product $((\cdot,\cdot))$ in $\widetilde{\mathcal{H}}$ (defining a norm equivalent to $\mathcal{H}^1$) and $\lambda > 0$ s.t. L is λ-coercive in the inner product $((\cdot,\cdot))$, i.e.

$$((h, Lh)) \geq \lambda((h,h)) \qquad \forall h \in \mathcal{H}^1/\mathcal{N}(L),$$

and hypocoercive on $\widetilde{\mathcal{H}}$. Furthermore, $((h,h)) = \|h\|^2 + a\|Ah\|^2 + 2b\langle Ah, Ch\rangle + c\|Ch\|^2$.

Proof. The proof of this theorem can be found in [71]. □

In what follows, we give the statement of result and proof tailored to our problem.

Theorem 18. *Assume a smooth solution $h(t,x,v)$ to the problem*

$$\partial_t h + v \cdot \nabla_x h - \nabla \mathcal{U}(x) \cdot \nabla_v h = \gamma(x) \triangle_v h - \gamma(x) v \cdot \nabla_v h,$$

with initial data $h_0 \in L^2(\mu)$ s.t $\int h_0 \, d\mu = 0$. We further assume a potential $\mathcal{U}(x) \in C^2(\mathbb{R}_x^{3N})$ and $\gamma(x) \in C^1(\mathbb{R}_x^{3N})$ that satisfy:

(i) $|\nabla^2 \mathcal{U}(x)| \leq C'$ (this condition can be relaxed to $|\nabla^2 \mathcal{U}(x)| \leq C'(1 + |\nabla \mathcal{U}(x)|)$) for $C' > 0$,

(ii) $\lambda_0 \leq \gamma(x) \leq \Lambda_0$ and $|\nabla_x \gamma(x)| \leq \Lambda_1$ with $\lambda_0, \Lambda_0, \Lambda_1 > 0$, and

(iii) the measure $e^{-\mathcal{U}(x)} dx$ satisfies a Poincaré inequality for a constant $\lambda > 0$.

It is proven that there exists a constant $C > 0$, such that for the given initial data

$$\|e^{-tL} h_0\|_{H^1(\mu)} \leq C e^{-\lambda t} \|h_0\|_{L^2(\mu)}.$$

Proof. The proof bears a striking resemblance to the strategy employed in the Hérau regularization. First, we consider the functional

$$\mathcal{E}(h) = \int h^2 \, d\mu + a \int |\nabla_v h|^2 \, d\mu + 2b \int \nabla_v h \cdot \nabla_x h \, d\mu + c \int |\nabla_x h|^2 \, d\mu,$$

and try to show that for carefully selected $a, b, c > 0$ we get

$$\frac{d}{dt}\mathcal{E}(h) \leq -C\Big(\int |\nabla_x h|^2 \, d\mu + \int |\nabla_v h|^2 \, d\mu + \int |\nabla^2_{vx} h|^2 \, d\mu + \int |\nabla^2_v h|^2 \, d\mu \Big)$$

for some $C > 0$.

Indeed, with computations very similar to the ones presented when proving regularization, one has that

$$\begin{aligned}
&\frac{d}{dt}\mathcal{E}(h) \leq \\
&- 2\lambda_0 \int |\nabla_v h|^2 \, d\mu - 2\lambda_0 a \int |\nabla^2_v h|^2 \, d\mu - 2b \int |\nabla_x h|^2 \, d\mu - 2\lambda_0 c \int |\nabla^2_{vx} h|^2 \, d\mu \\
&+ 2b \int \nabla_v h \cdot \nabla^2 \mathcal{U}(x) \nabla_v h \, d\mu - 2\lambda_0 a \int |\nabla_v h|^2 \, d\mu \\
&+ 2c \int \nabla_x h \cdot \nabla^2 \mathcal{U}(x) \nabla_v h \, d\mu - 2a \int \nabla_v h \cdot \nabla_x h \, d\mu - 2b \int \gamma(x) \nabla_v h \cdot \nabla_x h \, d\mu \\
&- 4b \int \gamma(x) \nabla^2_{vx} h : \nabla^2_v h \, d\mu \\
&- 2b \int \nabla_x \gamma(x) \cdot \nabla^2_v h \nabla_v h \, d\mu - 2c \int \nabla_x \gamma(x) \cdot \nabla^2_{vx} h \nabla_v h \, d\mu.
\end{aligned}$$

This implies that $\frac{d}{dt}\mathcal{E}(h)$ is bounded above by

$$
\begin{aligned}
&-2\lambda_0 \int |\nabla_v h|^2 \, d\mu - 2\lambda_0 a \int |\nabla_v^2 h|^2 \, d\mu - 2b \int |\nabla_x h|^2 \, d\mu - 2\lambda_0 c \int |\nabla_{vx}^2 h|^2 \, d\mu \\
&+ 2bC' \int |\nabla_v h|^2 \, d\mu - 2\lambda_0 a \int |\nabla_v h|^2 \, d\mu \\
&+ 2cC' \int |\nabla_x h||\nabla_v h| \, d\mu + 2a \int |\nabla_v h||\nabla_x h| \, d\mu + 2b\Lambda_0 \int |\nabla_v h||\nabla_x h| \, d\mu \\
&+ 4b\Lambda_0 \int |\nabla_{vx}^2 h||\nabla_v^2 h| \, d\mu \\
&+ 2b\Lambda_1 \int |\nabla_v^2 h||\nabla_v h| \, d\mu + 2c\Lambda_1 \int |\nabla_{vx}^2 h||\nabla_v h| \, d\mu.
\end{aligned}
$$

With a bit more work and putting terms together, we have

$$
\begin{aligned}
&\frac{d}{dt}\mathcal{E}(h) \leq \\
&\left(-2\lambda_0 + 2bC' - 2\lambda_0 a + \frac{(2cC' + 2a + 2b\Lambda_0)^2}{2b} + b\Lambda_1 + 2c\frac{\Lambda_1^2}{\lambda_0}\right) \int |\nabla_v h|^2 \, d\mu \\
&- \frac{3b}{2} \int |\nabla_x h|^2 \, d\mu + \left(-2\lambda_0 a + 4\frac{\Lambda_0^2}{\lambda_0}\frac{b^2}{c} + b\Lambda_1\right) \int |\nabla_v^2 h|^2 \, d\mu - \frac{\lambda_0 c}{2} \int |\nabla_{vx}^2 h|^2 \, d\mu.
\end{aligned}
$$

Under pretty much the same assumption on the parameters $a, b, c > 0$ like the ones we had when picking the parameters in Hérau technique, i.e. a, $\frac{b}{a}$, $\frac{c}{b}$, $\frac{a^2}{b}$, and $\frac{b^2}{ac}$ being chosen sufficiently small, it follows that

$$
\frac{d}{dt}\mathcal{E}(h) \leq -C\left(\int |\nabla_v h|^2 \, d\mu + \int |\nabla_x h|^2 \, d\mu\right).
$$

We now make use of the assumption that $e^{-\mathcal{U}(x)}dx$ satisfies a Poincaré inequality with constant λ. This implies that $d\mu(x, v)$ satisfies a Poincaré inequality with same constant, i.e.

$$
\int h^2 \, d\mu \leq \lambda \left(\int |\nabla_x h|^2 \, d\mu + \int |\nabla_v h|^2 \, d\mu\right),
$$

since $e^{-\frac{v^2}{2}}dv$ satisfies a Poincaré inequality with constant 1. This leads to

$$
\frac{d}{dt}\mathcal{E}(h) \leq -C\left(\int h^2 \, d\mu + \int |\nabla_x h|^2 \, d\mu + \int |\nabla_v h|^2 \, d\mu\right)
$$

for a new constant $C > 0$.

The inclusion of the mixed derivative term $\int \nabla_x h \cdot \nabla_v h \, d\mu$ in the functional $\mathcal{E}(h)$ has the same affect in proving hypocoercivity as it did for the proof of regularization. The evolution of this term provides the $-\int |\nabla_x h|^2 \, d\mu$ term necessary for closure. At the same time, it doesn't really alter the nature of the functional $\mathcal{E}(h)$. Indeed, if we choose $b^2/ac < 1$ (a condition that as we saw is necessary in the proof), then by a trivial C-S we get

$$
\begin{aligned}
-(1-\delta)\left(a\int |\nabla_v h|^2 \, d\mu + c\int |\nabla_x h|^2 \, d\mu\right) &\leq 2b\int \nabla_x h \cdot \nabla_v h \, d\mu \\
&\leq (1-\delta)\left(a\int |\nabla_v h|^2 \, d\mu + c\int |\nabla_x h|^2 \, d\mu\right)
\end{aligned}
$$

for sufficiently small $\delta > 0$.

Hence, it is shown that

$$
\frac{d}{dt}\mathcal{E}(h) \leq -K\mathcal{E}(h) \qquad \text{for some} \quad K > 0.
$$

The last part of the proof is to incorporate the regularity estimates we previously obtained to relax the assumptions of the initial data to $h_0 \in L^2(\mu)$, i.e.

$$
\|h\|_{H^1(\mu)} \leq e^{-K(t-t_0)}\|h(t_0)\|_{H^1(\mu)} \leq C\frac{e^{-K(t-t_0)}}{t_0^{3/2}}\|h_0\|_{L^2(\mu)} \qquad \text{for any} \quad 0 < t_0 < 1.
$$

$\square$

5.2 Relative Entropy

Another technique to study rate of convergence to the global stationary state $\mathcal{M}_{eq}(x,v)$, is via the use of relative entropy, see [15]. The relative entropy $H(f|g)$ between two probability density functions f, g is given by

$$H(f|g) = \iint f \log \frac{f}{g}\, dv\, dx.$$

The main idea behind the study of the convergence rate for the relative entropy $H(f|\mathcal{M}_{eq})$, is that due to the Csiszár-Kullback-Pinsker inequality for probability measures, the relative entropy controls the square of the L^1 distance between f and $\mathcal{M}_{eq}(x,v)$, i.e.

$$\|f - \mathcal{M}_{eq}\|_{L^1(\mathbb{R}^{3N}_x \times \mathbb{R}^{3N}_v)} \leq \sqrt{2H(f|\mathcal{M}_{eq})}.$$

The entropy dissipation rate (Fisher information) is

$$-\frac{d}{dt} H(f|\mathcal{M}_{eq}) = \iint f \left|\nabla_v \log \frac{f}{\mathcal{M}_{eq}}\right|^2 dv\, dx = \iint f \left|\nabla_v \log \frac{f}{\mathcal{M}}\right|^2 dv\, dx.$$

The dissipation rate vanishes iff $f = \rho\mathcal{M}(v)$, where $\rho = \int f\, dv$. This is exactly the difficulty in the entropy method approach as it appears in many types of kinetic equations with the most notable example that of the Boltzmann equation treated in [16]. In equations where the dissipative term acts only in the velocity space, the total entropy vanishes for states that belong in a subfamily of the Gaussian distribution. For this type of equations the transport term is responsible for driving the system away from these local equilibria states, thus making the entropy positive again and giving it space to dissipate more. The interplay between the transport and the dissipative terms drives the equation to a global equilibrium state with the relative entropy approaching 0 in a non monotone way.

To see the difference that the transport term makes consider this example of the Ornstein-Uhlenbeck equation borrowed from [52]. In this, there is a probability density $f(v,t) : \mathbb{R}_v^{3N} \times \mathbb{R}_+ \to \mathbb{R}_+$ that solves the equation

$$\partial_t f = \nabla_v \cdot (\nabla_v f + vf).$$

The unique global equilibrium state for this equation is the standard Gaussian in velocity space $\mathcal{M}(v)$. With the help of the log-Sobolev inequality (see appendix), one gets

$$-\frac{d}{dt}H(f|\mathcal{M}) \geq CH(f|\mathcal{M}),$$

which yields exponential decay of the entropy and subsequently of a solution to the equilibrium state.

In the above spirit, the log-Sobolev inequality here implies that

$$-\frac{d}{dt}H(f|\mathcal{M}_{eq}) \geq CH(f|\rho\mathcal{M})$$

which does not offer a closed inequality. The missing bit of information is the distance of a hydrodynamic variable ρ from $e^{-\mathcal{U}(x)}$

$$H(f|\mathcal{M}_{eq}) - H(f|\rho\mathcal{M}) = H(\rho|e^{-\mathcal{U}(x)}).$$

We assume that that if we focus in higher order time derivatives of the relative entropy of f with respect to local equilibria states we can provide a system of differential inequalities that is closed. It can be proven that

$$\frac{d^2}{dt^2}H(f|\rho\mathcal{M})\Big|_{\text{local eq.}} = \int \rho \left|\nabla_x \log \frac{\rho}{e^{-\mathcal{U}(x)}}\right|^2 dx.$$

The above relation is indicative of the closure we can obtain since it connects the second time derivative of $H(f|\rho\mathcal{M})$ with the Fisher information of $H(\rho|e^{-\mathcal{U}(x)})$.

Under suitable assumptions on the regularity of the solution and potential $\mathcal{U}(x)$, it can be shown after detailed analysis that the following system of inequalities holds,

$$
\begin{aligned}
&-\frac{d}{dt}H(f|\mathcal{M}_{eq}) \geq CH(f|\rho\mathcal{M}) \\
&\frac{d^2}{dt^2}H(f|\rho\mathcal{M}) \geq K\left(H(f|\mathcal{M}_{eq}) - H(f|\rho\mathcal{M})\right) - C(f_0)H(f|\rho\mathcal{M})^{1-\epsilon}
\end{aligned}
$$

for any $\epsilon > 0$, and some $K > 0$, $C(f_0) > 0$ that depend on ϵ. $C(f_0)$ is a constant that depends on initial solution profile f_0. This system of inequalities gives algebraic convergence rate as it is well explained in [15].

The precise statement of the theorem is:

Theorem 19. *Assume a probability density $f_0(x,v) \in L^1$ (initial data) that is controlled by Maxwellians, in the sense that there exists $a, A > 0$ s.t.*

$$a\mathcal{M}_{eq}(x,v) \leq f_0(x,v) \leq A\mathcal{M}_{eq}(x,v).$$

Assume also a globally smooth solution $f(x,v,t)$ to the F-P equation, and a quadratic potential $\mathcal{U}(x)$ with an added perturbation term, i.e.

$$\mathcal{U}(x) = \omega\frac{|x|^2}{2} + \Phi(x), \qquad \omega > 0, \quad \Phi(x) \in H^\infty(\mathbb{R}^n_x).$$

Then, for every $\epsilon > 0$, there exists a constant $C_\epsilon(f_0)$ s.t.

$$\|f - \mathcal{M}_{eq}\|_{L^1(\mathbb{R}^{3N}_x \times \mathbb{R}^{3N}_v)} \leq C_\epsilon(f_0)t^{-1/\epsilon}.$$

5.3 Entropy & Commutators

We have already modified the technique by Hérau to show hypocoercivity for the operator $Lh = v \cdot \nabla_x h - \nabla \mathcal{U}(x) \cdot \nabla_v h - (\nabla_v - v) \cdot \nabla_v h$, with $L^2(\mu)$ data. We can pretty much follow the same lines of proof, to relax the initial data to belong in the space $L \log L(\mu)$.

The Lyapunov functional used here takes the form

$$\begin{aligned}\mathcal{E}(h) = \int h \log h \, d\mu + a \int h |\nabla_v \log h|^2 \, d\mu \\ + 2b \int h \nabla_v \log h \cdot \nabla_x \log h \, d\mu + c \int h |\nabla_x \log h|^2 \, d\mu.\end{aligned}$$

At the same time, we have given a regularization result that allows for more general initial data (measure initial data), under the extra assumption of bounded first moments, which is used in:

Theorem 20. *Assume an initial profile $f_0(x, v) \in L^1(\mathbb{R}^n_v \times \mathbb{R}^n_x)$ that has finite second moments, i.e.*

$$\iint (|x|^2 + |v|^2) f_0 \, dv \, dx < \infty.$$

Assume also that the potential $\mathcal{U}(x) \in C^2(\mathbb{R}^{3N}_x)$ satisfies $|\nabla^2 \mathcal{U}(x)| \leq C$ for all $x \in \mathbb{R}^{3N}_x$, and that $e^{-\mathcal{U}(x)}$ satisfies a log-Sobolev inequality. It can be proven that

$$\iint f \log \frac{f}{\mathcal{M}_{eq}} \, dv \, dx \leq O(e^{-kt}) \quad \textit{for} \quad t \geq 1,$$

with $k > 0$, which implies exponential decay in L^1.

5.4 Hypocoercivity à la Dolbeault-Mouhot-Schmeiser

We present here another approach borrowed from [20], which strengthens slightly the Villani theory in some sense. The main framework will once again be the Hilbert space $\mathcal{H} = L^2(\mathcal{M}_{eq}^{-1} dvdx)$ with inner product denoted by $\langle \cdot, \cdot \rangle$ as in

$$\langle f, g \rangle = \iint \frac{f}{\mathcal{M}_{eq}} \frac{g}{\mathcal{M}_{eq}} \mathcal{M}_{eq} \, dv \, dx.$$

The idea is similar to the general idea of constructing an entropy $H(f)$ that allows a Gronwall type of inequality, like in the Villani theory, but with a difference. The functional entropy will satisfy $-\frac{d}{dt} H(f) \geq C H(f)$ for some $C > 0$ and is shown to be equivalent to the square of the $L^2(\mu)$ norm rather than a Sobolev type of norm. So this theory suffices to show convergence in $L^2(\mu)$ without any regularization properties for e^{-tL}.

We consider the general kinetic equation $\partial_t f + \mathcal{T} f = \mathcal{C} f$ with transport term $\mathcal{T} = v \cdot \nabla_x - \nabla \mathcal{U}(x) \cdot \nabla_v$ and a general collision operator $\mathcal{C}$ which is mass preserving and acting only in velocity space. The operator $\mathcal{T}$ is anti-symmetric and we make the extra assumption of a self-adjoint $\mathcal{C}$ ($\mathcal{C}^* = \mathcal{C}$). We further assume that $\mathcal{C}$ is dissipative in the sense $-\langle \mathcal{C} f, f \rangle \geq 0$ for all $f \in \mathcal{D}(\mathcal{C})$, but not coercive (no $C > 0$ s.t. $\langle \mathcal{C} f, f \rangle \geq C ||f||^2$).

We finally make the assumption of considering the k.e for $f - \mathcal{M}_{eq}$, so that the new density f corresponds to perturbations about the equilibrium zero state. The new f will be normalized as in $\iint f \, dv \, dx = 0$.

We now introduce the orthogonal projection P onto the null space of $\mathcal{C}$ (space of local equilibria $\mathcal{N}(\mathcal{C})$), i.e.

$$Pf = \left(\int f\,dv\right)\mathcal{M}(v) = \rho\mathcal{M}(v).$$

It is natural to assume at least one conservation law (mass) for the k.e, which would simply imply

$$\int \mathcal{C}f\,dv = 0 \quad \Longrightarrow \quad P\mathcal{C} = 0.$$

The entropy functional that will be used and was inspired by the work of [30] is

$$H(f) = \frac{1}{2}\|f\|^2 + \epsilon\langle Af, f\rangle \qquad \epsilon > 0,$$

for the operator $A = (\mathrm{I} + (\mathcal{T}P)^*\mathcal{T}P)^{-1}(\mathcal{T}P)^*$.

The dissipation rate $\mathcal{D}(f) = -\frac{d}{dt}H(f)$ is computed to be

$$\begin{aligned}\mathcal{D}(f) &= -\langle \mathcal{C}f, f\rangle - \epsilon\langle A(\mathcal{C}-\mathcal{T})f, f\rangle - \epsilon\langle Af, (\mathcal{C}-\mathcal{T})f\rangle \\ &= -\langle \mathcal{C}f, f\rangle + \epsilon\langle A\mathcal{T}Pf, f\rangle + \epsilon\langle A\mathcal{T}(I-P)f, f\rangle - \epsilon\langle \mathcal{T}Af, f\rangle - \epsilon\langle A\mathcal{C}f, f\rangle.\end{aligned}$$

The last term should equal $-\epsilon\langle (A+A^*)\mathcal{C}f, f\rangle$ but is simplified since $P\mathcal{C} = 0$ implies $\langle A^*\mathcal{C}f, f\rangle = 0$. In the attempt to bound all the terms of the dissipation rate and obtain a rate of convergence for $H(f)$ we make "natural" assumptions which are easily applicable in the case of a Fokker-Planck operator $\mathcal{C}$.

The first two are the assumptions of micro & macro coercivity which have already been employed in the Villani treatment of relative entropy when in the attempt to derive the system of differential inequalities that we presented.

The **microscopic** coercivity assumption states that $-\mathcal{C}$ is coercive on $\mathcal{N}(\mathcal{C})^{\perp}$, i.e.

$$-\langle \mathcal{C}f, f\rangle \geq \lambda_m \|(\mathrm{I}-P)f\|^2 \qquad \text{for some} \quad \lambda_m > 0. \tag{5.1}$$

Microscopic coercivity for $\mathcal{C} = \nabla_v \cdot (\nabla_v + v\cdot)$ boils down to the Poincaré inequality for $L^2(\mathcal{M}(v)dv)$ which holds trivially,

$$\begin{aligned} -\langle \mathcal{C}f, f\rangle &= \int \frac{1}{e^{-\mathcal{U}(x)}} \int \mathcal{M}(v) \Big|\nabla_v \frac{(\mathrm{I}-P)f}{\mathcal{M}(v)}\Big|^2 \, dv\, dx \\ &\geq \int \frac{1}{e^{-\mathcal{U}(x)}} \int \mathcal{M}(v) \Big|\frac{(\mathrm{I}-P)f}{\mathcal{M}(v)}\Big|^2 \, dv\, dx = \|(\mathrm{I}-P)f\|^2. \end{aligned}$$

The second condition assumed is the **macroscopic** coercivity condition

$$\|\mathcal{T}Pf\|^2 \geq \lambda_M \|Pf\|^2 \qquad \text{for some} \quad \lambda_M > 0, \tag{5.2}$$

which amounts to the validity of the Poincaré inequality with measure $L^2(e^{-\mathcal{U}(x)}dx)$ for the macroscopic variable ρ, i.e.

$$\int e^{-\mathcal{U}} \Big|\nabla_x \left(\frac{\rho}{e^{-\mathcal{U}}}\right)\Big|^2 dx \geq \lambda_M \int e^{-\mathcal{U}(x)} \Big|\frac{\rho}{e^{-\mathcal{U}}}\Big|^2 dx.$$

This assumption is not automatically satisfied for the equation at hand and boils down to finding the measures $e^{-\mathcal{U}(x)}$ for which a Poincaré inequality holds. In the appendix we present a sufficient condition for the Poincaré inequality to hold for the measure $e^{-\mathcal{U}(x)}$.

Employing (5.1), (5.2) and the extra assumption

$$P\mathcal{T}P = 0, \tag{5.3}$$

it can be shown that $A = PA$, as well as the fact that A and $\mathcal{T}A$ are bounded operators since it is easily proven that

$$(i) \quad \|Af\| \leq \frac{1}{2}\|(\mathrm{I}-P)f\| \qquad \& \qquad (ii) \quad \|\mathcal{T}Af\| \leq \|(\mathrm{I}-P)f\|.$$

Condition (i) is the reason that the entropy functional $H(f)$ is equivalent to $\|\cdot\|^2$.

The last assumption that is employed to give a decay rate for entropy is

$$\|\mathcal{T}A(\mathrm{I}-P)f\| + \|A\mathcal{C}f\| \leq C_M\|(\mathrm{I}-P)f\| \qquad \text{for some} \quad C_M > 0. \tag{5.4}$$

All the above can be combined in:

Theorem 21. *Consider the kinetic equation $\partial_t f + \mathcal{T}f = \mathcal{C}f$ where operators $\mathcal{T}, \mathcal{C}$ have the properties mentioned in the first two paragraphs.*

Assume that the following are satisfied

(i) $-\langle \mathcal{C}f, f\rangle \geq \lambda_m\|(\mathrm{I}-P)f\|^2$ for some $\lambda_m > 0$

(ii) $\|\mathcal{T}Pf\|^2 \geq \lambda_M\|f\|^2$ for some $\lambda_M > 0$

(iii) $P\mathcal{T}P = 0$

(iv) $\|A\mathcal{T}(\mathrm{I}-P)f\| + \|A\mathcal{C}f\| \leq C_m\|(\mathrm{I}-P)f\|$ for some $C_m > 0$.

Given initial data $f_0 \in L^2(d\mu)$ ($f_0 \geq 0$, $\iint f_0 = 1$), it can be shown that

$$-\frac{d}{dt}H(f) \geq \lambda H(f) \qquad \textit{for some} \quad \lambda > 0,$$

for the entropy we have defined. This implies that there exists $C > 0$ s.t.

$$\|e^{-tL}f_o\| \leq Ce^{-\lambda t}\|f_0\| \qquad t \geq 0.$$

Proof. See [20]. □

For the case of the inhomogeneous Fokker-Planck operator $\mathcal{C} = \nabla_v \cdot (\nabla_v + v\cdot)$ the assumptions mentioned in the theorem above are met in (see [20] for details) the following theorem.

Theorem 22. *Assume that the external potential* $\mathcal{U}(x)$ *has the following properties:*

(i) $\mathcal{U}(x) \in C^2(\mathbb{R}^{3N})$, *with* $\quad \int e^{-\mathcal{U}(x)} dx = 1$

(ii) $e^{-\mathcal{U}(x)}$ *satisfies a Poincaré type of inequality*

(iii) $\triangle \mathcal{U}(x) \leq c_1 + \frac{c_2}{2} |\nabla \mathcal{U}(x)|^2$ *for* $c_1 > 0$ *and* $c_2 \in (0,1)$, *and*

(iv) $|\nabla^2 \mathcal{U}(x)| \leq c_3(1 + |\mathcal{U}(x)|)$.

Then, the solution of $\partial_t f + \mathcal{T} f = \mathcal{C} f$ *with initial data in* $L^2(d\mu)$ *decays exponentially fast towards the global equilibrium state.*

Chapter 6: Diffusive Limit

In this last Chapter, we examine the hydrodynamic limit of the Cauchy problem $\partial_t f_\epsilon + L_\epsilon f_\epsilon = 0$ with $f_\epsilon(0,.,.) = f_\epsilon\Big|_{t=0}(.,.)$, as $\epsilon \to 0$, for the Fokker-Planck operator

$$L_\epsilon f = \frac{1}{\epsilon}(v \cdot \nabla_x f - \nabla \mathcal{U}(x) \cdot \nabla_v f) - \frac{1}{\epsilon^2} \nabla_v \cdot \left(G(x) \mathcal{M}_{eq} \nabla_v \left(\frac{f}{\mathcal{M}_{eq}} \right) \right).$$

There is an obvious analogy in this study, with the Kramers-Smouchowski limit for SDEs, see e.g. [23]. We begin with a formal argument based on the Hilbert expansion of f_ϵ, which gives an expansion for the hydrodynamic variable $\rho_\epsilon = \int f_\epsilon \, dv$ as well. After we explicitly compute the first terms of the expansion, i.e. f_0, $f_1, \ldots$, we obtain an equation for the limit hydrodynamic variable ρ (first term of the ρ_ϵ expansion), which is the Smoluchowski equation

$$\partial_t \rho = \nabla_x \cdot (G^{-1}(x)(\nabla_x \rho + \nabla \mathcal{U}(x)\rho)).$$

Two techniques are employed to establish the rigorous limit under different a priori energies. In the first, we prove weak convergence $\rho_\epsilon \rightharpoonup \rho$ for mild solutions, using weak compactness principles and initial data in weighted L^2 space. The second technique relies on a relative entropy argument, uses entropic initial data and gives convergence in L^1, uniformly on any time interval $[0, T]$, for $T > 0$.

6.1 Formal Result

6.1.1 Hilbert Expansion

The study of the hydrodynamic limit as $\epsilon \to 0$ will begin with the Hilbert expansion, and we will see the problems that arise in this analysis. We expand f_ϵ in powers of ϵ as in

$$f_\epsilon = f_0 + \epsilon f_1 + \epsilon^2 f_2 + \dots$$

and after we substitute in the FP equation $\partial_t f_\epsilon + L_\epsilon f_\epsilon = 0$, we balance powers of ϵ. This procedure leads to a cascade of equations which in our case can be solved explicitly at least for the initial terms. This hierarchy of equations is

$$\mathcal{C}(f_0) = 0 \qquad \text{(0'th order term)}$$

$$\mathcal{T}(f_0) = \mathcal{C}(f_1) \qquad \text{(1st order term)}$$

$$\partial_t f_{j-2} + \mathcal{T}(f_{j-1}) = \mathcal{C}(f_j) \qquad \text{(j'th order term for } j \geq 2\text{)}.$$

Solving the first equation, one trivially gets

$$f_0 = \rho_0 \mathcal{M},$$

where ρ_0 is the hydrodynamical variable of highest order in the expansion of ρ_ϵ, defined by $\rho_0 = \int f_0 \, dv$. A typical feature of the expansion is that the hydrodynamic variable $\rho_\epsilon = \int f_\epsilon \, dv$ is also expanded in a power series of ϵ where each term ρ_i is given by $\rho_i = \int f_i \, dv, \quad i \geq 0$.

The next two terms are,

$$f_1 = -v^T G^{-1} \mathcal{B} \rho_0 \mathcal{M}(v) + \rho_1 \mathcal{M}(v)$$

and

$$f_2 = \frac{1}{2}\Big(v^T G^{-1}\mathcal{B} v^T G^{-1}\mathcal{B}\rho_0 \mathcal{M}(v) - G^{-1}\mathcal{B}\cdot G^{-1}\mathcal{B}\rho_0\mathcal{M}(v)\Big)$$
$$- v^T G^{-1}\mathcal{B}\rho_1\mathcal{M}(v) + \rho_2\mathcal{M}(v),$$

for the vector field $\mathcal{B} = \nabla_x + \nabla_x \mathcal{U}(x)$.

A rigorous approach to the Hilbert expansion for f_ϵ and how it can be used to study the limit as $\epsilon \to 0$ is given by the following procedure. In general we truncate the expansion to some order i.e.

$$f_\epsilon = f_0 + \epsilon f_1 + \ldots + \epsilon^m f_m + \mathcal{R}_\epsilon$$

and set up the equation for the remainder term $\mathcal{R}_\epsilon(t, x, v)$.

In fact the truncated expansion could help us obtain a rigorous result for the hydrodynamic limit if

(i) $\mathcal{R}_\epsilon$ is a term of some order $l \geq 0$ in ϵ, i.e. $\mathcal{R}_\epsilon = \epsilon^l \mathcal{R}'_\epsilon$ for some function $\mathcal{R}'_\epsilon$ which we should rigorously be able to show is of order ϵ^0 (l does not necessarily have to be equal to $m+1$) and

(ii) we can prove that for an appropriate selected space equipped with a norm $\|\cdot\|$, the functions $\mathcal{R}'_\epsilon, f_1, \ldots f_m$ are sufficiently regular uniformly in time (or at least for a finite time interval $[0,T]$ for $T>0$), in the sense that

$$\sup_{\epsilon>0, t\in I} \{\|f_1\|, \|f_2\|, \ldots, \|f_m\|, \|\mathcal{R}'_\epsilon\|\} < \infty,$$

where $I = [0,\infty)$, or $[0,T]$. This would be enough to establish that $f_\epsilon \to \rho_0\mathcal{M}$ in the $\|\cdot\|$ norm, uniformly on the time interval I.

In our case, it is enough to truncate after the second term as in $f_\epsilon = f_0 + \epsilon f_1 + \epsilon^2 f_2 + \mathcal{R}_\epsilon$ and substitute in the F-P equation. The equation for the remainder term is

$$\partial_t \mathcal{R}_\epsilon + \frac{1}{\epsilon}\mathcal{T}(\mathcal{R}_\epsilon) - \frac{1}{\epsilon^2}\mathcal{C}(\mathcal{R}_\epsilon) = \mathcal{F}(t,x,v),$$

where $\mathcal{F}(t,x,v) = -\epsilon(\mathcal{T}(f_2) + \partial_t f_1) - \epsilon^2 \partial_t f_2$.

The study of the remainder equation in terms of its stability w.r.t. $\mathcal{F}$ will allow the rigorous justification of the limiting procedure. In reality, one should be able show that $\mathcal{R}_\epsilon$ is of order ϵ or less and be able to establish regularity results for the hydrodynamical variables that appear in $\mathcal{F}$. In this chapter we are going to use two methods different than the expansion we presented above based on two different types of a priori estimates. These methods either rely on compactness arguments, or functional “entropic” inequalities and the study of a relative entropy. Before that, we need to proceed with the formal argument and derive the limit equation of ρ_ϵ.

6.1.2 Equation for the limit Hydrodynamic Variable ρ

In the study of the limit case $\epsilon \to 0$, we want to derive the equation for the hydrodynamic term ρ, which at least formally should be the limit of ρ_ϵ.

Integrating the FP equation in velocity space we obtain

$$\partial_t \rho_\epsilon + \nabla_x \cdot J_\epsilon = 0, \tag{6.1}$$

where $J_\epsilon(t,x) = \frac{1}{\epsilon}\int v f_\epsilon \, dv$ is the flux vector.

To this end, we want to derive an expression for J_ϵ and see the terms that are involved in it. In the derivation of the equation for the first moment one multiplies the FP by v and integrates. The equation for J_ϵ is

$$\epsilon^2 \partial_t J_\epsilon(t,x) + \nabla_x \cdot \mathbb{P}_\epsilon(t,x) + \nabla \mathcal{U}(x)\rho_\epsilon = -G(x)J_\epsilon(t,x), \tag{6.2}$$

where $\mathbb{P}_\epsilon(t,x) = \int v \otimes v f_\epsilon \, dv$ is the pressure tensor.

As will be shown rigorously, the main contribution to the above equation comes from the r.h.s as well as the second and third terms of the l.h.s. Indeed, treating the pressure tensor $\mathbb{P}_\epsilon(t,x)$, we see that

$$\int v_i v_j f_\epsilon \, dv = -\int \partial_{v_i}(\mathcal{M}) v_j \frac{f_\epsilon}{\mathcal{M}} \, dv = \int \delta_{ij} f_\epsilon \, dv + \int \mathcal{M} \partial_{v_i} \left(\frac{f_\epsilon}{\mathcal{M}} \right) v_j \, dv,$$

which leads to a second term in (6.2) that equals

$$\nabla_x \cdot \mathbb{P}_\epsilon(t,x) = \nabla_x \rho_\epsilon + \nabla_x \cdot \int \mathcal{M} \nabla_v \left(\frac{f_\epsilon}{\mathcal{M}} \right) \otimes v \, dv.$$

With the above, (6.2) now becomes

$$\begin{aligned} J_\epsilon(t,x) = -G^{-1}(x) \left(\nabla_x \rho_\epsilon + \nabla \mathcal{U}(x) \rho_\epsilon \right) - \epsilon^2 G^{-1}(x) \partial_t J_\epsilon(t,x) \\ - G^{-1}(x) \nabla_x \cdot \int \mathcal{M}(v) \nabla_v \left(\frac{f_\epsilon}{\mathcal{M}} \right) \otimes v \, dv. \end{aligned} \tag{6.3}$$

The last term in (6.3) contains the second term in the expansion of $\mathbb{P}_\epsilon(t,x)$ which can be shown to be of order ϵ if one uses an appropriate a priori energy estimate e.g. in $L^2(\mu)$. This implies that in the limit $\epsilon \to 0$, we should be able to establish rigorously that $\mathbb{P}_\epsilon(t,x) \simeq \rho(t,x)\mathrm{I}$. The previous to last term $\epsilon^2 G^{-1}(x) \partial_t J_\epsilon$ should also be shown to be a term of order ϵ, as long as we give an appropriate interpretation to a solution $J_\epsilon(t,x)$ of equation (6.3). What is implied by this is that

$$J_\epsilon(t,x) = -G(x)^{-1} \left(\nabla_x \rho_\epsilon + \nabla \mathcal{U}(x) \rho_\epsilon \right) + \text{lower order terms in } \epsilon \ldots.$$

Finally, as $\epsilon \to 0$, the system of equations (6.1)-(6.2) converges to

$$\partial_t \rho + \nabla_x \cdot J = 0$$

$$J = -G(x)^{-1}(\nabla_x \rho + \nabla \mathcal{U}(x)\rho),$$

with $f_\epsilon(t,x,v) \simeq \rho(t,x)\mathcal{M}(v)$.

All this is enough to suggest that the equation for ρ is

$$\partial_t \rho = \nabla_x \cdot \left(G^{-1}(x)(\nabla_x \rho + \nabla \mathcal{U}(x)\rho)\right).$$

To understand how the formal argument can be used to establish the rigorous limit, we need to formulate in what way a solution f_ϵ to (6.3) makes sense. Here inspired by the analysis in the next section, we can consider that (6.3) is understood in the "weak" sense.

Assume a test function $\phi \in C_c^\infty(\mathbb{R}_+, \mathbb{R}^n_x, \mathbb{R}^n_v)$. Multiply the equation for J_ϵ by $\phi(t,x,v)$ and integrate in v, x, t. The goal is to provide a bound for the integral terms

$$\epsilon^2 \int_0^t \iint G^{-1} J_\epsilon \partial_t \phi \, dx\, dv\, ds \quad \text{and} \quad \int_0^t \iint \mathcal{M}\left(\nabla_v \left(\frac{f_\epsilon}{\mathcal{M}}\right) \otimes v\right) G^{-1} \nabla_x \phi \, dx\, dv\, ds$$

which we should be able to show are of order ϵ or less. In this direction it would be enough to elaborate on the use of an energy estimate (assuming bounded $L^2(\mu)$ initial data) which implies the bounded term

$$\frac{1}{\epsilon^2} \int_0^t \iint \left| G^{1/2} \nabla_v \left(\frac{f_\epsilon}{\mathcal{M}}\right)\right|^2 \mathcal{M}\, dx\, dv\, ds < \infty \qquad \text{for any } t > 0.$$

The finite energy itself implies, (since $J_\epsilon = \frac{1}{\epsilon} \int \mathcal{M} \nabla_v \left(\frac{f_\epsilon}{\mathcal{M}}\right) dv$)

$$\int_0^t \int \left| G^{1/2} J_\epsilon \right|^2 dx\, ds < \infty \qquad \text{for any } t > 0.$$

Of course, after using the energy estimate to find the order in ϵ of all terms in the equation for J_ϵ, and using compactness to get convergent subsequences, there still remains open the question of passing to the limit in ϵ.

6.2 Diffusive Limit via Weak Compactness and Proof of Theorem 1

6.2.1 Main Result

We are now in position to state and prove the main result about the limit of ρ_ϵ using compactness. Before we go into the statement of the theorem, we need to define a weak solution for (3.1).

Mild Solution

The discussion of weak formulation for (3.1) can start with a weak formulation of the stationary problem. The evolution equation can be written down in the form $\partial_t h + Lh = 0$, where the operator L is

$$L = v \cdot \nabla_x - \nabla \mathcal{U}(x) \cdot \nabla_v - \nabla_v \cdot G \nabla_v + v \cdot G \nabla_v.$$

The weak formulation of the stationary problem $Lh = 0$, gives a solution $h \in \mathcal{D}'(\mathbb{R}^{n,n}_{x,v})$. In fact, we shall assume that $-L$ generates the continuous semigroup $(e^{-tL})_{t \geq 0}$. We have also shown that solutions to the problem $\partial_t h + Lh = 0$ remain bounded in $L^2_{\mathcal{M}_{eq}} \cap L^\infty$.

Thus, a distributional solution $h(t, x, v)$ for the non-stationary problem will be a solution

$$h \in C(\mathbb{R}_+; \mathcal{D}'(\mathbb{R}^{n,n}_{x,v})) \cap L^\infty_{loc}(\mathbb{R}_+; L^2_{\mathcal{M}_{eq}} \cap L^\infty).$$

We now proceed to the proof of theorem 1.

6.2.2 A priori Energy & Weak Compactness

In order to study the limit $\epsilon \to 0$, we begin with the a priori estimate in $L^2_{\mathcal{M}_{eq}}(\mathbb{R}^{n,n}_{x,v})$. That is an energy estimate for $h_\epsilon(t)$ in $L^2(d\mu)$ (where $h_\epsilon(t)$ is an abbreviated notation for $h_\epsilon(t,x,v)$), i.e

$$\frac{1}{2}\int h_\epsilon^2(t)\,d\mu + \frac{1}{\epsilon^2}\int_0^t \int \left|G^{1/2}(x)\nabla_v h_\epsilon(s)\right|^2 d\mu\,ds = \frac{1}{2}\int h_\epsilon^2(0)\,d\mu.$$

We proceed in the decomposition of $f_\epsilon(t)$ in the following manner

$$f_\epsilon = \mathcal{M}(v)(\rho_\epsilon + \tilde{g}_\epsilon)$$

where the hydrodynamic variable ρ_ϵ has already been defined by

$$\rho_\epsilon = \int f_\epsilon\,dv$$

and a deviation $\tilde{g}_\epsilon$ from the local equilibrium state $\rho_\epsilon\mathcal{M}(v)$, that satisfies

$$\int \tilde{g}_\epsilon \mathcal{M}(v)\,dv = 0.$$

Integrating in velocity we obtain the hydrodynamic equation for ρ_ϵ,

$$\partial_t \rho_\epsilon + \frac{1}{\epsilon}\nabla_x \cdot \int \mathcal{M}(v)\nabla_v \tilde{g}_\epsilon\,dv = 0. \tag{6.4}$$

To simplify the analysis we consider the basic assumption $\inf \mathcal{U}(x) > -\infty$ which result in the following finite intervals for finite initial energy,

$$\int \rho_\epsilon^2\,dx < \infty, \qquad \iint \tilde{g}_\epsilon^2 \mathcal{M}(v)\,dv\,dx < \infty \qquad \forall t \geq 0.$$

The first bound is proven by a simple Jensen inequality on the $L^2(\mathcal{M}_{eq} dv\,dx)$ estimate of h_ϵ. We also have the energy estimate,

$$\frac{1}{\epsilon}\int_0^T \iint |G^{1/2}\nabla_v \tilde{g}_\epsilon|^2 \mathcal{M}(v)\,dv\,dx\,ds < \infty \quad \text{for any} \quad T > 0.$$

We can now, after picking a sequence of $\epsilon_i \to 0$, extract a subsequence which w.l.o.g. we still call ϵ_i s.t.

$$\rho_{\epsilon_i} \rightharpoonup \rho \quad \text{weakly in} \quad L^2(dx) \quad \forall t \geq 0,$$

$$\tilde{g}_{\epsilon_i} \rightharpoonup \tilde{g} \quad \text{weakly in} \quad L^2(\mathcal{M}(v)dvdx) \quad \forall t \geq 0,$$

$$\frac{1}{\epsilon_i} G^{1/2} \nabla_v \tilde{g}_{\epsilon_i} \rightharpoonup J \quad \text{weakly in} \quad L^2(\mathcal{M}(v)dvdxdt).$$

It is important to notice that we actually want something stronger than just ρ_ϵ being weakly compact in $L^2(dx) \quad \forall t \geq 0$. We actually want a uniform in time type of convergence so that we don't have a problem when we later pass to the limit in integrals of time. For this, we are actually proving that ρ_ϵ is compact in $C([0,T], \mathrm{w} - L^2(dx))$ in the lemma that follows.

Lemma 6. *ρ_ϵ is compact in $C([0,T], w - L^2(dx))$ i.e.*

$$\rho_\epsilon \rightharpoonup \rho \quad in \quad C([0,T], w - L^2(dx)).$$

Proof. Consider the functional $H(t) = \int \phi(x)\rho_\epsilon(x,t)dx$, $0 < t < T$, for a fixed $T > 0$ and $\phi \in C_c^\infty(\mathbb{R}_x^n)$. $H(t)$ can be proven pointwise finite for any $0 < t < T$, using Cauchy-Schwartz inequality and assuming always finite initial energy.

Now if we consider $t_1, t_2 > 0$ s.t. $0 \le t_1 \le t_2 \le T$, we get

$$
\begin{aligned}
&H(t_2) - H(t_1) = \int_{t_1}^{t_2} \int \phi(x) \partial_t \rho_\epsilon \, dx\, ds \qquad \text{(Using (6.4))} \\
&= \frac{1}{\epsilon} \int_{t_1}^{t_2} \iint \nabla_x \phi(x) \cdot \nabla_v \tilde{g}_\epsilon \mathcal{M}(v)\, dv\, dx\, ds \\
&\le \left(\int_{t_1}^{t_2} \iint |G^{-\frac{1}{2}} \nabla_x \phi(x)|^2 \mathcal{M}(v)\, dv\, dx\, ds \right)^{\frac{1}{2}} \left(\int_{t_1}^{t_2} \iint \frac{|G^{\frac{1}{2}} \nabla_v \tilde{g}_\epsilon|^2}{\epsilon^2} \mathcal{M}(v)\, dv\, dx\, ds \right)^{\frac{1}{2}} \\
&\le (t_2 - t_1)^{\frac{1}{2}} \left(\int |G^{-\frac{1}{2}} \nabla_x \phi(x)|^2 dx \right)^{\frac{1}{2}} \left(\int_{t_1}^{t_2} \iint \frac{|G^{\frac{1}{2}} \nabla_v \tilde{g}_\epsilon|^2}{\epsilon^2} \mathcal{M}(v)\, dv\, dx\, ds \right)^{\frac{1}{2}} \\
&\le C(t_2 - t_1)^{\frac{1}{2}}.
\end{aligned}
$$

By a density argument, it is shown that the above inequality is true for $\phi \in L^2(\mathbb{R}^n_x)$. The Arzelá-Ascoli theorem states that pointwise boundedness and equicontinuity suffice to show that ρ_ϵ is compact in $C([0,T], \mathrm{w} - L^2(dx))$. □

6.2.3 Passage to the Limit

Now that uniform convergence for ρ_ϵ has been established, we can proceed into deriving an equation for the remainder term $\tilde{g}_\epsilon$, i.e.

$$
\begin{aligned}
\epsilon \partial_t \tilde{g}_\epsilon - \nabla_x \cdot \int \mathcal{M} \nabla_v \tilde{g}_\epsilon \, dv + v \cdot (\nabla_x(\rho_\epsilon + \tilde{g}_\epsilon) + \nabla \mathcal{U}(x)(\rho_\epsilon + \tilde{g}_\epsilon)) \\
- \nabla \mathcal{U}(x) \cdot \nabla_v \tilde{g}_\epsilon = \frac{1}{\epsilon} \frac{1}{\mathcal{M}} \mathcal{C}(\mathcal{M} \tilde{g}_\epsilon).
\end{aligned} \tag{6.5}
$$

A weak solution of (6.5) will be in $C(\mathbb{R}_+, \mathcal{D}'(\mathbb{R}^{n,n}_{x,v}))$. For the sake of this proof we work with the choice of test functions $\varphi(x,v)$ s.t. $\varphi(x,v) \in C_c^\infty(\mathbb{R}^{n,n}_{x,v})$. The weak formulation is then given by

$$
\begin{aligned}
&\epsilon \iint \mathcal{M}(v)\varphi\ (\tilde{g}_\epsilon(t_2) - \tilde{g}_\epsilon(t_1))\ dv\, dx \\
&\qquad + \int_{t_1}^{t_2} \iint \mathcal{M}(v)\nabla_x\varphi \cdot \left(\int \mathcal{M}(v')\nabla_{v'}\tilde{g}_\epsilon\, dv' \right) dv\, dx\, ds \\
&\qquad\quad + \int_{t_1}^{t_2} \iint \mathcal{M}(v) v \cdot (-\nabla_x\varphi\, \rho_\epsilon + \varphi\nabla\mathcal{U}(x)\, \rho_\epsilon)\ dv\, dx\, ds \\
&\qquad\qquad + \int_{t_1}^{t_2} \iint \mathcal{M}(v)\, (-\nabla_x\varphi \cdot \nabla_v\tilde{g}_\epsilon + \varphi\nabla\mathcal{U}(x) \cdot \nabla_v\tilde{g}_\epsilon)\ dv\, dx\, ds \\
&- \int_{t_1}^{t_2} \iint \mathcal{M}(v)\varphi\nabla\mathcal{U}(x) \cdot \nabla_v\tilde{g}_\epsilon\, dv\, dx\, ds = -\frac{1}{\epsilon} \int_{t_1}^{t_2} \iint \mathcal{M}(v)\nabla_v\varphi \cdot G\nabla_v\tilde{g}_\epsilon\, dv\, dx\, ds.
\end{aligned}
$$

We use the notation I_j, for $1 \le j \le 6$ for the successive integrals that appear in the weak formulation. The study of the order of magnitude for each of them reveals that in the limit $\epsilon \to 0$ only terms $I_3\,\&\,I_6$ do not vanish. For now, we need to show the order of magnitude of each integral term and then consider the choice of test function that allows the coupling of (6.4) and the equation that we obtain in the limit $\epsilon \to 0$ in the weak formulation of (6.5).

Order of Magnitude for the I_j integral terms.

In what follows is the identification of the order of magnitude in ϵ of the I_j integrals.

Term I_1: The use of the Cauchy-Schwarz inequality, the a priori estimate and the fact that all Maxwellian moments are finite ($\int dv|v|^k\mathcal{M}(v) < \infty$) yields

$$
\begin{aligned}
I_1 &\le \epsilon \iint \mathcal{M}(v)\varphi\ (\tilde{g}_\epsilon(t_2) - \tilde{g}_\epsilon(t_1))\ dv\, dx \\
&\le \epsilon \left(\iint \varphi^2\mathcal{M}(v)\, dv\, dx \right)^{\frac{1}{2}} \left(\iint (|\tilde{g}_\epsilon(t_2)|^2 + |\tilde{g}_\epsilon(t_1)|^2)\mathcal{M}(v)\, dv\, dx \right)^{\frac{1}{2}} \\
&\le C\epsilon = O(\epsilon).
\end{aligned}
$$

Term I_2:

In similar fashion,

$$
\begin{aligned}
I_2 &= \int_{t_1}^{t_2} \iiint \mathcal{M}(v)\mathcal{M}(v')\nabla_x \varphi(x,v,s) \cdot \nabla_{v'} \tilde{g}_\epsilon(x,v',s)\, dv'\, dv\, dx\, ds \\
&\leq \epsilon \int_{t_1}^{t_2} \iiint \mathcal{M}(v)\mathcal{M}(v')|G^{-1/2}\nabla_x \varphi| \frac{|G^{1/2}\nabla_{v'}\tilde{g}_\epsilon|}{\epsilon}\, dv'\, dv\, dx\, ds \\
&\leq \epsilon \left(\int_{t_1}^{t^2} \iint \mathcal{M}(v)|G^{-1/2}\nabla_x \varphi|^2\, dv\, dx\, ds \right)^{1/2} \\
&\qquad\qquad \times \left(\frac{1}{\epsilon^2} \int_{t_1}^{t_2} \iint \mathcal{M}(v')|G^{1/2}\nabla_{v'}\tilde{g}_\epsilon|^2\, dv'\, dx\, ds \right)^{1/2} \\
&\qquad\qquad\qquad \leq C\epsilon = O(\epsilon).
\end{aligned}
$$

Term I_3:

Using the Cauchy-Schwartz once again ,

$$
\begin{aligned}
I_3 &\leq \int_{t_1}^{t_2} \iint \mathcal{M}(v) v \cdot \left(-\nabla_x \varphi\, \rho_\epsilon + \varphi \nabla \mathcal{U}(x)\, \rho_\epsilon\right)\, dv\, dx\, ds \\
&\leq \left(\int_{t_1}^{t_2} \iint \mathcal{M}(v)|v|^2 (|\nabla_x \varphi|^2 + |\varphi \nabla \mathcal{U}(x)|^2)\, dv\, dx\, ds \right)^{\frac{1}{2}} \left(\int_{t_1}^{t_2} \int \rho_\epsilon^2\, dx\, ds \right)^{\frac{1}{2}} \\
&\leq C \left(\iint \mathcal{M}(v)|v|^2 (|\nabla_x \varphi|^2 + |\varphi \nabla \mathcal{U}(x)|^2)\, dx \right)^{\frac{1}{2}} \left(\int \rho_\epsilon^2\, dx \right)^{\frac{1}{2}} = O(1).
\end{aligned}
$$

Term I_4: For this term

$$
\begin{aligned}
I_4 &\leq \int_{t_1}^{t_2} \iint \mathcal{M}(v) \left(-\nabla_x \varphi \cdot \nabla_v \tilde{g}_\epsilon + \varphi \nabla \mathcal{U}(x) \cdot \nabla_v \tilde{g}_\epsilon\right)\, dv dx ds \\
&\leq \epsilon \left(\int_{t_1}^{t_2} \iint \mathcal{M}(v)(|G^{-1/2}\nabla_x \varphi|^2 + |\varphi G^{-1/2} \nabla \mathcal{U}(x)|^2) dv dx ds \right)^{\frac{1}{2}} \\
&\qquad\qquad \times \left(\int_{t_1}^{t_2} \iint \frac{1}{\epsilon^2} |G^{1/2}\nabla_v \tilde{g}_\epsilon|^2 \mathcal{M}(v) dv dx ds \right)^{\frac{1}{2}} \\
&\qquad\qquad\qquad \leq C\epsilon = O(\epsilon).
\end{aligned}
$$

Term I_5: This term yields

$$
\begin{aligned}
I_5 &\leq -\int_{t_1}^{t_2} \iint \mathcal{M}(v)\varphi \nabla \mathcal{U}(x) \cdot \nabla_v \tilde{g}_\epsilon \, dv \, dx \, ds \\
&\leq \epsilon \left(\int_{t_1}^{t_2} \iint \mathcal{M}(v) |\varphi G^{-1/2} \nabla \mathcal{U}(x)|^2 \, dv \, dx \, ds \right)^{\frac{1}{2}} \\
&\qquad \times \left(\int_{t_1}^{t_2} \iint \frac{1}{\epsilon^2} |G^{1/2} \nabla_v \tilde{g}_\epsilon|^2 \mathcal{M}(v) \, dv \, dx \, ds \right)^{\frac{1}{2}} \\
&\leq C\epsilon = O(\epsilon).
\end{aligned}
$$

Term I_6: Finally for the I_6 term,

$$
\begin{aligned}
I_6 &\leq \frac{1}{\epsilon} \int_{t_1}^{t_2} \iint \mathcal{M}(v) \nabla_v \varphi \cdot G \nabla_v \tilde{g}_\epsilon \, dv \, dx \, ds \\
&\leq \left(\int_{t_1}^{t_2} \iint |G^{1/2} \nabla_v \varphi|^2 \mathcal{M}(v) \, dv \, dx \, ds \right)^{\frac{1}{2}} \\
&\qquad \times \left(\int_{t_1}^{t_2} \iint \frac{1}{\epsilon^2} |G^{1/2} \nabla_v \tilde{g}_\epsilon|^2 \mathcal{M}(v) \, dv \, dx \, ds \right)^{\frac{1}{2}} \\
&= O(1).
\end{aligned}
$$

Now we go back and consider the weak formulation of the (6.4), which is

$$
\int_{t_1}^{t_2} \int \phi \partial_t \rho \, dx \, ds = \int_{t_1}^{t_2} \iint \mathcal{M}(v) \nabla_x \phi \cdot G^{-1/2} J \, dv \, dx \, ds \qquad \text{for} \quad \phi \in C_c^\infty(\mathbb{R}_x^n). \tag{6.6}
$$

At the same time as the order analysis showed above, in the limit $\epsilon \to 0$, the (6.5) yields

$$
\int_{t_1}^{t_2} \iint \mathcal{M}(v) v \cdot (-\nabla_x \varphi \, \rho + \varphi \nabla \mathcal{U}(x) \, \rho) \, dv dx ds = -\int_{t_1}^{t_2} \iint \mathcal{M}(v) \nabla_v \varphi \cdot G^{1/2} J \, dv dx ds. \tag{6.7}
$$

These two equations are coupled for the choice of test function $\varphi(x,v) = \nabla_x \phi \cdot G^{-1} v$ where $\phi \in C_c^\infty(\mathbb{R}_x^n)$. The only problem is that this function is not smooth or compactly supported in $\mathbb{R}_v^n$ so we have to modify it slightly.

We begin by considering the cut-off function $\chi_{\delta_1}(v) = \chi(\delta_1 v)$, where $\chi(v) \in C_c^\infty(\mathbb{R}_v^n)$ is a function with values $0 \leq \chi(v) \leq 1$ s.t. $\chi(v) = 1 \, \text{for} \, |v| \leq 1$ and $\chi(v) = 0 \, \text{for} \, |v| \geq 1$. We also consider the standard mollification function,

$$\eta_{\delta_2}(v) = \frac{1}{\delta_2^n} \eta\left(\frac{v}{\delta_2}\right) \quad \text{for} \quad \eta \in C_c^\infty(\mathbb{R}_v^n) \text{ s.t. } \int \eta(v)\, dv = 1.$$

We now consider the function $\varphi_{\delta_1,\delta_2}(x,v) = (\chi_{\delta_1}(v) \nabla_x \phi \cdot G^{-1} v) \star \eta_{\delta_2}$. A standard result for the mollified function is that $\varphi_{\delta_1,\delta_2}$ converges to φ a.e. in $\mathbb{R}_v^n$ (as $\delta_1, \delta_2 \to 0$). Obviously $\nabla_x \varphi_{\delta_1,\delta_2}$ converges to $\nabla_x \varphi$ a.e. in $\mathbb{R}_v^n$, since the cut-off and mollification acts only in the v variable.

By substitution of φ with $\varphi_{\delta_1,\delta_2}$ in (6.7) one gets

$$\begin{aligned} \int_{t_1}^{t_2} \iint \mathcal{M}(v) v \cdot (-\nabla_x \varphi_{\delta_1,\delta_2}\, \rho + \varphi_{\delta_1,\delta_2} \nabla \mathcal{U}(x)\, \rho)\, dv\, dx\, ds = \\ - \int_{t_1}^{t_2} \iint \mathcal{M}(v) \nabla_v \varphi_{\delta_1,\delta_2} \cdot G^{1/2} J \, dv\, dx\, ds. \end{aligned}$$

We also have,

$$\begin{aligned} \nabla_v \varphi_{\delta_1,\delta_2}(x,v) &= \nabla_v \left((\nabla_x \phi \cdot G^{-1} v\, \chi_{\delta_1}(v)) \star \eta_{\delta_2} \right) \\ &= \nabla_v (\nabla_x \phi \cdot G^{-1} v\, \chi_{\delta_1}(v)) \star \eta_{\delta_2} \\ &= \left(\nabla_x \phi \cdot G^{-1} \chi_{\delta_1}(v) + \nabla_x \phi \cdot G^{-1} v\, \nabla_v \chi_{\delta_1}(v) \right) \star \eta_{\delta_2} \end{aligned}$$

where we made use of the fact that $\nabla_v (f \star \eta_\delta) = \nabla_v f \star \eta_\delta$.

A typical estimate for $\nabla_v \chi_{\delta_1}(v)$ is $|\nabla_v \chi_{\delta_1}(v)| \leq C\delta_1$. This can be easily seen by the definition of χ_{δ_1} and the fact that $|\nabla_v \chi| \leq C$ for some $C > 0$, since $\chi \in$

$C_c^\infty(\mathbb{R}_v^n)$. This estimate, together with the computation of $\nabla_v \varphi_{\delta_1,\delta_2}(x,v)$ above and the dominated convergence theorem imply that in the limit $\delta_1, \delta_2 \to 0$, one actually has that (6.7) holds with $\varphi(x,v) = \nabla_x \phi(x) \cdot G^{-1} v$. This choice of test function allows the coupling of (6.6) and (6.7) that yields

$$\int_{t_1}^{t_2} \int \phi \partial_t \rho \, dx \, ds = \int_{t_1}^{t_2} \int \left(\nabla_x \cdot (G^{-1} \nabla_x \phi) + \nabla_x \phi \cdot G^{-1} \nabla \mathcal{U}(x) \right) \rho \, dx \, ds$$

which completes the proof.

6.3 Diffusive Limit via Relative Entropy and Proof of Theorem 2

In this paragraph we are employing the relative entropy method in order to study the hydrodynamic limit. We have already used relative entropy for the study of the long time asymptotics of the equation. The technique goes back to the work by [74] for the Ginzburg-Landau model and S.Varadhan [68]. See for instance [27] (for Vlassov-Navier-Stokes equations) for a more elaborate instance of the method.

The relative entropy here measures the $LlogL$ distance between the distribution f_ϵ and the hydrodynamical equilibrium state $\rho\mathcal{M}(v)$. The idea is to study its rate of change and either establish that it has a sign in the leading order or that it satisfies a Gronwall type of inequality. Here, we are able to show that the entropy dissipates and that if initial data is prepared so that $\lim_{\epsilon\to 0} H(f_\epsilon(0)|\rho_0\mathcal{M}) = 0$ then $\lim_{\epsilon\to 0} H(f_\epsilon|\rho\mathcal{M}) = 0$ for all $t \in [0,T]$ for any $T > 0$.

Control of the relative entropy directly implies control of the L^1 norm $\|f_\epsilon - \rho\mathcal{M}\|$, by virtue of the Csiszár-Kullback-Pinsker inequality proven in the appendix.

Hydrodynamic variables play an important role in the calculations that follow so we re-introduce them here. The hydrodynamical density and flux are given by

$$\rho_\epsilon = \int f_\epsilon \, dv \quad \text{and} \quad J_\epsilon = \frac{1}{\epsilon}\int v f_\epsilon \, dv$$

respectively. The evolution of ρ_ϵ is governed by $\partial_t \rho_\epsilon = -\nabla_x \cdot J_\epsilon$, where the equation for the hydrodynamical flux was calculated to be $J_\epsilon(x,t) = -G^{-1}(\nabla_x\rho_\epsilon + \nabla\mathcal{U}(x)\rho_\epsilon) +$ lower order terms in ϵ. The equation for the formal limit ρ is $\partial_t\rho = \nabla\cdot(G^{-1}(\nabla_x\rho + \nabla\mathcal{U}(x)\rho))$.

We can start with an easy computation on the evolution of the $H(\rho_\epsilon|\rho)$ relative entropy. This computation becomes partly obsolete later when we show that the time derivative of the entropy $H(f_\epsilon|\rho\mathcal{M})$ satisfies a similar estimate. In fact, it is manifested by the following lemma that $H(\rho_\epsilon|\rho) \leq H(f_\epsilon|\rho\mathcal{M})$, since

Lemma 7. *Consider the hydrodynamic variables ρ_1, ρ_2 associated with the density functions f_1, f_2. Then, we have*

$$H(\rho_1|\rho_2) \leq H(f_1|f_2).$$

Proof. The proof is given in appendix. □

Nevertheless, we present the computation here, because it is a good starting point and it also contain parts helpful for the computation of $\frac{d}{dt}H(f_\epsilon|\rho\mathcal{M})$.

$$\begin{aligned}
&\frac{d}{dt}H(\rho_\epsilon|\rho) = \frac{d}{dt}\int \rho_\epsilon \log\frac{\rho_\epsilon}{\rho}\,dx = \frac{d}{dt}\int \rho_\epsilon \log \rho_\epsilon\,dx - \frac{d}{dt}\int \rho_\epsilon \log\rho\,dx \\
&= \int \partial_t\rho_\epsilon(\log\rho_\epsilon + 1)\,dx - \int \partial_t\rho_\epsilon \log\rho\,dx - \int \frac{\rho_\epsilon}{\rho}\partial_t\rho\,dx \\
&= \int J_\epsilon \cdot \frac{\nabla\rho_\epsilon}{\rho_\epsilon}\,dx - \int J_\epsilon\cdot\frac{\nabla\rho}{\rho}\,dx - \int \frac{\rho_\epsilon}{\rho}\nabla\cdot\left(G^{-1}(\nabla\rho + \nabla\mathcal{U}(x)\rho)\right)\,dx \\
&= \int J_\epsilon\cdot\left(\frac{\nabla\rho_\epsilon}{\rho_\epsilon} - \frac{\nabla\rho}{\rho}\right)dx + \int \rho\nabla\left(\frac{\rho_\epsilon}{\rho}\right)\cdot G^{-1}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x)\right)dx \\
&= \int J_\epsilon\cdot\left(\frac{\nabla\rho_\epsilon}{\rho_\epsilon} - \frac{\nabla\rho}{\rho}\right)dx + \int \left(\frac{\nabla\rho_\epsilon}{\rho_\epsilon} - \frac{\nabla\rho}{\rho}\right)\cdot G^{-1}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x)\right)\rho_\epsilon\,dx \\
&= \int \left(\frac{\nabla\rho_\epsilon}{\rho_\epsilon} - \frac{\nabla\rho}{\rho}\right)\cdot\left(\frac{J_\epsilon}{\rho_\epsilon} + G^{-1}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x)\right)\right)\rho_\epsilon\,dx \\
&= -\int G\left(\frac{J_\epsilon}{\rho_\epsilon} + G^{-1}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x)\right)\right)\cdot\left(\frac{J_\epsilon}{\rho_\epsilon} + G^{-1}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x)\right)\right)\rho_\epsilon\,dx \\
&+ r'_\epsilon = -\int \left|G^{1/2}\frac{J_\epsilon}{\rho_\epsilon} + G^{-1/2}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x)\right)\right|^2\rho_\epsilon\,dx + r'_\epsilon.
\end{aligned}$$

The remainder term r'_ϵ is found to be exactly equal to

$$-\int\left(\epsilon^2\partial_t J_\epsilon + \nabla_x\cdot\int \mathcal{M}\nabla_v\left(\frac{f_\epsilon}{\mathcal{M}}\right)\otimes v\,dv\right)\cdot\left(\frac{J_\epsilon}{\rho_\epsilon} + G^{-1}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x)\right)\right)dx.$$

Notice that r'_ϵ is a remainder term that should vanish as $\epsilon \to 0$. We don't bother with showing that $r'_\epsilon \to 0$ in rigorous manner, as we mainly work with the relative entropy $H(f_\epsilon|\rho\mathcal{M})$.

Yet, as we are about to remark promptly after the proof, the above computation of $\frac{d}{dt}H(\rho_\epsilon|\rho)$ alone can be used to establish the exact same result that we are about to prove!

At this point we compute the evolution of $H(f_\epsilon|\rho\mathcal{M})$ in similar manner. To make things easier we can introduce the global equilibrium state $\mathcal{M}_{eq}(x,v)$ in the computation that follows.

$$\begin{aligned} H(f_\epsilon|\rho\mathcal{M}) &= \iint f_\epsilon \log f_\epsilon \, dv\, dx - \iint f_\epsilon \log(\rho\mathcal{M})\, dv\, dx \\ &= \iint f_\epsilon \log \frac{f_\epsilon}{\mathcal{M}_{eq}}\, dv\, dx + \iint f_\epsilon \log \frac{\mathcal{M}_{eq}}{\rho\mathcal{M}}\, dv\, dx \\ &= \iint f_\epsilon \log \frac{f_\epsilon}{\mathcal{M}_{eq}}\, dv\, dx + \int \rho_\epsilon \log \frac{e^{-\mathcal{U}(x)}}{\rho}\, dx \\ &= H(f_\epsilon|\mathcal{M}_{eq}) - \int \rho_\epsilon \log \rho\, dx - \int \rho_\epsilon \mathcal{U}(x)\, dx. \end{aligned}$$

The reason we have introduced the global equilibrium state and $H(f_\epsilon|\mathcal{M}_{eq})$ is that the term involving $\frac{d}{dt}H(f_\epsilon|\mathcal{M}_{eq})$ can be easily bounded by an integral involving only hydrodynamical variables. Indeed, the time derivative of the first term $H(f_\epsilon|\mathcal{M}_{eq})$ is

$$\begin{aligned} \frac{d}{dt}\iint f_\epsilon \log \frac{f_\epsilon}{\mathcal{M}_{eq}}\, dv\, dx &= -\frac{1}{\epsilon^2}\iint f_\epsilon \Big| G^{1/2}\nabla_v \log \frac{f_\epsilon}{\mathcal{M}}\Big|^2 dv\, dx \\ &= -\frac{1}{\epsilon^2}\iint f_\epsilon \Big| G^{1/2}\left(\frac{\nabla_v f_\epsilon}{f_\epsilon} + v\right)\Big|^2 dv\, dx \leq -\int \frac{|G^{1/2}J_\epsilon|^2}{\rho_\epsilon}\, dx. \end{aligned}$$

The last inequality is in fact due to the Hölder's inequality

$$\int \frac{|G^{1/2}J_\epsilon|^2}{\rho_\epsilon}\,dx = \frac{1}{\epsilon^2}\int \frac{\left(\int G^{1/2}\left(v+\frac{\nabla_v f_\epsilon}{f_\epsilon}\right) f_\epsilon\,dv\right)^2}{\rho_\epsilon}\,dx$$
$$\le \frac{1}{\epsilon^2}\iint \left|G^{1/2}\left(\frac{\nabla_v f_\epsilon}{f_\epsilon}+v\right)\right|^2 f_\epsilon\,dv\,dx.$$

We are now in position to compute the time evolution of the $H(f_\epsilon|\rho\mathcal{M})$ entropy by means of computing the sum of its terms. Indeed,

$$\frac{d}{dt}H(f_\epsilon|\rho\mathcal{M}) = \frac{d}{dt}H(f_\epsilon|\mathcal{M}_{eq}) - \frac{d}{dt}\int \rho_\epsilon \log\rho\,dx - \frac{d}{dt}\int \rho_\epsilon \mathcal{U}(x)\,dx$$
$$\le -\int \frac{|G^{1/2}J_\epsilon|^2}{\rho_\epsilon}\,dx - \frac{d}{dt}\int \rho_\epsilon\log\rho\,dx - \int J_\epsilon\cdot\nabla\mathcal{U}(x)\,dx.$$

The computation of $\frac{d}{dt}\int\rho_\epsilon\log\rho\,dx$ has been performed as a part of the computation of $\frac{d}{dt}H(\rho_\epsilon|\rho)$ above. We thus get,

$$\frac{d}{dt}H(f_\epsilon|\rho\mathcal{M}) \le -\int \frac{|G^{1/2}J_\epsilon|^2}{\rho_\epsilon}\,dx - \int J_\epsilon\cdot\nabla\mathcal{U}(x)\,dx - \int J_\epsilon\cdot\frac{\nabla\rho}{\rho}\,dx$$
$$+\int\left(\frac{\nabla\rho_\epsilon}{\rho_\epsilon}-\frac{\nabla\rho}{\rho}\right)\cdot G^{-1}\left(\frac{\nabla\rho}{\rho}+\nabla\mathcal{U}(x)\right)\rho_\epsilon\,dx$$
$$= -\int J_\epsilon\cdot\left(G\frac{J_\epsilon}{\rho_\epsilon}+\frac{\nabla\rho}{\rho}+\nabla\mathcal{U}(x)\right)dx$$
$$-\int\left(G\frac{J_\epsilon}{\rho_\epsilon}+\frac{\nabla\rho}{\rho}+\nabla\mathcal{U}(x)\right)\cdot G^{-1}\left(\frac{\nabla\rho}{\rho}+\nabla\mathcal{U}(x)\right)\rho_\epsilon\,dx + r_\epsilon$$
$$= -\int\left(G\frac{J_\epsilon}{\rho_\epsilon}+\frac{\nabla\rho}{\rho}+\nabla\mathcal{U}(x)\right)\cdot\left(\frac{J_\epsilon}{\rho_\epsilon}+G^{-1}\left(\frac{\nabla\rho}{\rho}+\nabla\mathcal{U}(x)\right)\right)\rho_\epsilon\,dx + r_\epsilon$$
$$= -\int\left|G^{1/2}\frac{J_\epsilon}{\rho_\epsilon}+G^{-1/2}\left(\frac{\nabla\rho}{\rho}+\nabla\mathcal{U}(x)\right)\right|^2\rho_\epsilon\,dx + r_\epsilon. \tag{6.8}$$

Now that we have computed the dissipation of the relative entropy, we can proceed to the proof of Theorem 2.

Proof. We have already given the formal computation for $\frac{d}{dt}H(f_\epsilon|\rho\mathcal{M})$. What is left is to find the exact expression for the remainder term r_ϵ and show that it is indeed of lower order using the a priori estimate.

Let us begin with the observation that the quantity

$$d_\epsilon = G^{1/2}(x)(v\sqrt{f_\epsilon} + 2\nabla_v\sqrt{f_\epsilon}) = 2\sqrt{\mathcal{M}}G^{1/2}(x)\nabla_v\sqrt{\frac{f_\epsilon}{\mathcal{M}}}$$

is of order $O(\epsilon)$ in L^2, as implied by the boundedness of the total energy, i.e.

$$\int_0^T \iint |d_\epsilon|^2 \, dv \, dx \, dt \le C\epsilon^2, \qquad T > 0.$$

Indeed, consider the free energy associated with the FP equation

$$\mathcal{E}(f_\epsilon) = \iint f_\epsilon \left(\ln f_\epsilon + \frac{|v|^2}{2} + \mathcal{U}(x) \right) dv \, dx.$$

This quantity is dissipated since

$$\frac{d}{dt}\mathcal{E}(f_\epsilon) = -\frac{1}{\epsilon^2} \iint |d_\epsilon|^2 \, dv \, dx,$$

which implies

$$\mathcal{E}(f_\epsilon(T,\cdot,\cdot)) + \frac{1}{\epsilon^2}\int_0^T \iint |d_\epsilon|^2 \, dv \, dx \, ds = \mathcal{E}(f_\epsilon(0,\cdot,\cdot)).$$

We now have to compute the remainder term r_ϵ that appears in (6.8). By (6.3), it is implied that

$$\begin{aligned} \frac{\nabla\rho_\epsilon}{\rho_\epsilon} - \frac{\nabla\rho}{\rho} = -G\frac{J_\epsilon}{\rho_\epsilon} - \frac{\nabla\rho}{\rho} &- \nabla\mathcal{U}(x) - \epsilon^2\frac{\partial_t J_\epsilon}{\rho_\epsilon} \\ &- \frac{1}{\rho_\epsilon}\nabla_x \cdot \int \mathcal{M}\nabla_v\left(\frac{f_\epsilon}{\mathcal{M}}\right) \otimes v \, dv. \end{aligned}$$

Hence, by direct substitution in (6.8) the remainder term is

$$r_\epsilon = -\int \left(\epsilon^2 \partial_t J_\epsilon + \nabla_x \cdot \left(\int \mathcal{M}\nabla_v\left(\frac{f_\epsilon}{\mathcal{M}}\right) \otimes v \, dv \right) \right) \cdot G^{-1}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x) \right) dx.$$

This consists of two parts $r_{1,\epsilon}$ and $r_{2,\epsilon}$, which integrated in time are

$$\int_0^T r_{1,\epsilon} \, dt = -\epsilon \int_0^T \iint \partial_t f_\epsilon \, v \cdot G^{-1}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x) \right) dv \, dx \, dt,$$

$$\int_0^T r_{2,\epsilon}\,dt = \int_0^T \iint \left(\mathcal{M}\,v \otimes \nabla_v \left(\frac{f_\epsilon}{\mathcal{M}}\right)\right) : \nabla \left(G^{-1}\left(\frac{\nabla\rho}{\rho} + \nabla\mathcal{U}(x)\right)\right)\,dv\,dx\,dt.$$

It will be our task to show that both integrals vanish as $\epsilon \to 0$. An important step towards this is showing that $|v|^2 f_\epsilon \in L^\infty((0,T), L^1(\mathbb{R}^n \times \mathbb{R}^n))$ for any $T > 0$.

• **Bound for** $\iint |v|^2 f_\epsilon\,dv\,dx$ **in** $L^\infty(0,T)$.

The bound on the kinetic energy (uniform in time) is a straightforward consequence of the elementary, yet general, Young's inequality

$$ab \le h(a) + h^*(b),$$

where h,h^* are a Young's convex pair (h^* is explicitly computed by the Legendre transform of the convex function h). Here we use $h(z) = z\log z$, and $h^*(z) = e^{z-1}$, i.e.

$$\frac{1}{2}\iint |v|^2 f_\epsilon\,dv\,dx \le \iint f_\epsilon \log\frac{f_\epsilon}{\mathcal{M}_{eq}}\,dv\,dx + \iint e^{-\frac{|v|^2}{2}-1}\mathcal{M}_{eq}\,dv\,dx.$$

This implies that

$$\iint \frac{|v|^2}{2} f_\epsilon(t,v,x)\,dv\,dx \le C \qquad \text{for some } C>0,\ \forall t \in [0,T],$$

since $e^{-\mathcal{U}(x)} \in L^1(\mathbb{R}^n)$ and the entropy integral is bounded by the a priori estimate.

It is now time to control the residual terms. Let tensor D and vectors E,F be shorts for $D = \nabla(G^{-1}(\nabla\log\rho + \nabla\mathcal{U}(x)))$, $E = G^{-1}(\nabla\log\rho + \nabla\mathcal{U}(x))$, and $F = G^{-1}\nabla\partial_t\log\rho$. The easiest term to control is

$$\begin{aligned}
\int_0^T r_{2,\epsilon}\,dt &= \int_0^T \iint \mathcal{M}v \otimes \nabla_v\left(\frac{f_\epsilon}{\mathcal{M}}\right) : D\,dv\,dx\,dt \\
&= \int_0^T \iint \sqrt{f_\epsilon}v \otimes \gamma(x)^{-\frac{1}{2}} d_\epsilon : D\,dv\,dx\,dt \\
&\le C\epsilon \left(\int_0^T \iint \lambda_0 \frac{|d_\epsilon|^2}{\epsilon^2}\,dv\,dx\,dt\right)^{\frac{1}{2}} \left(\int_0^T \iint \|D\|_\infty |v|^2 f_\epsilon\,dv\,dx\,dt\right)^{\frac{1}{2}}.
\end{aligned}$$

Finally, the term

$$\int_0^T r_{1,\epsilon}\,dt = -\,\epsilon \iint (f_\epsilon(T,v,x) - f_\epsilon(0,v,x))\,v\cdot E\,dv\,dx$$
$$+\,\epsilon\int_0^T \iint f_\epsilon\, v\cdot F\,dv\,dx\,dt$$

is treated with the use of the Cauchy-Schwartz inequality which concludes the proof.

□

Remark 14. *Let us mention here why showing L^1 convergence of $\rho_\epsilon(t,x)$ to the limiting distribution $\rho(x)$ is enough to imply that f_ϵ converges to $\rho\mathcal{M}$ (in L^1) by the following simple argument. Indeed, we decompose $f_\epsilon - \rho\mathcal{M}$ as in*

$$f_\epsilon - \rho\mathcal{M} = f_\epsilon - \rho_\epsilon\mathcal{M} + (\rho_\epsilon - \rho)\mathcal{M}.$$

It is trivial showing that the second term $(\rho_\epsilon - \rho)\mathcal{M}$ of the decomposition $\to 0$ in L^1, by assumption. For the first term $f_\epsilon - \rho_\epsilon\mathcal{M}$, we use the Csiszár-Kullback-Pinsker and log-Sobolev inequalities in that order, and finally the a priori energy bound to get

$$\|f_\epsilon - \rho_\epsilon\mathcal{M}\|_{L^1} \le \sqrt{2}\left(\iint f_\epsilon \log\frac{f_\epsilon}{\rho_\epsilon\mathcal{M}}\,dv\,dx\right)^{1/2} \le \sqrt{2}\epsilon\left(\iint \frac{|G^{-\frac{1}{2}}d_\epsilon|^2}{\epsilon^2}\,dv\,dx\right)^{1/2}$$
$$\le \sqrt{2}\epsilon C \to 0 \qquad as \quad \epsilon\to 0,$$

which concludes the argument.

Chapter A: Appendix

This appendix is devoted to some individual topics that are standard in various literature. We made an effort to include them here for the sake of a full presentation.

We begin by reviewing the energy dissipation for the Stokes system that includes the particles and the medium. The energy or "variational" formulation is in fact another way to present the N particle Stokes problem. The reason this formulation was important in our study was because it allowed the derivation of a non-negative hydrodynamic and friction tensor.

Next, the Stokes flow solution for a single particle is presented both for the homogeneous and non homogeneous Dirichlet boundary conditions. The understanding of the Stokes' solution operator to a single particle problem coupled with the method of reflections sheds light to the behavior of the N particle system in the so called dilute regime.

As we begin with the analysis of the FP operator L, we are overwhelmed by the amount of computations on derivations and especially those on commutators. All these calculations are squeezed here in the paragraph with the name 'Commutator Algebra'. These calculations have not been included in the main text for economy of space.

Finally, there is a paragraph devoted to sufficient conditions for the Poincaré and log-Sobolev inequalities. These conditions are important in giving decay rates to Kramers-Smoluchowski type of equations and they are employed in many cases when the question of convergence to a global equilibrium state for the FP Cauchy problem is involved.

A.1 Energetics of Particle System

In order to understand the variational formulation that gives the energy dissipation used for the RPY approximation, we need to study the energy of the particle system. The physical energy $\mathcal{E}(t)$ at time t is the sum

$$\mathcal{E}(t) = \mathcal{E}_{par}(t) + \mathcal{E}_{fl}(t)$$

of the particles' energy

$$\mathcal{E}_{par}(t) = \sum_{i=1}^{N} \frac{1}{2} m|v_i|^2 + \mathcal{U}(x_1, x_2, \ldots, x_N),$$

and the energy of the medium

$$\mathcal{E}_{fl}(t) = \frac{1}{2} \int_{\mathbb{R}\setminus\cup_i B_i} |u(x)|^2 \, dx,$$

when one considers a fluid with unit density ($\rho = 1$).

The rate of change in the above terms is

$$\frac{d}{dt}\mathcal{E}_{par}(t) = \sum_{i=1}^{N} \left(m v_i \cdot \frac{dv_i}{dt} + \nabla_{x_i} \mathcal{U}(x_1, x_2, \ldots, x_N) \cdot v_i \right).$$

Using the balance of forces, the last term yields

$$\frac{d}{dt}\mathcal{E}_{par}(t) = \sum_{i=1}^{N} \int_{S_i} v_i^T \sigma \cdot n_i \, dS = \sum_{i=1}^{N} \int_{S_i} u(x)^T \sigma \cdot n_i \, dS,$$

for n_i being the outward normal on S_i.

The rate of energy change for the fluid motion is

$$\frac{d}{dt}\mathcal{E}_{fl}(t) = \int_{\mathbb{R}^3\setminus\cup_i B_i} u(x) \cdot \frac{Du(x)}{Dt} \, dx,$$

as follows by the transport theorem. The notation $\frac{D}{Dt}$ stands for the material derivative involving a time derivative plus a transport term

$$\frac{D}{Dt} = \frac{d}{dt} + u(x) \cdot \nabla \quad .$$

Since $\frac{Du}{Dt} = div\sigma$ by the equation of motion, we get

$$\begin{aligned}
\frac{d}{dt}\mathcal{E}_{fl}(t) &= \int_{\mathbb{R}^3 \setminus \cup_i B_i} u(x)^T div\sigma(x)\, dx \\
&= \int_{\mathbb{R}^3 \setminus \cup_i B_i} div(u^T \sigma)\, dx - \int_{\mathbb{R}^3 \setminus \cup_i B_i} \sigma : \nabla u\, dx \\
&= -\int_{S_i} u^T \sigma \cdot n_i\, dS - \sum_{i,j} \int \sigma_{ij} \frac{\partial u_i}{\partial x_j}\, dx,
\end{aligned}$$

where indices in the second integral designate vector and tensor components.

The expression in the last integral, for an incompressible fluid, yields

$$\sum_{i,j} \sigma_{ij} \frac{\partial u_i}{\partial x_j} = \frac{1}{2} \sum_{i,j} \sigma_{ij} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \frac{1}{2} \eta_s \sum_{i,j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)^2,$$

where we have made use of the symmetry of σ.

As a result, the rate of change for the total energy of the system is

$$\frac{d}{dt}\mathcal{E}(t) = -\frac{1}{2}\eta_s \sum_{i,j} \int_{\mathbb{R}^3 \setminus \cup_k B_k} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)^2 dx \leq 0.$$

A.2 Stokes Flow Past a Sphere

We have presented a methodology for solving the Stokes N particle problem in a complex domain D, by virtue of an iterative scheme. In order to proceed with the technique, we need to solve the particle problem for a single sphere.

For now, we present the solution of Stokes' problem for a sphere of radius R, with center located at the origin $x = 0$. The velocity on the surface of the sphere is $u(x) = U + \Omega \wedge x$ ($|x| = R$) and vanishes at infinity. Observe that there is a translational U, and angular Ω component to the velocity. The choice of frame of reference with a static sphere simplifies the problem. The resulting solution (outer solution) can be found in any standard textbook in fluid mechanics e.g. [48,56] etc., and is also presented in [36,37] i.e.

$$u(x) = \frac{3}{4}\frac{R}{|x|}\left(I + \frac{x \otimes x}{|x|^2}\right)U + \frac{1}{4}\frac{R^3}{|x|^3}\left(I - 3\frac{x \otimes x}{|x|^2}\right)U + \frac{R^3}{|x|^3}\Omega \wedge x,$$

$$p(x) = \frac{3}{2}R\eta\frac{U \cdot x}{|x|^3} \qquad \text{for} \quad |x| \geq R.$$

Computing the fluid forces on the surface of the particle, we end up with

$$F_f = 6\pi R\,\eta\,U,$$

$$T_f = 8\pi R^2 \eta\,\Omega.$$

One notices that translational and rotational components are fully decoupled. Also, in the leading order the solution is

$$u(x) \approx \frac{3}{4}\frac{R}{|x|}\left(I + \frac{x \otimes x}{|x|^2}\right)U,$$

which is exactly the Stokeslet approximation for $R = 1$.

Although this solution does not cover the case of an inhomogeneous BC on the surface of the sphere, it is highly indicative that the velocity field at a point in the medium is inversely proportional to its distance from the sphere. For a solution to the inhomogeneous BCs problem one can see e.g. [36, 56].

A.3 Fundamental Solutions

In this part of the appendix we give the fundamental solution to the equation

$$\partial_t f + v \cdot \nabla_x f = \gamma \Delta_v f \quad \text{for} \quad x, v \in \mathbb{R}^n,\ \gamma > 0,$$

and initial data $f(0,x,v) = \delta(x_0, v_0)$. We present here two solutions. The first one is "analytic" in nature, and the second is "algebraic" in the sense that is takes advantage of the Lie algebra structure. We should mention that many special solutions to the Fokker-Planck equation can be found in [64].

Starting with the analytic solution, we consider the transformed equation

$$\partial_t \hat{f} - \xi \cdot \nabla_\eta \hat{f} + \gamma |\eta|^2 \hat{f} = 0,$$

for the Fourier transform $\hat{f}(t,\xi,\eta) = \iint e^{-i(x\cdot\xi + v\cdot\eta)} f(t,x,v)\, dv\, dx$. The initial condition is transformed to

$$\hat{f}(0,\xi_0,\eta_0) = e^{-i(x_0\cdot\xi_0 + v_0\cdot\eta_0)}.$$

The solution to the characteristic system of the transformed equation is $\xi(t) = \xi_0$, and $\eta(t) = -\xi_0 t + \eta_0$, and hence the solution of the transformed equation is

$$\begin{aligned}
\hat{f}(t,\xi,\eta) &= \hat{f}(0,\xi_0,\eta_0) e^{-\gamma \int_0^t |\eta(t-s)|^2\, ds} \\
&= e^{-i(x_0\cdot\xi_0 + v_0\cdot\eta_0)} e^{-\gamma \int_0^t |\eta(t-s)|^2\, ds} \\
&= e^{-i(x_0\cdot\xi + v_0\cdot(\eta+\xi t))} e^{-\gamma\left(\frac{\xi^2}{3} t^3 + \xi\cdot\eta t^2 + \eta^2 t\right)}.
\end{aligned}$$

Taking the inverse transform, one gets

$$\begin{aligned} f(t,x,v) &= \frac{1}{(2\pi)^{2n}} \iint e^{i(x\cdot\xi+v\cdot\eta)} \hat{f}(t,\xi,\eta)\, d\eta\, d\xi \\ &= \frac{1}{(2\pi)^{2n}} \iint e^{i((x-x_0-v_0t)\cdot\xi+(v-v_0)\cdot\eta)} e^{-\gamma\left(\frac{\xi^2}{3}t^3+\xi\cdot\eta t^2+\eta^2 t\right)}\, d\eta\, d\xi. \end{aligned}$$

Last integration when performed yields

$$\begin{aligned} f(t,x,v) = \frac{1}{(3\pi^2\gamma)^n}\frac{1}{t^{2n}} \exp\Big[-\frac{1}{\pi^2\gamma}\Big(3\frac{|x-x_0-v_0t|^2}{t^3} \\ - 3\frac{(x-x_0-v_0t)\cdot(v-v_0)}{t^2} + \frac{|v-v_0|^2}{t}\Big)\Big]. \end{aligned}$$

The second method was suggested to me by C.D.Levermore and makes use of the Baker-Campbell-Hausdorff formula for elements of a Lie algebra. We write the equation as $\partial_t f = Lf$ for $L = \gamma A - D$, with $A = \Delta_v$ and $D = v\cdot\nabla_x$.

Here $[A,D] = 2\nabla_v\cdot\nabla_x = 2B$ and $[B,D] = \Delta_x$. At the same time we have $[A,B] = [A,C] = [B,C] = [C,D] = 0$. So as a result, $A = \Delta_v$, $B = \nabla_v\cdot\nabla_x$, $C = \Delta_x$, and $D = v\cdot\nabla_x$ form a Lie algebra. The idea is to write e^{tL} as $e^{a(t)A}e^{b(t)B}e^{c(t)C}e^{d(t)D}$, for the differentiable functions $a(t), b(t), c(t), d(t)$, with $a(0) = b(0) = c(0) = d(0) = 0$. Those functions are to be computed explicitly.

We need to compute commutations of the fields A, B, C, D with $e^{a(t)A}$ etc. For that, we first have to find $[A^n, D]$, $[B^n, D]$, ... We begin with,

$$
\begin{aligned}
[A^n, D] &= A^n D - DA^n = A^n D - (DA)A^{n-1} = A^n D - (AD - [A, D])A^{n-1} \\
&= A^n D - ADA^{n-1} + 2BA^{n-1} = A^n D - A(DA)A^{n-2} + 2BA^{n-1} \\
&= A^n D - A(AD - [A, D])A^{n-2} + 2BA^{n-1} = A^n D - A^2 DA^{n-2} + 4BA^{n-1} \\
&= (n - 2 \text{ commutations of } A, D) \ldots = A^n D - A^n D + 2nBA^{n-1} \\
&= 2nBA^{n-1}.
\end{aligned}
$$

In the same manner, it is shown $[B^n, D] = nCB^{n-1}$ with the rest similar terms being 0 e.g. $[C^n, D] = 0$ etc. This implies

$$
\begin{aligned}
[e^{a(t)A}, D] &= [\sum_{n=0}^{\infty} \frac{(a(t)A)^n}{n!}, D] = \sum_{n=0}^{\infty} \frac{a(t)^n}{n!} [A^n, D] = \sum_{n=0}^{\infty} \frac{a(t)^n}{n!} 2nBA^{n-1} \\
&= 2a(t)B \sum_{n=1}^{\infty} \frac{a(t)^{n-1}}{(n-1)!} A^{n-1} = 2a(t)Be^{a(t)A}.
\end{aligned}
$$

We also calculate $[e^{b(t)B}, D] = b(t)Ce^{b(t)B}$, and $[e^{a(t)A}, B] = [e^{a(t)A}, C] = [e^{b(t)B}, C] = \ldots = 0$.

Differentiating e^{tL} in time yields

$$
\begin{aligned}
Le^{tL} &= a'(t)Ae^{a(t)A}e^{b(t)B}e^{c(t)C}e^{d(t)D} + e^{a(t)A}b'(t)Be^{b(t)B}e^{c(t)C}e^{d(t)D} \\
&+ e^{a(t)A}e^{b(t)B}c'(t)Ce^{c(t)C}e^{d(t)D} + e^{a(t)A}e^{b(t)B}e^{c(t)C}d'(t)De^{d(t)D} \\
&= (a'(t)A + b'(t)B + c'(t)C)e^{tL} + d'(t)e^{a(t)A}e^{b(t)B}De^{c(t)C}e^{d(t)D} \\
&= (a'(t)A + b'(t)B + c'(t)C)e^{tL} + d'(t)e^{a(t)A}(b(t)C + D)e^{b(t)B}e^{c(t)C}e^{d(t)D} \\
&= (a'(t)A + b'(t)B + c'(t)C + d'(t)b(t)C)e^{tL} + 2d'(t)a(t)Be^{tL} + d'(t)De^{tL} \\
&= (a'(t)A + (b'(t) + 2d'(t)a(t))B + (c'(t) + d'(t)b(t))C + d'(t)D)e^{tL}.
\end{aligned}
$$

This implies the system of equations:

$$a'(t) = \gamma, \quad b'(t) + 2d'(t)a(t) = 0, \quad c'(t) + d'(t)b(t) = 0, \quad d'(t) = -1,$$

with solution $a(t) = \gamma t, \quad b(t) = \gamma t^2, \quad c(t) = \gamma \frac{t^3}{3}, \quad d(t) = -t.$

Hence, the semi-group can be written as $e^{tL} = e^{\gamma\left(tA+t^2B+\frac{t^3}{3}C\right)}e^{-tD}$, which gives the exact solution $f(t,x,v)$ that we presented with the method above.

Remark 15. *The equation that we solved above is not the toughest example with a possible exact solution that we can compute using the Lie algebra structure. In fact, equation*

$$\partial_t f + v \cdot \nabla_x f = \gamma(\Delta_v f + v \cdot \nabla_v f + nf),$$

which now has a unique non-zero equilibrium state can be solved with same tricks. We write the equation in form $\partial_t f = Lf$ for $L = \gamma(A+E) - D$, with $E = v \cdot \nabla_v + n$ and the rest of vector fields A, B, C, D same as before. The new commutations introduced here are the ones with E, which are

$$[A,E] = 2A, \quad [B,E] = B, \quad [C,E] = 0, \quad [D,E] = -D.$$

Like above, one computes with help of induction

$$[A^n,E] = 2nA^n, \quad [B^n,E] = nB^n, \quad [C^n,E] = 0, \quad [D^n,E] = -nD^n,$$

which in return yields the relations

$$[e^{aA},E] = 2aAe^{aA}, \quad [e^{bB},E] = bBe^{bB}, \quad [e^{cC},E] = 0, \quad [e^{dD},E] = -dDe^{dD}.$$

With the above computations at hand and using the same idea of expanding e^{tL} as $e^{aA}e^{bB}e^{cC}e^{dD}e^{eE}$ for functions $a(t), \ldots, e(t)$ with $a(0) = \ldots = e(0) = 0$, after *differentiating in time and commuting fields we end up with the system of equations*

$$
\begin{aligned}
2a(t)e'(t) + a'(t) &= \gamma \\
e'(t)(b(t) - 2a(t)d(t)) + b'(t) + 2a(t)d'(t) &= 0 \\
c'(t) + d'(t)b(t) - b(t)d(t)e'(t) &= 0 \\
d'(t) - d(t)e'(t) &= -1 \\
e'(t) &= \gamma.
\end{aligned}
$$

The above system has solution $a(t) = \frac{1}{2}(1 - e^{-2\gamma t})$, $\quad b(t) = \frac{1}{\gamma} - \frac{2}{\gamma}e^{-\gamma t} + \frac{1}{\gamma}e^{-2\gamma t}$, $c(t) = -\frac{3}{2\gamma^2} + \frac{t}{\gamma} - \frac{1}{2\gamma^2}e^{-2\gamma t} + \frac{2}{\gamma^2}e^{-\gamma t}$, $\quad d(t) = \frac{1}{\gamma}(1 - e^{\gamma t})$ *and* $e(t) = \gamma t$.

A.4 Strong Solutions with Regular Coefficients

The link between an SDE and its corresponding Fokker-Planck (forward Kolmogorov) equation has been discussed massively in literature and it can be viewed as the evolution to the idea of the link between a transport equation and its characteristic ODE. Here, we review the existence of strong solutions in its connection to the corresponding SDE problem.

Let us consider the SDE for $X_t \in \mathbb{R}^n$,

$$dX_t = b(X_t)\,dt + \sqrt{2}\sigma(X_t)\,dW_t$$

with initial condition $X_0 \in \mathbb{R}^n$. The drift $b(x)$ is a vector in $\mathbb{R}^n$ and the dispersion matrix $\sigma(x) \in \mathbb{R}^{n\times n}$. Also W_t is the standard Brownian vector.

The definition that holds for a strong solution of the SDE is

Definition 3. *A strong solution to the SDE is a solution X_t that exists for a given probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, the given Brownian motion W_t, and initial data X_0. Furthermore, uniqueness of a strong solution is meant to be understood in the pathwise sense.*

The following theorem is a well known result (see [39]) and has its analog in the Lipschitz-Cauchy theory for ODEs. The theorem, as presented in [50], states

Theorem 23. *Assume that the drift and dispersion satisfy the following growth conditions,*

$$\frac{b(x)}{1+|x|} \in (L^\infty(\mathbb{R}^n))^n, \qquad \frac{\sigma(x)}{1+|x|} \in (L^\infty(\mathbb{R}^n))^{n\times n}.$$

Also, assume they satisfy the Lipschitz condition

$$\|b(x) - b(y)\|_{\mathbb{R}^n} + \|\sigma(x) - \sigma(y)\|_{\mathbb{R}^{n\times n}} \leq C\|x - y\|_{\mathbb{R}^n}.$$

Then, there exists a unique strong solution to the SDE given X_0 initial data.

The same assumptions are actually enough to establish a unique continuous solution to

$$\partial_t f - b \cdot \nabla f - \sigma\sigma^T : \nabla^2 f = 0$$

for continuous initial data $f_0 \in C^1(\mathbb{R}_+, \mathbb{R}^n)$. The solution to the backward Kolmogorov equation is given by $f(t, x) = \mathbb{E}(f_0(X_t^{-1}))$ as a result of the famous Feynmann-Kac formula.

Going back to our original problem in this thesis, the SDE under study is

$$d\begin{pmatrix} x_t \\ v_t \end{pmatrix} = b(x_t, v_t)\, dt + \sqrt{2}\sigma(x_t, v_t)\, dW_t$$

for $b(x, v) = \begin{pmatrix} v \\ -Gv - \nabla\mathfrak{U}(x) \end{pmatrix}$ and $\sigma(x, v) = \begin{pmatrix} 0 & 0 \\ 0 & G^{1/2}(x) \end{pmatrix}$.

Conditions of Theorem 15 are trivial for the above drift and dispersion. The growth conditions are,

$$\frac{G(x)}{1 + |x|}, \quad \frac{G^{1/2}(x)}{1 + |x|} \in (L^\infty(\mathbb{R}^{3N}))^{3N\times 3N}, \qquad \frac{\nabla\mathfrak{U}(x)}{1 + |x|} \in (L^\infty(\mathbb{R}^{3N}))^{3N}$$

and the Lipschitz regularity condition is

$$\|G(x) - G(y)\|_{\mathbb{R}^{3N\times 3N}} + \|G^{1/2}(x) - G^{1/2}(y)\|_{\mathbb{R}^{3N\times 3N}} + \|\nabla\mathfrak{U}(x) - \nabla\mathfrak{U}(y)\|_{\mathbb{R}^{3N}}$$
$$\leq C\|x - y\|_{\mathbb{R}^{3N}}.$$

A.5 Commutator Algebra

We present some of the computations for the commutators of the operators we deal with in Chapter 4. In these computations assume summation over repeated indices. The operators L, & A, B are:

$$Lh = v \cdot \nabla_x h - \nabla \mathcal{U}(x) \cdot \nabla_v h - \nabla_v \cdot G(x) \nabla_v h + v \cdot G(x) \nabla_v h,$$

$$A = G^{1/2}(x) \nabla_v, \quad A^* = -G^{1/2}(x) \nabla_v + G^{1/2}(x) v, \quad \text{and} \quad B = v \cdot \nabla_x - \nabla \mathcal{U}(x) \cdot \nabla_v$$

with $B^* = -B$, where the adjoint is understood in the $L^2(\mu)$ sense.

First, we start with some easy ones e.g.

$$[\partial_{v_i}, -\partial_{v_j} + v_j]_{i,j} h = \partial_{v_i}(-\partial_{v_j} + v_j) h - (-\partial_{v_j} + v_j) \partial_{v_i} h$$

$$= -\partial^2_{v_i v_j} h + \delta_{ij} h + v_j \partial_{v_i} h + \partial^2_{v_i v_j} h - v_j \partial_{v_i} h = \delta_{ij} h.$$

$$[\partial_{v_i}, v_k \partial_{x_k} - \partial_{x_k} \mathcal{U}(x) \partial_{v_k}]_i h = \partial_{v_i}(v_k \partial_{x_k} h$$

$$- \partial_{x_k} \mathcal{U}(x) \partial_{v_k} h) - (v_k \partial_{x_k} h - \partial_{x_k} \mathcal{U}(x) \partial_{v_k} h) \partial_{v_i} h = \delta_{ik} \partial_{x_k} h + v_k \partial^2_{v_k x_k} h$$

$$- \partial_{x_k} \mathcal{U}(x) \partial^2_{v_i v_k} h - v_k \partial^2_{v_k x_k} h + \partial_{x_k} \mathcal{U}(x) \partial^2_{v_i v_k} h = \partial_{x_i} h.$$

$$[\partial_{v_i}, \partial_{x_j}]_{i,j} h = 0.$$

$$[-\partial_{v_i} + v_i, \partial_{x_j}]_{i,j} h = 0.$$

In other words,

$$[\nabla_v, \nabla_v^*] = I,\ [\nabla_v, B] = \nabla_x,\ [\nabla_v, \nabla_x] = 0,\ [\nabla_v^*, \nabla_x] = 0.$$

$$[A_i,A_j^*]_{i,j}h = -G_{ik}^{1/2}(x)\partial_{v_k}G_{jm}^{1/2}(x)(\partial_{v_m}-v_m)h + G_{jm}^{1/2}(x)(\partial_{v_m}-v_m)G_{ik}^{1/2}(x)\partial_{v_k}h$$

$$= -G_{ik}^{1/2}(x)G_{jm}^{1/2}(x)\partial_{v_kv_m}h + G_{ik}^{1/2}(x)G_{jm}^{1/2}(x)\delta_{km}h + G_{ik}^{1/2}(x)G_{jm}^{1/2}(x)v_m\partial_{v_k}h$$

$$- G_{ik}^{1/2}(x)G_{jm}^{1/2}(x)v_m\partial_{v_k}h + G_{jm}^{1/2}(x)G_{ik}^{1/2}(x)\partial_{v_mv_k}h = G_{ij}(x)h.$$

$$Ch = [A_i,B]_ih = G_{ik}^{1/2}(x)\partial_{v_k}(v_m\partial_{x_m} - \partial_{x_m}\mathcal{U}(x)\partial_{v_m})h$$

$$-(v_m\partial_{x_m} - \partial_{x_m}\mathcal{U}(x)\partial_{v_m})G_{ik}^{1/2}(x)\partial_{v_k}h = G_{ik}^{1/2}(x)\delta_{km}\partial_{x_m}h + G_{ik}^{1/2}(x)v_m\partial_{x_m}\partial_{v_k}h$$

$$- G_{ik}^{1/2}(x)\partial_{x_m}\mathcal{U}(x)\partial_{v_mv_k}h - G_{ik}^{1/2}(x)v_m\partial_{x_m}\partial_{v_k}h + G_{ik}^{1/2}(x)\partial_{x_m}\mathcal{U}(x)\partial_{v_mv_k}h$$

$$- v_m(\partial_{x_m}G_{ik}^{1/2}(x))\partial_{v_k}h = G_{ik}^{1/2}(x)\partial_{x_k}h - v_m(\partial_{x_m}G_{ik}^{1/2}(x))\partial_{v_k}h.$$

Now it is time to compute commutators of L in directions of certain derivatives e.g.

$$(\partial_t + L)\partial_{x_i}h = -\partial_{x_i}Lh + L\partial_{x_i}h = [L,\partial_{x_i}]h.$$

Here,

$$-\partial_{x_i}Lh = -v_j\partial^2_{x_ix_j}h + \partial^2_{x_ix_j}\mathcal{U}(x)\partial_{v_j}h + \partial_{x_j}\mathcal{U}(x)\partial^2_{x_iv_j}h + (\partial_{x_i}G_{jk}(x))\partial^2_{v_jv_k}h$$

$$+ G_{jk}(x)\partial^3_{x_iv_jv_k}h - v_j(\partial_{x_i}G_{jk}(x))\partial_{v_k}h - v_jG_{jk}(x)\partial^2_{x_iv_k}h.$$

$$L\partial_{x_i}h = v_j\partial^2_{x_ix_j}h - \partial_{x_j}\mathcal{U}(x)\partial^2_{x_iv_j}h - G_{jk}(x)\partial^3_{x_iv_jv_k}h + v_jG_{jk}(x)\partial^2_{x_iv_k}h.$$

Hence,

$$[L,\partial_{x_i}]h = \partial^2_{x_ix_j}\mathcal{U}(x)\partial_{v_j}h + (\partial_{x_i}G_{jk}(x))\partial^2_{v_jv_k}h - v_j(\partial_{x_i}G_{jk}(x))\partial_{v_k}h.$$

In vector notation this is

$$(\partial_t + L)\nabla_x h = \nabla^2 \mathcal{U}(x)\nabla_v h + \nabla_x G(x)\nabla_v^2 h - (\nabla_x G(x)\cdot v)\nabla_v h. \tag{A.1}$$

We now compute the commutator $(\partial_t + L)\partial_{v_i} h = [L, \partial_{v_i}]h$.

$$-\partial_{v_i} Lh = -\partial_{x_i} h - v_j \partial^2_{x_j v_i} h + \partial_{x_j}\mathcal{U}(x)\partial^2_{v_j v_i} h + G_{jk}(x)\partial^3_{v_j v_k v_i} h$$
$$- G_{ik}(x)\partial_{v_k} h - v_j G_{jk}(x)\partial^2_{v_k v_i} h,$$

also

$$L\partial_{v_i} h = v_j \partial^2_{x_j v_i} h - \partial_{x_j}\mathcal{U}(x)\partial^2_{v_j v_i} h - G_{jk}(x)\partial^3_{v_k v_i v_j} h + v_j G_{jk}(x)\partial^2_{v_k v_i} h.$$

Hence,

$$[L, \partial_{v_i}]h = -\partial_{x_i} h - G_{ij}(x)\partial_{v_j} h \quad ,$$

or in vector notation

$$(\partial_t + L)\nabla_v h = -\nabla_x h - G(x)\nabla_v h. \tag{A.2}$$

The evolution of the mixed derivative

$$(\partial_t + L)\partial_{x_k} h \cdot \partial_{v_k} h = -\partial_{x_k} Lh \cdot \partial_{v_k} h - \partial_{x_k} h \cdot \partial_{v_k} Lh + L(\partial_{x_k} h \cdot \partial_{v_k} h)$$

will require the computation of the following:

$$-\partial_{x_k} Lh \cdot \partial_{v_k} h = -v_j \partial^2_{x_k x_j} h \partial_{v_k} h + \partial^2_{x_k x_j}\mathcal{U}(x)\partial_{v_j} h \partial_{v_k} h + \partial_{x_j}\mathcal{U}(x)\partial^2_{x_k v_j} h \partial_{v_k} h$$
$$+ (\partial_{x_k} G_{jm}(x))\partial^2_{v_j v_m} h \partial_{v_k} h + G_{jm}(x)\partial^3_{x_k v_j v_m} h \partial_{v_k} h$$
$$- v_j(\partial_{x_k} G_{jm}(x))\partial_{v_m} h \partial_{v_k} h - v_j G_{jm}(x)\partial^2_{x_k v_m} h \partial_{v_k} h,$$

$$-\partial_{x_k}h\cdot\partial_{v_k}Lh = -|\partial_{x_k}h|^2 - v_j\partial^2_{x_jv_k}h\partial_{x_k}h + \partial_{x_j}\mathcal{U}(x)\partial^2_{v_jv_k}h\partial_{x_k}h$$

$$+ G_{jm}(x)\partial^3_{v_jv_mv_k}h\partial_{x_k}h - G_{kj}(x)\partial_{v_j}h\partial_{x_k}h - v_jG_{jm}(x)\partial^2_{v_mv_k}h\partial_{x_k}h,$$

and

$$L(\partial_{x_k}h\cdot\partial_{v_k}h) = v_j\partial_{x_j}(\partial_{x_k}h\cdot\partial_{v_k}h) - \partial_{x_j}\mathcal{U}(x)\partial_{v_j}(\partial_{x_k}h\cdot\partial_{v_k}h)$$

$$- G_{jm}(x)\partial^2_{v_jv_m}(\partial_{x_k}h\cdot\partial_{v_k}h) + v_jG_{jm}(x)\partial_{v_m}(\partial_{x_k}h\cdot\partial_{v_k}h).$$

The sum of the three terms above is

$$(\partial_t + L)\partial_{x_k}h\cdot\partial_{v_k}h = \partial^2_{x_kx_j}\mathcal{U}(x)\partial_{v_j}h\partial_{v_k}h - |\partial_{x_k}h|^2 - G_{kj}(x)\partial_{v_j}h\partial_{x_k}h$$

$$+ (\partial_{x_k}G_{jm}(x))\partial^2_{v_jv_m}h\partial_{v_k}h - v_j(\partial_{x_k}G_{jm}(x))\partial_{v_m}h\partial_{v_k}h - 2G_{jm}(x)\partial^2_{v_mx_k}h\partial^2_{v_jv_k}h.$$

Finally, the sum of the above in vector notation yields

$$(\partial_t + L)\nabla_xh\cdot\nabla_vh = \nabla_vh\cdot\nabla^2\mathcal{U}(x)\nabla_vh - |\nabla_xh|^2 - \nabla_vh\cdot G(x)\nabla_xh \tag{A.3}$$

$$+ \nabla_vh\cdot(\nabla_xG(x)\nabla^2_vh) - \nabla_vh\cdot(\nabla_xG(x)\cdot v)\nabla_vh - 2(G(x)\nabla^2_{vx}h) : \nabla^2_vh.$$

The third order derivative in velocity, when commuted with L i.e. $[L, \partial^2_{v_iv_jv_k}]$, gives (for the case of an identity diffusion matrix $G(x) = \mathrm{I}$)

$$(\partial_t + L)\partial^3_{v_iv_jv_k}h = -3\partial^3_{v_iv_jv_k}h - \partial^3_{x_iv_jv_k}h - \partial^3_{v_ix_jv_k}h - \partial^3_{v_iv_jx_k}h.$$

In the case of a matrix $G(x)$, this is given by

$$(\partial_t + L)\partial^3_{v_iv_jv_k}h = -G_{km}(x)\partial^3_{v_iv_jv_m}h - G_{jm}(x)\partial^3_{v_iv_mv_k}h - G_{im}(x)\partial^3_{v_mv_jv_k}h$$

$$- \partial^3_{x_iv_jv_k}h - \partial^3_{v_ix_jv_k}h - \partial^3_{v_iv_jx_k}h.$$

A.6 Csiszár-Kullback-Pinsker Inequality & other Inequalities based on Convexity of Entropy

We have employed the Csiszár-Kullback-Pinsker inequality in Chapters 5 & 6. This inequality gives a bound on the total variation between two probability measures μ, ν in terms of the relative entropy of μ w.r.t. measure ν. The original derivation of the inequality can be found in [12, 47, 59]. The exact CKP inequality (with optimal constant) for two measures μ,ν is $\|\mu-\nu\|_{TV} \leq \sqrt{\frac{1}{2}H(\mu|\nu)}$, where $\|\mu-\nu\|_{TV}$ is the total variation between μ, ν and $H(\mu|\nu)$ the Kullback information of μ w.r.t. ν as will be defined shortly after. In the case of measures μ, ν with corresponding densities g_1, g_2 ($g_1 = \frac{d\mu}{dx}$ etc.) the CKP inequality is $\|g_1 - g_2\|_{L^1} \leq \sqrt{2H(g_1|g_2)}$ (since $\|\mu-\nu\|_{TV} = \frac{1}{2}\|g_1 - g_2\|_{L^1}$).

Here we follow a proof which can be found e.g. in [4] and is attributed to Talagrand. To fix things, we assume a Polish space X, $P(X)$ is the set of Borel probability measures on X, and μ,ν two measures in $P(X)$.We define the Kullback information of μ w.r.t. ν by

$$H(\mu|\nu) = \int_X f \log f \, d\nu, \qquad f = \frac{d\mu}{d\nu}$$

if $\mu \ll \nu$, and $H(\mu|\nu) = +\infty$ otherwise. Remember also that the total variation is defined by $\|\mu-\nu\|_{TV} = \sup\limits_{A\subset X} |\mu(A) - \nu(A)|$.

Now consider the function $h(t) = (1+t)\log(1+t) - t$, for $t > -1$. With the help of h, we can write

$$H(\mu|\nu) = \int_X h(u) \, d\nu \qquad \text{for} \quad u := f - 1.$$

Notice that $h'(t) = \log(1+t)$ & $h''(t) = \frac{1}{1+t}$, implying $h(0) = h'(0) = 0$. The Taylor expansion of h about 0 with the remainder term in integral form yields

$$h(t) = \int_0^t (t-x)h''(x)\,dx = t^2 \int_0^1 \frac{1-s}{1+st}\,ds.$$

We now infer from a simple CS inequality that

$$\int_X |u|(x)\,d\nu \int_0^1 (1-s)\,ds \le \left(\int_X \int_0^1 \frac{u^2(x)(1-s)}{1+su(x)}\,ds\,d\nu\right)^{\frac{1}{2}} \times \left(\int_X \int_0^1 (1-s)(1+su(x))\,ds\,d\nu\right)^{\frac{1}{2}}.$$

The last inequality implies

$$\int_X |1-f|\,d\nu \le C\left(\int_X \int_0^1 \frac{u^2(x)(1-s)}{1+su(x)}\,ds\,d\nu\right)^{\frac{1}{2}},$$

for the constant $C = \frac{\left(\int_X \int_0^1 (1-s)(1+su(x))\,ds\,d\nu\right)^{\frac{1}{2}}}{\int_0^1 (1-s)\,ds}$ which can be explicitly computed and has the exact value $C = \sqrt{2}$. Since $\|\mu-\nu\|_{TV} = \frac{1}{2}\int_X |1-f|\,d\nu$, the CKP inequality follows.

It is now time to prove the inequality $H(\rho_1|\rho_2) \le H(f_1|f_2)$ mentioned earlier in this study, where ρ_1, ρ_2 are hydrodynamical variables associated to f_1, f_2. The relative entropy mentioned here is the one in $LlogL$ sense.

Proof. For the probability measures ρ_1, ρ_2, we have

$$\begin{aligned} H(\rho_1|\rho_2) &= \int \rho_1 \log\frac{\rho_1}{\rho_2}\,dx = \int \left(\rho_1 \log\frac{\rho_1}{\rho_2} - \rho_1 + \rho_2\right)dx \\ &= \int \rho_2 \left(\frac{\rho_1}{\rho_2}\log\frac{\rho_1}{\rho_2} - \frac{\rho_1}{\rho_2} + 1\right)dx. \end{aligned}$$

Once again, we consider the convex function $\phi(x) = x\log x - x + 1$. The above expression can now be written as

$$
\begin{aligned}
H(\rho_1|\rho_2) &= \int \rho_2\, \phi\left(\frac{\rho_1}{\rho_2}\right) dx = \int \rho_2\, \phi\left(\int \frac{f_1}{f_2}\frac{f_2}{\rho_2}\, dv\right) dx \\
&\leq \iint \rho_2\, \phi\left(\frac{f_1}{f_2}\right)\frac{f_2}{\rho_2}\, dv\, dx = \iint f_2 \left(\frac{f_1}{f_2}\log\frac{f_1}{f_2} - \frac{f_1}{f_2} + 1\right) dv\, dx \\
&= H(f_1|f_2).
\end{aligned}
$$

The inequality employed here is a Jensen's inequality for the measure $\frac{f_2}{\rho_2}$ in $L^1(dv)$.

□

A.7 Poincaré & Log-Sobolev Inequalities

We give sufficient conditions for a probability measure μ with density $e^{-\mathcal{U}(x)}$ ($d\mu = e^{-\mathcal{U}(x)}dx$), in order that it satisfies Poincaré and log-Sobolev inequalities. Since the need for such results appeared in the study for convergence rates, we will typically introduce these inequalities in their connection with convergence rates for

$$\partial_t \rho = \nabla \cdot \left(D(x)(\nabla \rho + \nabla \mathcal{U}(x)\rho) \right), \tag{A.4}$$

for a function $\rho(x,t)$ on $\mathbb{R}^n \times \mathbb{R}_+$, with initial data $\rho_0(x) \geq 0$ satisfying $\int \rho_0 \, dx = 1$ and $D(x) \geq 0$ being symmetric, uniformly strictly positive $n \times n$ matrix.

There is a broad choice of functionals (entropies) for the study of convergence rates. Assume a strictly convex function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ ($\phi''(x) \geq 0$), with the additional restrictions

$$\phi(1) = \phi'(1) = 0.$$

In order to construct an admissible relative entropy, one considers the extra condition

$$2(\phi''')^2 \leq \phi'' \phi^{(4)}.$$

We introduce the relative entropy related to ϕ, i.e

$$H_\phi(\rho | e^{-\mathcal{U}(x)}) = \int \phi(h) \, d\mu,$$

where $h = \rho / e^{-\mathcal{U}(x)}$.

The entropy dissipation rate $\mathcal{D}_\phi(\rho | e^{-\mathcal{U}(x)}) = -\frac{d}{dt} H_\phi(\rho | e^{-\mathcal{U}(x)})$, after computation yields

$$\mathcal{D}_\phi(\rho | e^{-\mathcal{U}(x)}) = \int \phi''(h) \nabla_x h \cdot D(x) \nabla_x h \, d\mu. \tag{A.5}$$

One such choice of functions ϕ is

$$\phi_p(x) = \frac{x^p - 1 - p(x-1)}{p-1}, \qquad p \in (1,2].$$

These functions satisfy all the conditions mentioned, including the admissibility condition. The two obvious choices for entropies are the $L^2(d\mu)$ space for $\phi_2(h) = (h-1)^2$, and the $LlogL(d\mu)$ for $\phi_1(h) = h\log h - h + 1$ (which corresponds to the limiting case $p \downarrow 1$).

The dissipation rates for these two functionals are

$$\begin{aligned}
\mathcal{D}_2(\rho|e^{-\mathcal{U}(x)}) &= 2\int \nabla_x h \cdot D(x)\nabla_x h \, d\mu \quad , \\
\mathcal{D}_1(\rho|e^{-\mathcal{U}(x)}) &= \int h \left| D(x)^{1/2}\nabla_x \log h \right|^2 d\mu \\
&= 4\int \left| D(x)^{1/2}\nabla_x \sqrt{h} \right|^2 d\mu.
\end{aligned}$$

Restricting ourselves to $D(x) = \mathrm{I}$, the Poincaré inequality with constant $\lambda > 0$ for a measure $e^{-\mathcal{U}(x)}\,dx$ in this setting, is

$$\mathcal{D}_2(\rho|e^{-\mathcal{U}(x)}) \geq 2\lambda H_2(\rho|e^{-\mathcal{U}(x)}).$$

Integrating from initial time $t = 0$ to time t we can prove exponential decay for the H_2 relative entropy

$$H_2(\rho(t)|e^{-\mathcal{U}(x)}) \leq H_2(\rho(0)|e^{-\mathcal{U}(x)})e^{-2\lambda t}.$$

In this case, we say that μ admits a spectral gap with constant λ. The following theorem gives a sufficient condition for the measure μ to satisfy a Poincaré inequality.

Theorem 24. *Let $\mathcal{U}(x) \in C^2(\mathbb{R}^n)$ s.t μ is a probability measure with density $e^{-\mathcal{U}(x)}$. If*

$$\frac{|\nabla\mathcal{U}(x)|^2}{2} - \Delta\mathcal{U}(x) \to \infty \quad \textit{as} \quad |x| \to \infty,$$

then a Poincaré inequality holds for measure μ and some constant $\lambda > 0$.

Proof. The proof can be found in [71] pg 132. $\square$

The log-Sobolev inequality with constant $\lambda > 0$ for the measure $e^{-\mathcal{U}(x)}\,dx$ is by definition

$$\mathcal{D}_1(\rho|e^{-\mathcal{U}(x)}) \geq 2\lambda H_1(\rho|e^{-\mathcal{U}(x)}).$$

As previously, integration in time implies exponential decay for the H_1 relative entropy which by means of the Csiszar-Kullback-Pinsker inequality gives exponential convergence to the stationary solution in the L^1 norm.

In standard formulation, a log-Sobolev inequality with constant $\lambda > 0$ for a measure μ is satisfied for a function h on $\mathbb{R}^n$ iff

$$\int h^2 \log h^2\,d\mu - \left(\int h^2\,d\mu\right)\log\left(\int h^2\,d\mu\right) \leq \frac{2}{\lambda}\int |\nabla_x h|^2\,d\mu.$$

The original version of a log-Sobolev inequality states that the inequality is satisfied with constant 1 i.e.

$$\int h^2 \log h^2\,d\nu - \left(\int h^2\,d\nu\right)\log\left(\int h^2\,d\nu\right) \leq 2\int |\nabla_x h|^2\,d\nu,$$

for the standard Gaussian measure $d\nu = (2\pi)^{-n/2}e^{-x^2/2}\,dx$ for any real function. This allows the embedding $H^1(d\nu) \subset L^2 \log L^2(d\nu)$. Proof of the original log-Sobolev inequality for a standard Gaussian can be found in many sources e.g [8, 28] etc.

Sufficient conditions exist that guarantee the validity of a log-Sobolev inequality, when ρ solves (A.4). The Bakry-Emery condition that appeared in [2] is the original mention of such a condition. We now present this condition, as well as two simplified versions of the condition for a diagonal and an identity matrix $D(x)$.

Theorem 25. *Consider a probability measure* μ *with density* $e^{-\mathcal{U}(x)}$.

(i) Assume a symmetric, uniformly positive matrix $D(x)$ *and that there exists* $\lambda_1 > 0$ *s.t.*

$$Ric \geq \lambda_1 D(x).$$

Then, for any function ρ *that solves* (A.4), ρ *satisfies a log-Sobolev inequality with constant* λ_1. *Here, Ric is the Ricci curvature tensor for the manifold*
$M = (\mathbb{R}^n, D(x)^{-1})$.

(ii) The condition for a scalar diffusion $D(x) = d(x)\mathrm{I}$ *is*

$$\left(\frac{1}{2} - \frac{n}{4}\right)\frac{1}{d(x)}\nabla d(x) \otimes \nabla d(x) + \frac{1}{2}(\Delta d(x) - \nabla d(x) \cdot \nabla \mathcal{U}(x))\mathrm{I}$$
$$+ d(x)\nabla^2\mathcal{U}(x) + \frac{\nabla\mathcal{U}(x) \otimes \nabla d(x) + \nabla d(x) \otimes \nabla \mathcal{U}(x)}{2} - \nabla^2 d(x) \geq \lambda_1 \mathrm{I}.$$

Finally,

(iii) the condition for identity diffusion $D(x) = \mathrm{I}$ *is*

$$\nabla^2\mathcal{U}(x) \geq \lambda_1 \mathrm{I}.$$

Remark 16. *The reason for the different versions of the Bakry-Emery condition is due to the following observation. Let* $\mathcal{D}_\phi^D$ *be the entropy dissipation as seen in* (A.5) *for the dissipation rate associated with the F-P operator* $L\rho = \nabla\cdot(D(\nabla\rho + \nabla\mathcal{U}(x)\rho))$, $D(x) \geq 0$. *Assume two matrices* $D_1(x)$, $D_2(x)$ *that satisfy*

$$D_1(x) \leq D_2(x) \qquad \forall x \in \mathbb{R}^n$$

in the sense of positive definite matrices. The dissipation rates for these two matrices satisfy

$$\mathcal{D}_\phi^{D_2}(\rho|e^{-\mathcal{U}(x)}) \leq \mathcal{D}_\phi^{D_1}(\rho|e^{-\mathcal{U}(x)}).$$

An immediate consequence is the following. If for any uniformly positive $D(x)$ there exists a function $d(x) > 0$ s.t $\forall x \in \mathbb{R}^n$

$$0 < d(x)\mathrm{I} \leq D(x),$$

then condition (ii) settles the exponential decay for $D(x)$ in the $H_1(h)$ entropy. A similar scenario holds for the condition (iii), if there exists some $d > 0$ s.t. $\forall x \in \mathbb{R}^n$

$$0 < d\mathrm{I} \leq D(x).$$

Proof. The theorem can be proved with an "inverted" point of view. We differentiate the dissipation rate and establish the inequality

$$-\frac{d}{dt}\mathcal{D}_\phi(\rho|e^{-\mathcal{U}(x)}) \geq 2\lambda_1\mathcal{D}_\phi(\rho|e^{-\mathcal{U}(x)})$$

which yields exponential decay for the entropy dissipation functional. The remaining steps are :

(a) Integrating w.r.t. time from t to $+\infty$ to establish the log-Sobolev inequality

(b) Since the computation is done formally, a density argument should be performed in the end.

We omit the proof of the full Bakry-Emery condition $Ric \geq \lambda_1 D(x)$ because it requires the language of differential geometry for its full understanding. The reader is directed to the original source [2] for the proof of (i), or even [1]. We restrict ourselves to the interesting case (ii) of the condition, which is more general than (iii).

We start by computing the time derivative of the dissipation rate i.e.

$$\frac{d}{dt}\mathcal{D}_\phi(\rho|e^{-\mathcal{U}(x)}) = -\int \phi'''(h)\nabla\cdot(D\nabla h e^{-\mathcal{U}(x)})\nabla h^T D\nabla h\, dx \\ -2\int \phi''(h)\nabla h^T D\nabla h_t\, d\mu.$$

Since $\nabla h_t = \nabla(e^{\mathcal{U}(x)}\nabla\cdot(De^{-\mathcal{U}(x)}\nabla h))$ and using the fact that D is diagonal $D = d(x)\mathrm{I}$, after a tedious algebra which can be traced in [1], the grouping of different terms will give

$$-\frac{d}{dt}\mathcal{D}_\phi(\rho|e^{-\mathcal{U}(x)}) = \int tr(XY)\, d\mu + 2\lambda_1\int \phi''(h)d(x)|\nabla h|^2\, d\mu,$$

where the matrices X, Y are

$$X = \begin{pmatrix} 2\phi''(h) & 2\phi'''(h) \\ 2\phi'''(h) & \phi^{(4)}(h) \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

for

$a = \sum_{i,j}\left(d(x)\nabla^2_{x_ix_j}h + \frac{1}{2}\nabla_{x_i}d(x)\nabla_{x_j}h + \frac{1}{2}\nabla_{x_i}h\nabla_{x_j}d(x) - \frac{1}{2}\delta_{ij}\nabla d(x)\cdot\nabla h\right)^2$,

$b = d(x)^2\nabla h\cdot\nabla^2 h\nabla h + \frac{1}{2}d(x)|\nabla h|^2\nabla h\cdot\nabla d(x)$, and

$c = d(x)^2|\nabla h|^4$.

It can be shown that both X and Y are non negative matrices, which yields

$$-\frac{d}{dt}\mathcal{D}_\phi(\rho|e^{-\mathcal{U}(x)}) \geq 2\lambda_1\mathcal{D}_\phi(\rho|e^{-\mathcal{U}(x)}),$$

and concludes the proof.

$\square$

Bibliography

[1] ARNOLD, A., MARKOWICH, P., TOSCANI, G., AND UNDERREITER, A. *On Convex Sobolev inequalities and rate of convergence to equilibrium for Fokker-Planck type equations.* Commun. Part. Diff. Eq. **26**, 1-2 (2001), 43-100.

[2] BAKRY, D., AND EMERY, M. *Diffusions hypercontractives.* Lect. Notes Math. **19** (1985), 177-206.

[3] BIRD, B., ARMSTRONG, R., CURTISS, C., AND HASSAGER, O. *Dynamics of polymeric liquids: Kinetic Theory vol* **2.** John Wiley & Sons, Second edition, 1994.

[4] BOLLEY, F., AND VILLANI, C. *Weighted Csiszár-Kullback-Pinsker inequalities and applications to trasportation inequalities.* Ann. Fac. Sci. Toulouse Math. **14**, (3) 2005, 331-352.

[5] BOUCHUT, F. *Existence and uniqueness of a global smooth solution for the Vlassov -Poisson-Fokker-Planck system in three dimensions.* J.Funct. Anal. **111**, 1 (1993), 239-258.

[6] BOUCHUT, F. *Smoothing effect for the non-linear Vlassov-Poisson-Fokker-Planck system.* J. Differential Equations. **122**, 2 (1995), 225-238.

[7] BOUCHUT, F. *Hypoelliptic regularity in kinetic equations.* J. Math. Pure. Appl. **81**, 11 (2002), 1135-1159.

[8] CARLEN, E., AND LOSS, M. *Logarithmic Sobolev Inequalitites and Spectral Gaps, in Recent advances in the theory and applications of mass transport.* Contemp. Math. **353** (2004), 53-60.

[9] Carrilo, J., and Soler, J. *On the initial value problem for the Vlassov-Poisson-Fokker -Planck system with initial data in L^p spaces.* Math. Method. Appl. Sci. **18**, 10 (1995), 825-839.

[10] Chandrasekhar, S. *Stochastic problems in Physics and Astronomy.* Rev. Mod. Phys. **15**, 1 (1943), 1-89.

[11] Champagnat, N., and Jabin, P-E. *Strong solutions to stochastic differential equations with rough coefficients.* Submitted Preprint, 2013.

[12] Csiszár, I. *Information-type measures of difference of probability distributions and indirect observations.* Stud. Sci. Math. Hung. **2**, (1967), 299-318.

[13] Degond, P. *Global existence of smooth solutions for the Vlassov-Fokker-Planck equation in* 1 *and* 2 *space dimensions.* Ann. Sci. Ecole Norm. Sup. **19**, 4 (1986), 519-542.

[14] Degond, P., and Liu, H. *Kinetic models for polymers with inertial effects.* Networks and Heterogeneous Media **4**, 4 (2009), 625-647.

[15] Desvillettes, L., and Villani, C. *On the trend to global equilibrium in spatially inhomogeneous entropy dissipating systems: the linear Fokker-Planck equation.* Commun. Pur. Appl. Math. **54**, 1 (2001), 1-42.

[16] Desvillettes, L., and Villani, C. *On the trend to global equilibrium for spatially inhomogeneous kinetic systems: the Boltzmann equation.* Invent. Math. **159**, 2 (2005), 245-316.

[17] Diperna, R.J., and Lions, P.L. *Ordinary differential equations, transport theory and Sobolev spaces.* Invent. Math. **98**, 3 (1989), 511-547.

[18] Doi, M. *Introduction to Polymer Physics.* Oxford University Press, 1996.

[19] Doi, M., and Edwards, S.F. *The theory of polymer dynamics.* Oxford University Press, New York, 1986.

[20] Dolbeault, J., Mouhot, C., and Schmeiser, C. *Hypocoercivity for Linear Kinetic Equations Conserving Mass*, 2010.

[21] Ebeling, W., Gudowska-Nowak, E., and Sokolov, I.M. *On Stochastic Dynamics in Physics-Remarks on History and Terminology.* Acta Phys. Pol. B **39** (2008), 1003-1019.

[22] FOKKER, A.D. *Die mittlere Energie rotierender elektrischer Dipole im Strahlungsfeld.* Ann. Phys. **348**, 5 (1914), 810-820.

[23] FREIDLIN, M. *Some remarks on the Smoluchowski-Kramers Approximation.* J. Stat. Phys. **117**, 3-4 (2004), 617-634.

[24] GHANI, N., AND MASMOUDI, N. *Diffusion Limit of The Vlassov-Poisson-Fokker-Planck System.* Commun. Math. Sci. **8**, 2 (2010), 463-479.

[25] GOUDON, T. *Hydrodynamic limit for the Vlasov-Poisson-Fokker-Planck system: Analysis of the two-dimensional case.* Math. Mod. Meth. Appl. S. **15**, 737 (2005), 737-752.

[26] GOUDON, T., JABIN, P-E., AND VASSEUR, A. *Hydrodynamic Limit for the Vlasov-Navier-Stokes Equation. Part I:Light Particles Regime.* Indiana U. Math. J. **53**, 6 (2004), 1495-1515.

[27] GOUDON, T., JABIN, P-E., AND VASSEUR, A. *Hydrodynamic Limit for the Vlasov-Navier-Stokes Equation. Part II:Fine Particles Regime.* Indiana U. Math. J. **53**, 6 (2004), 1517-1536.

[28] GROSS, L. *Logarithmic Sobolev Inequalities.* Am. J. Math. **97**, 4 (1975), 1061-1083.

[29] HELFFER, B., AND NIER, F. *Hypoelliptic Estimates and Spectral Theory for Fokker-Planck Operators and Witten Laplacians.* Lect. Notes Math. **1862**, Springer, 2005.

[30] HÉRAU, F. *Hypocoercivity and exponential time decay for the linear inhomogeneous relaxation Boltzmann equation.* Asymptotic Anal. **46** , 3-4 (2006), 349-359.

[31] HÉRAU, F. *Short and long time behavior of the Fokker-Planck equation in a confining potential and applications.* J. Funct. Anal. **244**, 1 (2007), 95-118.

[32] HÉRAU, F., AND NIER, F. *Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with high degree potential.* Arch. Ration. Mech. An. **171**, 2 (2004), 151-218.

[33] HOLLEY, R., AND STROOCK, D. *Logarithmic Sobolev inequalities and Stochastic Ising models.* J. Stat. Phys. **46**, 5-6 (1987), 1159-1194.

[34] Hörmander, L. *Hypoelliptic Second Order Differential Equations.* Acta Math. **119** (1967), 147-171.

[35] Jabin, P-E. *Private communication.*

[36] Jabin, P-E., and Otto, F. *Identification of the Dilute Regime in Particle Sedimentation.* Commun. Math. Phys. **250**, 2 (2004), 415-432.

[37] Jabin, P-E., and Perthame, B. *Notes on mathematical problems on the dynamics of dispersed particles interacting through a Fluid in Modeling in applied sciences, a kinetic theory approach.* N. Bellomo, M. Pulvirenti Eds, Birkhäuser (2000), 111-147.

[38] Jannick, G., and des Cloizeaux, J. *Polymers in Solution : Their Modelling and Structure.* Oxford University Press, 1990.

[39] Karatzas, I., and Shreve, S. *Brownian motion and Stochastic calculus.* Graduate Texts in Mathematics **113**, Springer, 1988.

[40] Khoklov, A., and Grosberg, A. *Statistical Physics of macromolecules (Polymers and Complex Materials).* AIP Series in Polymers and Complex Materials, AIP Press, New York, 1994.

[41] Kirkwood, J.G. *John Gamble Kirkwood Collected Works : Macromolecules vol* **3.** Documents on modern physics, Gordon and Breach, 1967.

[42] Klein, O. *Zur statistischen Theorie der Suspensionen und Lösungen.* Arkiv för matematik, Astronomi och Fysik **16**, 5 (1921), 1-51.

[43] Kohn, J.J. *Lectures on Degenerate Elliptic Problems.* Pseudodifferential Operators with Applications. C.I.M.E Summer Schools **75** (2011), 89-151.

[44] Kolmogorov, A. *Zur Theorie der Stetigen Zufällige Prozesse.* Math. Ann. **108** (1933), 149-160.

[45] Kolmogorov, A. *Zufällige Bewegungen (Zur Theorie der Brownschen Bewegung).* Ann. Math. **35** (1934), 116-117.

[46] Kramers, H.A. *Brownian motion in a field of force and the diffusion model of chemical reactions.* Physica **7**, 4 (1940), 284-304.

[47] Kullback, S. *A lower bound for discrimination information in terms of variation.* IEEE T. Inform. Theory **13**, 1 (1967), 126-127.

[48] Lamb, H. *Hydrodynamics.* Cambridge University Press, Sixth Edition, 1975.

[49] Le Bris, C., and Lions, P-L. *Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications.* Ann. Mat. Pur. Appl. **183**, 1 (2004), 97-130.

[50] Le Bris, C., and Lions, P-L. *Existence and Uniqueness of Solutions to Fokker-Planck Type Equations with Irregular Coefficients.* Commun. Part. Diff. Eq. **33**, 7 (2008), 1272-1317.

[51] Lerner, N., and Pravda-Starov K. *Sharp hypoelliptic estimates for some Kinetic equations.* Preprint, 2011.

[52] Markowich, P., and Villani, C. *On the Trend to Equilibrium for the Fokker-Planck Equation: An Interplay Between Physics and Functional Analysis.* Math. Contemp. **19** (2000), 1-29.

[53] Nagy, G. *Essentials of Pseudodifferential Operators.* Preprint, 2004.

[54] Nier, F. *Hypoellipticity for Fokker-Planck operators and Witten Laplacians.* Preprint, 2009.

[55] Øksendal, B. *Stochastic differential equations. An introduction with applications.* Universitext Springer-Verlag, Sixth Edition, Berlin, 2003.

[56] Oseen, C.H. *Neure Methoden und Ergebnisse in der Hydrodynamic.* Akademische Verlagsgesellschaft, Leipzig, 1927.

[57] Otto, F., and Tzavaras, A. *Continuity of Velocity Gradients in Suspensions of Rod-like Molecules.* Commun. Math. Phys. **277**, 3 (2008), 729-758.

[58] Pazy, A. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Appl. Math. Sci. **44**, Springer, 1983.

[59] Pinsker, M. *Information and information stability of random variables and processes.* Holden-Day, San Francisco, 1964.

[60] Planck, M. *An essay on statistical dynamics and its amplification in the quantum theory.* Sitzber. Preuss. Akad. Wiss., (1917), 324-341.

[61] Poupaud, F., and Soler, J. *Parabolic limit and Stability of the Vlasov-Fokker-Planck system.* Math. Mod. Meth. Appl. S. **10**, 7 (2000), 1027-1045.

[62] Pulvirenti, M., and Simeoni, C. *L^∞ estimates for the Vlassov-Poisson-Fokker-Planck equation.* Math. Method. Appl. Sci. **23**, 10 (2000), 923-935.

[63] Reed, M., and Simon, B. *Methods of Modern Mathematical Physics I & II.* Academic Press, New York, 1975.

[64] Risken, H. *The Fokker-Planck Equation. Methods of Solution and Applications.* Springer Series in Synergetics **18**, Second edition, Berlin, 1989.

[65] Rotne, J., and Prager, S. *Variational Treatment of Hydrodynamic Interactions in Polymers.* J. Chem. Phys. **50**, 4831 (1969), 4831-4837.

[66] Smoluchowski, M. *Drei Vortrage über Diffusion Brownsche Bewegung and Koagulation von Kolloidteilchen.* Phys. Zeit. **17** (1916), 557-585.

[67] Taylor, M. *Pseudodifferential Operators.* Princeton University Press, 1981.

[68] Varadhan, S. *Entropy methods in hydrodynamic scaling.* Proceedings of the International Congress of Mathematicians,Birkhäuser,Basel, (1995), 196-208.

[69] Victory, H. *On the existence of global weak solutions for Vlassov-Poisson-Fokker-Planck systems.* J. Math. Anal. Appl. **160**, 2 (1991), 525-555.

[70] Victory, H., and O'Dwyer, B. *On classical solutions of Vlassov-Poisson-Fokker-Planck systems.* Indiana U. Math. J. **39**, 1 (1990), 105-156.

[71] Villani, C. *Hypocoercivity.* Mem. Am. Math. Soc. **202**, 950 (2009).

[72] Villani, C. *Hypocoercive diffusion operators.* Proceedings of the International Congress of Mathematicians, Madrid, 2006.

[73] Yamakawa, H. *Tranport Properties of Polymer Chains in Dilute Solutions: Hydrodynamic Interactions.* J. Chem. Phys. **53**, 436 (1970), 436-443.

[74] Yau, H.T. *Relative entropy and hydrodynamics of Ginzburg-Landau models.* Lett. Math. Phys. **22**, 1 (1991), 63-80.