ABSTRACT

Title of Dissertation: PART I: ON THE STABILITY THRESHOLD OF COUETTE FLOW IN A UNIFORM MAGNETIC FIELD

PART II: QUANTITATIVE CONVERGENCE TO EQUILIBRIUM FOR HYPOELLIPTIC STOCHASTIC DIFFERENTIAL EQUATIONS WITH SMALL NOISE

Kyle Liss
Doctor of Philosophy, 2021

Dissertation Directed by: Professor Jacob Bedrossian
Department of Mathematics

This dissertation contains two parts. In Part I, we study the stability of the Couette flow $(y,0,0)^T$ in the presence of a uniform magnetic field $\alpha(\sigma,0,1)$ on $\mathbb{T}\times\mathbb{R}\times\mathbb{T}$ using the 3D incompressible magnetohydrodynamics (MHD) equations. We consider the inviscid, perfect conductor limit $\mathbf{Re}^{-1} = \mathbf{R}_m^{-1} \ll 1$ and prove that for strong and suitably oriented background fields the Couette flow is asymptotically stable to perturbations that are $\mathcal{O}(\mathbf{Re}^{-1})$ in the Sobolev space H^N. More precisely, we establish the decay estimates predicted by a linear stability analysis and show that the perturbations $u(t,x+yt,y,z)$ and $b(t,x+yt,y,z)$ remain $\mathcal{O}(\mathbf{Re}^{-1})$ in $H^{N'}$ for some $1 \ll N'(\sigma) < N$. In the Navier-Stokes case, high regularity control on the perturbation in a coordinate system adapted to the mixing of the Couette flow is known only under the stronger assumption of $\mathcal{O}(\mathbf{Re}^{-3/2})$

data [17]. The improvement in the MHD setting is possible because the magnetic field induces time oscillations that partially suppress the lift-up effect, which is the primary transient growth mechanism for the Navier-Stokes equations linearized around Couette flow.

In Part II, we study the convergence rate to equilibrium for a family of Markov semigroups $\{\mathcal{P}_t^\epsilon\}_{\epsilon>0}$ generated by a class of hypoelliptic stochastic differential equations on $\mathbb{R}^d$, including Galerkin truncations of the incompressible Navier-Stokes equations, Lorenz-96, and the shell model SABRA. In the regime of vanishing, balanced noise and dissipation, we obtain a sharp (in terms of scaling) quantitative estimate on the exponential convergence in terms of the small parameter ϵ. By scaling, this regime implies corresponding optimal results both for fixed dissipation and large noise limits or fixed noise and vanishing dissipation limits. As part of the proof, and of independent interest, we obtain uniform-in-ϵ upper and lower bounds on the density of the stationary measure. Upper bounds are obtained by a hypoelliptic Moser iteration, the lower bounds by a De Giorgi-type iteration (both uniform in ϵ). The spectral gap estimate on the semigroup is obtained by a weak Poincaré inequality argument combined with quantitative hypoelliptic regularization of the time-dependent problem.

PART I: ON THE STABILITY THRESHOLD OF COUETTE FLOW IN A UNIFORM MAGNETIC FIELD
PART II: QUANTITATIVE CONVERGENCE TO EQUILIBRIUM FOR HYPOELLIPTIC STOCHASTIC DIFFERENTIAL EQUATIONS WITH SMALL NOISE

by

Kyle Liss

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
2021

Advisory Committee:
Professor Jacob Bedrossian, Chair
Professor Sandra Cerrai
Professor Johan Larsson
Professor Charles Levermore
Professor Antoine Mellet

Acknowledgments

Here I take the opportunity to express my gratitude for the many people that helped me through graduate school and to become the person that I am today.

First and foremost I would like to thank my advisor, Jacob Bedrossian. I owe the successful completion of this dissertation largely to his guidance, support, and collaboration. He suggested to me the research direction pursued in Part I, co-authored the paper comprising Part II,[1] and offered countless insights and ideas throughout my mathematical journey to the present work. I am equally grateful for Jacob's contributions to my development and graduate school experience that extend beyond the research making up my dissertation. He taught me about the process of doing research and greatly impacted what kinds of questions I find interesting. Jacob's deep understanding of PDEs is something I admire and aspire to. Much of my intuition for PDEs and stochastics that I've built I can attribute to learning from him. It was a pleasure and honor to be his first PhD student. I know that I will be the first of many.

I also want to thank the members of my committee, Sandra Cerrai, Johan Larsson, Dave Levermore, and Antoine Mellet, for taking the time to read my dissertation, participate in my defense, and offer useful suggestions for future thinking. I am particularly grateful for Professor Levermore. Not only did he serve additionally on

[1]As such, Jacob made important mathematical contributions to this project. He played a particularly leading role in the Moser iterations of Chapters 10 and 11, and the proofs of Lemmas 11.1 and 6.3.

my preliminary oral exam committee, but he also organized the RIT on Applied PDE, in which I gave many talks and gained a lot of confidence. I appreciate the feedback he gave on my talks. Professor Levermore's philosophy on the key components of any good talk is something I will always remember. It is in my mind whenever preparing a presentation.

Next I want thank Professor Dionisios Margetis. Professor Margetis mentored me in an undergraduate research program during the summer after my junior year. He encouraged me throughout the entire experience and helped me believe that I was a person that could be successful in graduate school and beyond. Ultimately, he wrote me a recommendation letter that helped me get accepted to the University of Maryland. My time working with Professor Margetis came at an important crossroads in my life. If not for his support and the positive summer experience, I may very well have not pursued academia.

I am also grateful for guidance from a number of postdocs and elder graduate students: Jacky Chong, Sam Punshon-Smith, Siming He, Alex Blumenthal, Kasun Fernando, and others. Jacky was my graduate student mentor while I worked with Professor Margetis as an undergraduate. We frequently discussed for multiple hours at a time that summer and he continued to act as a mentor to me when I entered graduate school. I am also particularly grateful to Alex Blumenthal, with whom discussions about SPDEs and ergodic theory played an important role in my studies leading up to the research in Part II.

Thank you also to my colleagues and friends Nick, Stavros, and Sven. Our frequent discussions about both math and life enriched my graduate experience. Your friendship

and support mean a lot to me.

Finally, I want to express my gratitude and love for my parents, brother Zach, best friend John, and loving girlfriend Yasmin for always supporting me even in the most difficult of times. I would not be where I am today if not for your love and care.

Table of Contents

Chapter 1: Introduction

At the broadest level, this dissertation is motivated by a desire to understand the long-time behavior of fluids and plasmas in weakly dissipative regimes. There is a lengthy list of problems falling under this general umbrella whose rigorous mathematical analysis is of great importance. A select few include understanding the decay of electric field perturbations in collisionless plasmas [12,28,108], quantifying enhanced dissipation timescales in fluids caused by mixing [24, 125, 130–132], and describing the long-time statistical properties of weakly nonlinear, chaotic waves [29,49,109,129].

The present dissertation contains two parts, each studying a particular problem within the general class of questions introduced above. In Part I (Chapters 2-5, which recount the work from [93]) we consider the nonlinear stability of a magnetized shear flow in a conducting fluid using the equations of 3D magnetohydrodynamics. Our main result is a quantitative estimate (with respect to the small dissipation parameters) on the size of the equilibrium's basin of attraction. In Part II (Chapters 6-11, which recount the work from [20]) we turn away from regimes of stability and are motivated instead by *fully developed turbulence*. We prove here results on the convergence rate to statistical equilibrium and properties of the stationary measure for a fairly general stochastic differential equation (SDE) covering prototypical turbulent systems such as Lorenz-96 and Galerkin truncations of the forced Navier-Stokes equations. In the remainder of this chapter, we briefly introduce the two separate parts. For detailed introductions, see Chapters 2 and 6.

1.1 Part I: Stability of a magnetized shear flow

The central problem in hydrodynamic/hydromagnetic stability is to understand how equilibrium fluid and plasma configurations respond to small perturbations, a question with numerous applications to areas such as geophysics, aerospace engineering, solar physics, controlled fusion, and atmospheric/climate science. It is of particular interest to study how the onset of instabilities depends on magnetic backgrounds and dimensionless quantities, known as *Reynolds numbers*, that measure the relative importance of advection and diffusion. In fluid mechanics,[1] the Reynolds number inuitively quantifies the degree to which turbulent behavior is expected. It is defined as $\mathbf{Re} = LU/\nu$, where ν is the kinematic viscosity of the fluid and L and U are characteristic length and velocity scales. An overview of hydrodynamic/hydromagnetic stability theory can be found in the textbooks [35, 45, 128].

The classical approach to hydrodynamic stability pioneered in the nineteenth century focuses (naturally) on linear theory, i.e., the behavior of infinitesimal perturbations. While linearized analysis is a crucial part of understanding any given stability problem, it is an experimental fact that it often does not give the full story. Indeed, there are flows that are linearly stable at all Reynolds number, but experimentally unstable at sufficiently high Reynolds number [Section 2.3; [45]], and flows that

[1] For plasmas, in addition to $\mathbf{Re}$, there is the magnetic Reynolds number $\mathbf{Re}_{\mathrm{m}}$.

transition to turbulence at Reynolds numbers much lower than what is predicted by the linear theory [3, 128]. Collectively, the phenomenon wherein a linearly stable flow is nevertheless experimentally unstable and transitions to turbulence at sufficiently high Reynolds number is known as *subcritical transition*. When subcritical instabilities drive the route to turbulence, a more sophisticated analysis accounting for nonlinear effects is needed. This challenge has been faced by a number of works over the past decade for the planar *Couette flow* $\mathbf{u} = (y, 0, 0)$ (and zero pressure gradient), which is perhaps the simplest nontrivial stationary solution to the Navier-Stokes equations and moreover observed to undergo subcritical transition. In particular, [16, 17, 22, 23, 81, 102, 124] proved nonlinear asymptotic stability of Couette flow in the Euler equations or in a quantitative sense in the limit of zero viscosity, and [19, 40, 92] analyze the sharpness of certain results by constructing solutions that exhibit transient growth or do not decay. See Section 2.1 for a list of additional related references and a more detailed description of previous results on Couette flow in the Navier-Stokes/Euler equations. See also the review article [18].

In Part I of this dissertation, we extend the rigorous study of Couette flow in the infinite Reynolds number limit to the setting of a plasma surrounded by a strong, uniform background magnetic field. Our main result (Theorem 3.3) shows in the mathematical setting of [17, 124] that there is a precise sense in which the background field has a stabilizing effect on the flow. This contributes to the various examples in the literature on how a magnetic field can stabilize a plasma (see e.g. [33,34,42,43,74,85,117,120]). At the level of the physics, the main step in the proof is a careful linear analysis to understand how the wavelike transport created by the magnetic field interacts with both the linear stabilizing and destabilizing mechanisms of the shear. At the more technical level, the proof requires adapting the Fourier side energy method techniques developed for shear flows (see e.g. [17]) to our particular setting.

1.2 Part II: Convergence to equilibrium for a class of hypoelliptic SDEs

While understanding perturbative regimes of equilibrium flow configurations is of fundamental importance, flows in nature more commonly exist in *turbulent* states. Turbulence seems quite difficult to define precisely, but experiments and life experience suggest that the following properties are fundamental: (A) chaotic dynamics in the sense of extreme sensitivity with respect to initial conditions, (B) unique and reproducible long-time statistical properties (in damped-driven settings), and (C) inherently infinite dimensional behavior, namely nonlinear cascades that send information to smaller and smaller length scales in the limit of infinite Reynolds number. Due to (A) and (B), predicting the details of a single experiment is both impossible and unnecessary. Instead, what is meaningful to study are statistical properties defined through averaging procedures. In fact, experiments show that many statistical properties of turbulent systems are universal in that they do not depend on the particular experimental setup (see e.g. [55, 109]).

A natural mathematical framework for studying turbulence is to introduce a white-in-time random forcing into the fluid equations and use techniques from stochastic analysis. In this setting, there has been a significant effort put forward to understand (A)-(C) at a mathematically rigorous level. The most satisfactory results are concerned with (B). Indeed, unique and geometric ergodicity are well-understood for Markov processes generated by a variety of stochastically forced systems (see e.g. [20, 53, 54, 60, 62, 69, 71, 72, 126]). Far less is known about (A) and (C); however recent progress on chaos has been made in simplified settings such as passive scalar turbulence [13] and finite dimensional models [15]. Moreover, various conditional theorems on flux laws in hydrodynamics are known [25, 26, 112] and Batchelor's law of passive scalar turbulence was recently proven [14].

In spite of the progress discussed above, many fundamental questions remain open. A particularly important example, and the subject of part II of this dissertation, is the quantitative understanding of (B). While tools for proving unique and geometric ergodicity are fairly well understood, estimating the rate of convergence to statistical equilibrium in the limit $\mathbf{Re} \to \infty$ is much more difficult. For infinite dimensional models, the current estimates on the convergence rate are far from optimal with respect to scaling in **Re**, and improving them seems currently out of reach. Interestingly, the problem of quantitative ergodic properties is quite difficult even in finite dimensions. In fact, until the recent work [20] by the author and his advisor, a near-optimal (with respect to scaling in **Re**) estimate on the convergence rate to equilibrium for the stochastically forced Galerkin Navier-Stokes equations was unknown.

The second part of this dissertation details the work from [20]. More specifically, in Part II we consider a broad class of hypoelliptic SDEs on $\mathbb{R}^d$ containing not only Galerkin truncations of the Navier-Stokes equations but also other prototypical turbulent systems such as Lorenz-96 and the shell model SABRA. The term *hypoelliptic* refers to how the stochastic forcing only acts directly on a few low modes, but can spread to all modes through the nonlinearity. This is the main physical situation of interest, since in turbulence one usually thinks of the external forcing acting only at the large length scales [55, 109]. We consider the regime of vanishing, balanced forcing and dissipation described by the parameter $\epsilon \to 0$ (which should be thought of as an inverse Reynolds number). Our main result (Theorem 6.5) is an optimal estimate, with respect to scaling in $0 < \epsilon \ll 1$, on the convergence of transition probabilities to the system's unique invariant measure. As mentioned above, this greatly improves upon the previously known bounds in the literature. We also prove a quantitative probabilistic smoothing estimate (Lemma 6.3) and some uniform-in-ϵ pointwise estimates on the stationary density (Theorem 6.2). These results were mainly obtained as lemmas needed to prove Theorem 6.5, but we believe that they are also of independent interest.

Chapter 2: Introduction to Part I

In Part I of this dissertation we consider the 3D magnetohydrodynamic (MHD) equations, which are a widely applicable model for conducting fluids, e.g. plasmas and liquid metals. The MHD equations couple a reduced form of Maxwell's equations of electromagnetism with the Navier-Stokes equations of fluid mechanics. We will consider the situation where the fluid is incompressible with a constant density (normalized to unity). In this case, the equations read

$$\begin{cases} \partial_t \tilde{b} = -\nabla \times E, \\ \nabla \times \tilde{b} = \mu_0 J, \\ J = \sigma(E + \tilde{u} \times \tilde{b}), \\ \partial_t \tilde{u} + \tilde{u} \cdot \nabla \tilde{u} + \nabla \tilde{p} = \nu \Delta \tilde{u} + J \times \tilde{b}, \\ \nabla \cdot \tilde{u} = \nabla \cdot \tilde{b} = 0. \end{cases} \tag{2.1}$$

Here, $\tilde{u}$ and $\tilde{p}$ are the fluid velocity field and pressure, and J, $\tilde{b}$, and E denote respectively the current density, magnetic field, and electric field. The pressure is a scalar, while $\tilde{u}$, J, $\tilde{b}$, and E are vector fields valued in $\mathbb{R}^3$. The physical parameters are the kinematic viscosity ν, the conductivity σ, and the vacuum permeability μ_0. The first two equations in (2.1) are Faraday's law and Ampere's law, respectively. Together with $\nabla \cdot \tilde{b} = 0$, they are a simplified version of Maxwell's equations that neglects the displacement current in Ampere's law and the presence of any net charge density. The third equation is Ohm's law, which describes how currents form in response to the forces on free charges, namely the Lorentz force and the electric force from the induced field in Faraday's law. The penultimate equation in (2.1) and the incompressibility condition $\nabla \cdot \tilde{u} = 0$ are the Navier-Stokes equations with the volumetric Lorentz force $J \times \tilde{b}$. Physically, the evolution equation for $\tilde{u}$ describes conservation of momentum in the fluid. For a discussion of the applicability of (2.1) and an introduction to plasma physics and MHD, see the textbooks [28, 38].

One typically expresses (2.1) in terms of $\tilde{u}$ and $\tilde{b}$ alone. Using Ohm's law and Ampere's law to eliminate J and E from the system and then applying vector calculus identites yields

$$\begin{cases} \partial_t \tilde{u} + \tilde{u} \cdot \nabla \tilde{u} - \tilde{b} \cdot \nabla \tilde{b} = -\nabla \tilde{p} + \nu \Delta \tilde{u}, \\ \partial_t \tilde{b} + \tilde{u} \cdot \nabla \tilde{b} - \tilde{b} \cdot \nabla \tilde{u} = \mu \Delta \tilde{b}, \\ \nabla \cdot \tilde{u} = \nabla \cdot \tilde{b} = 0. \end{cases} \tag{2.2}$$

Here, we have defined the magnetic diffusivity $\mu = (\sigma \mu_0)^{-1}$. Assuming that the characteristic velocity and length scales in (2.2) have already been normalized to unity, $\nu = \mathbf{Re}^{-1}$ is the inverse Reynolds number and $\mu = \mathbf{R}_{\mathrm{m}}^{-1}$ is the inverse magnetic Reynolds number. In general, (2.2) must be supplemented with appropriate boundary conditions. In the present dissertation we will only consider the initial value problem for (2.2) on the boundary-less domain $\mathbb{T} \times \mathbb{R} \times \mathbb{T}$, where $\mathbb{T}$ is the periodized interval $[0, 2\pi]$. Our notation is $(t, x, y, z) \in \mathbb{R}^+ \times \mathbb{T} \times \mathbb{R} \times \mathbb{T}$.

It is of general importance to understand the stability of equilibrium solutions to (2.2), and in particular how the stability of a given equilibrium depends on the Reynolds numbers. Perhaps the simplest stationary solution with a nonzero velocity field is the Couette flow $u_s = (y, 0, 0)^T$ in any uniform magnetic field $b_s = \alpha(\sigma, 0, 1)^T$. Here, α and σ are real numbers that without loss of generality we take to be positive (henceforth, σ will never denote conductivity). Analyzing the stability of (u_s, u_s) in the limit **Re**, $\mathbf{Re}_\mathrm{m} \to \infty$ serves as a model problem for understanding shear flow stability in magnetized plasmas, an area that has received a lot of attention in the past [33–35, 42, 78, 85, 95, 106, 117, 120]. To investigate the stability of (u_s, b_s) we introduce the perturbations u and b defined by $\tilde{u} = u + u_s$ and $\tilde{b} = b + b_s$. They satisfy the system (summation over repeated indices is implied)

$$\begin{cases} \partial_t u + u \cdot \nabla u - b \cdot \nabla b + y\partial_x u - \alpha \partial_\sigma b + \begin{pmatrix} u^2 \\ 0 \\ 0 \end{pmatrix} = -\nabla p^{\text{NL}} + 2\nabla \Delta^{-1} \partial_x u^2 + \nu \Delta u, \\ \partial_t b + u \cdot \nabla b - b \cdot \nabla u + y\partial_x b - \alpha \partial_\sigma u - \begin{pmatrix} b^2 \\ 0 \\ 0 \end{pmatrix} = \mu \Delta b, \\ p^{\text{NL}} = (-\Delta)^{-1}(\partial_j u^i \partial_i u^j - \partial_j b^i \partial_i b^j), \\ \nabla \cdot u = \nabla \cdot b = 0, \\ u\big|_{t=0} = u_{\text{in}}(x, y, z), \quad b\big|_{t=0} = b_{\text{in}}(x, y, z), \end{cases} \tag{2.3}$$

where we have introduced the notation $u = (u^1, u^2, u^3)^T$, $b = (b^1, b^2, b^3)^T$, and $\partial_\sigma = \sigma\partial_x + \partial_z$.

It is an interesting question in its own right how to formulate the nonlinear stability problem for (2.3). One possible formulation is motivated by the phenomenon in 3D hydrodynamics known as *subcritical transition*. This refers to when a linearly stable flow (see [18,45] for precise definitions) is nevertheless experimentally unstable at sufficiently high Reynolds number. Perhaps the most famous example is flow through a pipe, studied by Reynolds in his original experiments. He found that laminar pipe flow becomes turbulent at sufficiently high Reynolds number, and yet numerical calculations suggest that the linearized system is spectrally stable for any Reynolds number [45]. Distinct from this example are flows (e.g., plane Poiseuille flow) that are linearly unstable for high enough Reynolds number, but typically transition to turbulence in experiments well below the critical Reynolds number predicted by the linear theory [31,128]. An idea dating back to Kelvin [83] to reconcile these instances of linear stability and experimental instability is that while a given flow might be nonlinearly stable for any fixed Reynolds number, its basin of attraction could be shrinking as $\nu \to 0$. The equilibrium is then unstable in practice at sufficiently high Reynolds number due to the inevitable presence of finite amplitude perturbations in experiments.

Regarding the Couette flow on $\mathbb{T} \times \mathbb{R} \times \mathbb{T}$ one can show that both (2.3), and the corresponding Navier-Stokes system obtained by setting $\alpha = b = 0$ are asymptotically linearly stable in an appropriate sense for any $\sigma, \alpha \geq 0$ and $\mu, \nu > 0$. For precise statements, see [Proposition 1.2; [17]] and Proposition 3.2 below; see also [chapter 8, [45]] and [78, 106] for classical results on the linear stability of parallel shear flows in the Navier-Stokes and MHD equations. Despite being linearly stable, Couette flow is experimentally unstable, undergoing subcritical transition at sufficiently high Reynolds number (see e.g. [115, 118] and the references therein). Kelvin's suggestion above is particular sensible for Couette flow. Indeed, as we will discuss below in Chapter 3, a transient growth mechanism in general amplifies solutions by $\min(\nu, \mu)^{-p}$ for some $p > 0$ up until a suitable dissipation timescale, and hence unless the initial data is small with respect to μ and ν it should be expected that the linear approximation eventually breaks down, opening up the possibility to transition into turbulence. Transient growth mechanisms can lead to transition in magnetized shears as well (see e.g. [42,85]). In light of the discussion above, a natural nonlinear stability problem for (2.2) is the following

[18]:

Problem 1. *Let $0 < \mu, \nu \ll 1$ and $\alpha \geq 1$. Given an initial norm X_i and a final norm X_f, determine the smallest $\beta(X_i, X_f), \gamma(X_i, X_f) \geq 0$ such that if the initial perturbations u_{in} and b_{in} satisfy*

$$\mu^{-\beta}\|b_{in}\|_{X_i} + \nu^{-\gamma}\|u_{in}\|_{X_i} \leq c_0$$

for c_0 sufficiently small (and independent of μ and ν), then the solution is global in time, does not transition away from (u_s, b_s), and converges back to (u_s, b_s) as $t \to \infty$ in the sense that

$$\|(u, b)\|_{L^\infty X_f} \lesssim c_0, \quad \lim_{t\to\infty} \|(u(t), b(t))\|_{X_f} = 0.$$

In the hydrodynamics literature, the number $\gamma \geq 0$ is referred to as the *transition threshold*. It is not known a priori that the basin of attraction necessarily shrinks as a power law, but (at least in the hydrodynamic case) this is what tends to be observed numerically [46, 98, 115]. Note moreover that studies of Couette flow in the 2D [22, 23, 102] and 3D [16, 17, 19, 124] Navier-Stokes equations suggest that γ might depend in a complicated way on the norms X_i and X_f.

The transient growth mechanism referred to above which is most responsible for the subcritical transition of 3D Couette flow in the Navier-Stokes equations is the *lift-up effect* (first observed in [50]). It arises from the linear u^2 term in the equation for u^1, and causes linear in time growth of u^1 before the dissipation timescale $t \sim \nu^{-1}$ (and arbitrarily large growth in the Euler equations). Other growth mechanisms that play a role in the stability analysis include an algebraic growth in time of derivatives caused by the mixing, and an amplification of streamwise dependent modes due to a transient unmixing of high frequency information to large scales. This latter effect was first noticed by Orr in [111] and is known as the *Orr mechanism*. One needs to contend with these same effects to study the MHD shear problem. There is a crucial difference, however, in that the presence of a background magnetic field partially suppresses the lift-up effect. This observation plays a fundamental role in the proof of our main result and is discussed in Section 3.1.

Our goal in this part of the disseration is to contribute to the study of Problem 1 in the case where X_i and X_f are Sobolev spaces adapted to the linear dynamics and in the special case where the ideal limit is taken with $\mu = \nu$. Our main result shows that due to the magnetic field's influence on the lift-up instability there is a precise mathematical sense in which a strong background magnetic field has a stabilizing effect on the Couette flow. The fact that a large background field can have a stabilizing effect on a conducting fluid is classically known in the literature. For example, the works [33, 34, 120] predict that a sufficiently large magnetic field can delay the onset of Taylor vortices in the Taylor-Couette flow, a fact that has been observed experimentally [43]. Morever, it is known that a magnetic field parallel to the main flow in a free shear layer can have a stabilizing influence on the Kelvin-Helmholtz stability; see e.g. [94] and the references therein. Regarding rigorous mathematical results, He, Xu, and Wu proved in [74] that if one takes the stationary solution $u_s = (0,0,0)$, $b_s = (0,0,1)$ to the 3D, *ideal* MHD equations and introduces a smooth perturbation that is sufficiently small and well localized, then the resulting solution is smooth and global in time. Such a global-in-time result for the 3D Euler equations of course famously remains open. As a final example, and perhaps most relevant to our present study, it has been observed numerically that a magnetic field can suppress subcritical transition in planar shear flows [42, 85]. In fact, the work [85] has already observed that a spanwise magnetic field (the z-direction in our setup) can suppress the lift-up transient growth (see Section 2.3 for more discussion).

Part I is organized as follows. In the remainder of this chapter, we review known results for the stability of Couette flow in the Navier-Stokes/Euler equations, summarize our main result, and explain its relationship with the existing literature and some key aspects of the proof. In Chapter 3 we discuss the linear effects and heuristics that form the basis of how to treat the fully nonlinear problem. At the end of Chapter 3 we also state precisely our main nonlinear stability result. In Chapter 4 set up preliminary aspects of the proof. Finally, in Chapter 5 we carry out the requisite energy estimates.

2.1 Known results for the Navier-Stokes and Euler equations

There is a substantial body of mathematical results on the analog of Problem 1 for the Couette flow in the 3D Navier-Stokes equations. The works most related to our present study are [17] and [124], which prove, in distinct senses, that $\gamma \leq 3/2$ and $\gamma \leq 1$, respectively, when the domain is $\mathbb{T} \times \mathbb{R} \times \mathbb{T}$ and X_i and X_f are Sobolev spaces. More specifically, the result in [17] shows that if $\|u_{\text{in}}\|_{H^s} \ll \nu^{3/2}$ for any $s > 9/2$ then there holds

$$\|U\|_{L^\infty H^{s-2}} = \mathcal{O}(\nu^{1/2}), \quad \lim_{t\to\infty} \|U_{\neq}\|_{H^{s-2}} = 0,$$

where the subscript $\neq$ denotes the projection onto nonzero frequencies in x (see Section 2.4 below) and $U(t, x, y, z) = u(t, x + yt + t\psi(t, y, z), y - \psi(t, y, z), z)$ for a solution-dependent function ψ that remains $\mathcal{O}(\nu^{1/2})$ in H^s. The leading order effect of the coordinate transformation is to unwind by the mixing induced by the Couette flow, which amounts to modding out by the main component of the linear evolution. We thus refer to U, borrowing terminology from dispersive PDE, as the *profile*. High regularity control on the profile gives quantitative information on the dynamics. For example, one can deduce that the mixing effect that characterizes the linear behavior persists as the leading order effect at the nonlinear level. Hence, the result in [17] shows that for sufficiently regular initial data the solution looks essentially linear. On the other hand, the authors in [124] consider relatively low regularity (H^2 on the velocity variables) and prove that $\gamma \leq 1$ when the derivatives are measured in the original coordinates. This result partially improves those in [17] due to the weaker assumption on the initial data. It might be possible to extend the methods in [124] to obtain high regularity profile estimates for $\mathcal{O}(\nu)$ Sobolev data and thereby obtain a strict improvement on the results of [17], however, $\gamma \leq 1$ in the sense of profile estimates on $\mathbb{T} \times \mathbb{R} \times \mathbb{T}$ is currently only known for infinite regularity perturbations lying in a Gevrey space [19]. In fact, for Gevrey data it is possible to partially follow the lift-up instability, and hence even more precisely characterize the nonlinear dynamics [16, 19]. Here we take the approach of [16, 17, 19] and prove profile estimates.

The known stability results in 2D are stronger because the lift-up effect is eliminated. In particular, for the Navier-Stokes equations on $\mathbb{T} \times \mathbb{R}$ it is known that $\gamma \leq 1/3$ for Sobolev data [102] (see also the earlier and simpler work [23] that proves $\gamma \leq 1/2$) and $\gamma = 0$ for Gevrey data [22]. These papers improve the analogous 3D results in [17, 124] (Sobolev) and [16, 19] (Gevrey). Asymptotic stability of the Couette flow is also known in the 2D Euler equations provided that the initial perturbations are small in a suitable Gevrey space; see [21] and [81], which treat the cases of $\mathbb{T} \times \mathbb{R}$ and $\mathbb{T} \times [0, 1]$, respectively. Interestingly, it is known that the theorem in [21] does not extend to Gevrey spaces weaker than those that it originally considered. This is a consequence of the recent work [40], and leads to the conclusion that for Couette flow in the 2D Euler equations the dynamics of perturbations depend importantly on their regularity.

We should mention here that the mathmatics literature on the stability of shear flows extends far beyond Couette flow in the Navier-Stokes or MHD equations. It includes the analysis of both other shears, for instance the Poisueille and Kolmogorov flows, and other systems, such as the Boussinesq equations. There has also been significant work on the stabilizing effects of general background shears in linear advection diffusion equations. For a small subset of the literature, see for example [24, 39, 125, 131, 132] and the references therein.

2.2 Summary of main result

We defer the complete statement of our main result until the end of Chapter 3, after we have discussed in detail the linearization of (2.3). However, it can be summarized as follows.

Theorem 1 (Summary)**.** *Let $\mu = \nu \in (0, 1]$ and $\sigma > 0$ be an irrational number that satisfies a generic Diophantine condition (see (3.40) and (3.12) for precise statements). For α and N sufficiently large (depending only on σ) there exists a constant $c_0(N) \ll 1$*

so that if

$$\|u_{in}\|_{H^{N+2}} + \|b_{in}\|_{H^{N+2}} \leq c_0 \nu$$

then the solution to (2.3) *is global in time and the linear profile* (U, B) *defined by*

$$(U(t,x,y,z), B(t,x,y,z)) = (u(t,x+yt,y,z), b(t,x+yt,y,z))$$

satisfies

$$\|U\|_{L^\infty H^N} + \|B\|_{L^\infty H^N} \lesssim \nu^{2/3}.$$

In other words, we have $\gamma(X_i = H^N, X_f = H^{N-2}) \leq 1$ *for* X_f *measuring derivatives on the linear profile.*

Remark 1. While taking $\mu = \nu$ is mathematically natural, it is usually not the case for real physical applications. Thus, it is of interest to consider $\mu \neq \nu$ and the more general double limit stated in Problem 1. However, even for the simpler stationary solution $u_s = (0,0,0)$, $b_s = (0,0,1)$, studying the 3D MHD equations with $\mu \neq \nu$ involves substantial mathematical difficulties (see [123] and the references therein), and our current proof makes heavy use of the $\mu = \nu$ structure.

2.3 Brief discussion of results and ideas of the proof

Theorem 1 shows that a sufficiently strong magnetic field with an appropriate irrational tilt has a stabilizing effect on the Couette flow. Indeed, our result establishes high regularity profile estimates for Sobolev space perturbations on $\mathbb{T} \times \mathbb{R} \times \mathbb{T}$ akin to those in [17], but under the weaker assumption of $\mathcal{O}(\nu)$ initial data. Moreover, while (2.3) is more complicated than the Navier-Stokes equations, our proof in many respects is much simpler than those in [17, 124]. Most notably, our proof does not require a solution-dependent nonlinear change of coordinates. We are also able to treat the various estimates more generally because without the lift-up effect u^1 and u^3 behave essentially the same.

As mentioned above, the magnetic field's stabilizing effect comes mostly from partially suppressing the lift-up effect. There is a physical heuristic to explain this. The lift-up effect in the Navier-Stokes equations occurs as the fluid circulates in planes normal to the direction of the streamwise flow, which redistributes the mean streamwise velocity and can drastically alter the shear profile [50]. When the fluid is electrically conducting and a sufficiently strong magnetic field is present in the spanwise direction (the z-direction for our setup), the field lines provide a restoring force via the frozen-in-law that resists the rotation of fluid layers. Thus, instead of growth, in the MHD setting oscillations occur and are transmitted in the form of Alfvén waves. As mentioned earlier, the fact that a spanwise magnetic field can weaken the lift-up instability due to its influence on modes that depend on the spanwise variable has appeared previously in the literature [85]. For a detailed mathematical discussion of the linear stabilization for our present setting, see Section 3.1. Note that a physical heuristic similar to the one above can also be used to understand (in part, at least) the suppression of Taylor vortices by a co-axial magnetic field (see [Section 7, [33]]).

In the proof of Theorem 1 we capture the oscillations induced by the magnetic field by integrating by parts in time using the basic identity

$$e^{i\omega(\mathbf{k})t} = \frac{1}{i\omega(\mathbf{k})}\partial_t e^{i\omega(\mathbf{k})t}. \tag{2.4}$$

It turns out that for a Fourier mode with $\mathbf{k} = (k, \eta, \ell) \in \mathbb{Z} \times \mathbb{R} \times \mathbb{Z}$ the oscillation frequency behaves like $\omega(\mathbf{k}) \approx |\sigma k + \ell|$. One of the main difficulties in the proof is finding a way to utilize the oscillations for modes with $|\sigma k + \ell| \approx 0$, i.e., frequencies with wave vectors approximately perpendicular to the background magnetic field. This challenge underlies our idea to choose σ irrational, as it provides a kind of non-resonance condition that ensures $|\sigma k + \ell| \neq 0$ for all $(k, \ell) \in \mathbb{Z} \times \mathbb{Z}$. By Dirichlet's approximation theorem, for any choice of $\sigma \in \mathbb{R} \setminus \mathbb{Q}$ there exist infinitely many $(k, \ell) \in \mathbb{Z} \times \mathbb{Z}$ such that $|\sigma k+\ell| \leq |k|^{-1}$, and so we still incur losses when $\omega(\mathbf{k}) \to 0$ as $|k| \to \infty$. However, using results in Diophantine approximation we can absorb such losses by paying regularity. For more details, we refer to Sections 3.2 and 3.4 below.

Our proof also relies on the same stabilizing effects of the Couette flow utilized in the works [16,17,19,22,23,124] on the Navier-Stokes equations. In particular, the mixing induced by the Couette flow results in an improved dissipation timescale (with respect to the heat equation) for the x-dependent modes. This is referred to as *enhanced dissipation*. A second stabilizing mechanism is the *inviscid damping*, which was first discovered by Orr [111] and causes decay on a timescale uniform in ν. The linear analysis for the 3D Navier-Stokes equations predicts an inviscid damping timescale of $\langle t\rangle^{-2}$, while in the MHD setting we have the difficulty that this is slowed to $\langle t\rangle^{-1}$ and is significantly harder to access (see Section 3.3). To exploit the stabilizing properties at the nonlinear level we use the Fourier multiplier methods employed in [16, 17, 19, 22, 23]. In this respect we follow most closely the ideas of [17].

A few natural questions arise from Theorem 1. First, it is true for the linearization of (2.3) that the high norms are at worst amplified by a factor $\nu^{-1/3}$ in the sense that for any $s \geq 0$ there holds

$$\sup_{t\geq 0} \|(u(t, x+yt, y, z), b(t, x+yt, y, z))\|_{H^s} \lesssim \nu^{-1/3}\|(u_{in}, b_{in})\|_{H^{s+2}},$$

which suggests that the threshold estimate in Theorem 1 is far from optimal. It is of great interest to consider if there is a sense in which $\gamma < 1$ for (2.3). This would be a significant theorem in that the lift-up effect implies that no analogous result is possible for the Navier-Stokes equations. The $\gamma \leq 1$ estimate in our main result reflects a treatment of the nonlinear terms that does not carefully analyze the possible growth mechanisms and instead relies entirely on the dissipation to absorb the loss of derivative. A more precise treatment of the nonlinearities using the techniques in [16, 19, 21, 22, 81] may yield $\gamma < 1$ for perturbations small in a suitable Gevrey space. In fact, given the recent work [102], a detailed analysis of the nonlinearity could possibly yield $\gamma < 1$ even in a Sobolev topology. Due to the result in [124], it is also reasonable to ask if the Sobolev threshold improves if one considers only low regularity and derivatives measured in the original coordinates. On the face of it, however, this seems unlikely because utilizing the inviscid damping in the MHD setting costs more regularity than in the Navier-Stokes case, and moreover the amount of regularity depends on the choice of σ (see Sec. 3.4). Studying the MHD problem in high regularity may be the most natural. The next follow up question to Theorem 1 is to determine whether or not the threshold estimate $\gamma \leq 1$ holds in the case that σ is either rational or violates (3.40). Currently, for general $\sigma \in \mathbb{R}$ we can prove only that $\gamma \leq 4/3$ (see Corollary 3.4). Lastly, as mentioned above, it is of interest to consider the physical case $\mu \neq \nu$ and the general double scaling limit suggested in Problem 1.

2.4 Notations and conventions for Part I

Given a vector $v = (v_j)_{j=1}^n$, $v_j \in \mathbb{C}$, we write $|v|$ or $|v_1, \ldots, v_n|$ to denote the ℓ^1 norm. For $a \in \mathbb{R}$ we use the standard notation $\langle a \rangle = \sqrt{1 + a^2}$. For two quantities a and b we write $a \lesssim b$ to mean that there exists $C \geq 0$ such that $a \leq Cb$. The constant C may depend on N (as defined in Theorem 3.3), but is always independent of ν, α, t_1 and t_2 (both t_1 and t_2 are defined in Section 4.4). Sometimes we will write $a \lesssim_\beta b$ if we want to emphasize that the implicit constant depends on some parameter β. All unlabeled integrals are assumed to be taken over $(x, y, z) \in \mathbb{T} \times \mathbb{R} \times \mathbb{T}$ and we use the shorthand notation $dV = dx\, dy\, dz$. For two functions f and g and a norm $\|\cdot\|_X$ we write

$$\|(f, g)\|_X = \sqrt{\|f\|_X^2 + \|g\|_X^2}.$$

Unless specified otherwise, in the rest of this section f and g denote functions from $\mathbb{T} \times \mathbb{R} \times \mathbb{T}$ into $\mathbb{R}^n$ for some $n \in \mathbb{N}$. We define the Fourier transform $\hat{f} : \mathbb{Z} \times \mathbb{R} \times \mathbb{Z} \to \mathbb{C}^n$ by

$$\mathcal{F}(f) = \hat{f}(k, \eta, \ell) = \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{T} \times \mathbb{R} \times \mathbb{T}} e^{-i(kx + \eta y + \ell z)} f(x, y, z) dV.$$

The function f is then recovered via the Fourier inversion formula

$$f(x,y,z) = \frac{1}{(2\pi)^{3/2}} \sum_{k\in\mathbb{Z}} \int_{\eta\in\mathbb{R}} \sum_{\ell\in\mathbb{Z}} e^{i(kx+\eta y+\ell z)} \hat{f}(k,\eta,\ell) d\eta.$$

We denote the projection of f onto the zero frequencies in x by

$$f_0 = \int_{\mathbb{T}} f(x,y,z)dx.$$

Then, we write

$$f_{\neq} = f - f_0$$

for the projection onto the nonzero frequencies in x. At times it will also be convenient to project onto the nonzero frequencies in z. For this, we use the alternate notation

$$P_{\ell\neq 0} f = f - \int_{\mathbb{T}} f(x,y,z)dz.$$

To avoid conflicting with the subscript notation just defined, when f is vector valued we use a superscript to denote the components. For example, if f is valued in $\mathbb{R}^3$ then we write $f = (f^1, f^2, f^3)$.

For a general Fourier multiplier with symbol $m(k,\eta,\ell)$ we write mf to denote $\mathcal{F}^{-1}(m(k,\eta,\ell)\hat{f})$. Also, when T is an operator represented by the multiplier $m_T : \mathbb{Z} \times \mathbb{R} \times \mathbb{Z} \to \text{Codomain}(m_T)$ and $h : \text{Codomain}(m_T) \to \mathbb{C}$ we write $h(T)$ for the operator with symbol $h(m_T(k,\eta,\ell))$. For example,

$$|\nabla| \text{ has symbol } m(k,\eta,\ell) = |k| + |\eta| + |\ell|,$$
$$\partial_\sigma^{-1} \text{ has symbol } m(k,\eta,\ell) = \frac{1}{i(\sigma k + \ell)}.$$

For simplicity we drop the ℓ^1 norm absolute value sign in $\langle|\nabla|\rangle$. That is, we write $\langle\nabla\rangle^s$ for the operator with symbol

$$m(k,\eta,\ell) = [1 + (|k| + |\eta| + |\ell|)^2]^{s/2}, \quad s \geq 0.$$

Since σ is fixed in the proof, we also use for any $a \in \mathbb{R}$ the shorthand notation

$$T_a^t = e^{a(\sigma\partial_x + \partial_z)t}$$

to denote the multiplier with symbol $e^{ia(\sigma k+\ell)t}$. We then write $\partial_t T_a^t$ to denote the Fourier multiplier with symbol $ia(\sigma k + \ell)e^{ia(\sigma k+\ell)t}$.

For $s \geq 0$ we define the Sobolev space H^s using the norm

$$\|f\|_{H^s} := \| \langle\nabla\rangle^s f\|_{L^2},$$

and we write the associated inner product as

$$\langle f, g\rangle_{H^s} = \int \langle\nabla\rangle^s f \cdot \langle\nabla\rangle^s g \, dV.$$

For functions $f(t,x,y,z)$ of space and time defined on the time interval (a,b) we define the Banach space $L^p(a,b;H^s)$ for $1 \leq p \leq \infty$ by the norm

$$\|f\|_{L^p(a,b;H^s)} = \|\|f\|_{H^s}\|_{L^p(a,b)}.$$

When the time interval is clear from context or mentioned explicitly elsewhere we use the shorthand notation $\|f\|_{L^p(a,b;H^s)} = \|f\|_{L^pH^s}$. For a Banach space X with norm $\|\cdot\|_X$ and a time interval $[a,b]$ we write $C([a,b];X)$ for the Banach space of continuous functions $h:[a,b] \to X$ equipped with the norm

$$\|h\|_{C([a,b];X)} = \sup_{a\leq t\leq b} \|h(t)\|_X.$$

Chapter 3: Linear Theory

Understanding the linearization of (2.3) is in some sense the key step in the proof of Theorem 1. Indeed, any result that estimates the stability threshold is ultimately considering a regime where the solution looks approximately linear. In this chapter we discuss the three linear effects that are crucial in the upcoming nonlinear analysis: the suppression of the lift-up effect due to the magnetic field, inviscid damping, and enhanced dissipation. When studying the inviscid damping, we discuss how to quantify the losses from integration by parts in time using results in Diophantine approximation.

The linearization of (2.3) reads

$$\begin{cases} \partial_t u + y\partial_x u - \alpha\partial_\sigma b + \begin{pmatrix} u^2 \\ 0 \\ 0 \end{pmatrix} = 2\nabla\Delta^{-1}\partial_x u^2 + \nu\Delta u, \\ \partial_t b + y\partial_x b - \alpha\partial_\sigma u - \begin{pmatrix} b^2 \\ 0 \\ 0 \end{pmatrix} = \nu\Delta b. \end{cases} \tag{3.1}$$

Since solutions to $\partial_t f + y\partial_x f = 0$ in general have derivative growth like $\|f(t)\|_{H^s} \gtrsim$

$\langle t \rangle^s \|f_{in}\|_{H^s}$, we must make a coordinate transform that unwinds by the mixing of the Couette flow if we want to obtain global in time estimates for (3.1). We introduce

$$\begin{aligned} X &= x - yt, \\ Y &= y, \\ Z &= z. \end{aligned} \tag{3.2}$$

Denoting $B(t, X, Y, Z) = b(t, x, y, z)$ and $U(t, X, Y, Z) = u(t, x, y, z)$, (3.1) then becomes

$$\begin{cases} \partial_t U - \alpha \partial_\sigma B + \begin{pmatrix} U^2 \\ 0 \\ 0 \end{pmatrix} = 2\nabla_L \Delta_L^{-1} \partial_X U^2 + \nu \Delta_L U, \\ \partial_t B - \alpha \partial_\sigma U - \begin{pmatrix} B^2 \\ 0 \\ 0 \end{pmatrix} = \nu \Delta_L B, \end{cases} \tag{3.3}$$

where $\nabla_L = (\partial_X^L, \partial_Y^L, \partial_Z^L) = (\partial_X, \partial_Y - t\partial_X, \partial_Z)$, $\Delta_L = \nabla_L \cdot \nabla_L$, and it is understood that $\partial_\sigma = \sigma \partial_X + \partial_Z$ when acting on functions in the new coordinates. In general, for a function $g(t, x, y, z)$ we will denote $G(t, X, Y, Z) = g(t, x, y, z)$.

3.1 Lift-up effect

We first review the lift-up effect for the Navier-Stokes equations. Recall that we write $f_0 = \int_{\mathbb{T}\times\mathbb{R}\times\mathbb{T}} f(x,y,z)dx$ for the x-average of a function. Setting $\alpha = 0$ and integrating the velocity equations in x yields the system

$$\partial_t u_0 + \begin{pmatrix} u_0^2 \\ 0 \\ 0 \end{pmatrix} = \nu \Delta u_0, \tag{3.4}$$

which we can solve explicitly to obtain

$$\begin{aligned} u_0^1(t) &= e^{\nu t \Delta}(u_0^1(0) - t u_0^2(0)), \\ u_0^2(t) &= e^{\nu t \Delta} u_0^2(0), \\ u_0^3(t) &= e^{\nu t \Delta} u_0^3(0). \end{aligned}$$

The lift-up effect refers to the linear in time growth of u_0^1 predicted by the formula above for $t \lesssim \nu^{-1}$. In general, the best global in time estimate that one can hope for is

$$\|u_0^1\|_{L^\infty H^s} + \nu^{1/2} \|\nabla u_0^1\|_{L^2 H^s} \lesssim \nu^{-1} \|u_0\|_{H^s}. \tag{3.5}$$

Now we turn to the MHD case. The stabilizing effect of the magnetic field holds just as well in the idealized equations, so to keep the equations as short as possible we set

$\nu = 0$ in this section for simplicity. We introduce the Elsässer variables

$$w^{\pm} = u \mp b, \tag{3.6}$$

which one easily checks from (3.3) are transported either parallel or antiparallel to the background magnetic field. It is natural to unwind by this transport and define the profiles (recall the definition $T_a^t = e^{at\partial_\sigma}$)

$$z^{\pm} = T_{\pm\alpha}^t w^{\pm}. \tag{3.7}$$

Computing from (3.3) we find that $z_0^{\pm}$ solves the system

$$\partial_t z_0^{\pm} + \begin{pmatrix} e^{\pm 2\alpha t \partial_z} P_{\ell \neq 0} z_0^{\mp,2} \\ 0 \\ 0 \end{pmatrix} = 0, \tag{3.8}$$

where we have noted that $P_{\ell\neq 0} z_0^{\pm,2} = z_0^{\pm,2}$ by incompressibility, and also that $T_a^t g_0 = e^{at\partial_z} g_0$ for any function g. While (3.8) has a similar structure to (3.4), the forcing term now exhibits oscillations that nullify the previously observed growth. By direct integration on the Fourier side we obtain the solution

$$\begin{aligned}
\hat{z}^{\pm,1}(t,0,\eta,\ell) &= \hat{w}^{\pm,1}(0,0,\eta,\ell) - \mathbf{1}_{\ell\neq 0} \frac{e^{\pm i\alpha\ell t}}{\alpha\ell} \sin(\alpha\ell t) \hat{w}^{\mp,2}(0,0,\eta,\ell), \\
\hat{z}^{\pm,2}(t,0,\eta,\ell) &= \hat{w}^{\pm,2}(0,0,\eta,\ell), \\
\hat{z}^{\pm,3}(t,0,\eta,\ell) &= \hat{w}^{\pm,3}(0,0,\eta,\ell).
\end{aligned}$$

This immediately yields the estimate

$$\|(u_0(t), b_0(t))\|_{H^s} \lesssim \|(u_0(0), b_0(0))\|_{H^s} \quad \forall\, s \geq 0, \tag{3.9}$$

which is a tremendous gain over (3.5) for $\nu \ll 1$. Proving an estimate like (3.9) for the nonlinear equations will require integration by parts in time. The procedure is fairly straightforward though because the zero mode lift-up term oscillates with frequency $|\alpha\ell| \gg 1$, and hence we can integrate by parts with no losses. See Section 5.5.2 for the calculation.

Remark 2. The growth of the x-averages caused by the lift-up effect is a distinctly 3D phenomena. Indeed, in 2D the divergence free constraint implies that $u_0^2 = z_0^{\pm,2} = 0$, so the discussion above becomes irrelevant.

3.2 Diophantine approximation

In our analysis we gain decay in various oscillatory terms using (2.4). To quantify the possible losses from $\omega(\mathbf{k})$ in the denominator we need facts about Diophantine approximation. For our purposes the following result, which is a consequence of Roth's theorem [32], suffices.

Lemma 3.1. *Let t be an irrational algebraic number and fix any $r > 0$. Then, there exists a constant $C(t, r) > 0$ such that*

$$\left| t - \frac{p}{q} \right| > \frac{C}{|q|^{2+r}} \tag{3.10}$$

for all rational p/q.

From Lemma 3.1, it follows that if $\sigma \in \mathbb{R}^+ \setminus \mathbb{Q}$ is algebraic then for any $r > 0$ and $s \geq 0$ there holds

$$\|\partial_\sigma^{-1} g_{\neq}\|_{H^s} \lesssim_{\sigma,r} \|g_{\neq}\|_{H^{s+1+r}}. \tag{3.11}$$

For $n = n(\sigma)$ and $c = c(\sigma)$ as defined below in Theorem 3.3, the inequality above reads

$$\|\partial_\sigma^{-1} g_{\neq}\|_{H^s} \leq \frac{1}{c}\|g_{\neq}\|_{H^{s+n}}, \tag{3.12}$$

which for the sake of consistency of notation is the form that we will employ in all that follows. As we will see below in Section 3.4, (3.12) in essence says that for the terms involving nonzero modes in x we can integrate by parts in time at the cost of a losing n derivatives.

Remark 3. It is interesting to note that the set of real numbers for which there exists an $r > 0$ such that (3.10) fails to hold for any constant $C > 0$ has Lebesgue measure zero. Moreover, by Louiville's theorem on Diophantine approximation, any irrational number that is algebraic of order two (i.e., is the root of a second degree polynomial with integer coefficients) satisfies (3.10) with $r = 0$ and a constant C that is easy to quantify [32]. This result is in some sense sharp since, as mentioned in Section 2.3, Dirichlet's approximation theorem implies that for any irrational number t there exist infinitely many rational p/q that satisfy $|t - p/q| < 1/q^2$.

3.3 Inviscid damping

3.3.1 3D Euler equations

It is well known that in the Navier-Stokes/Euler setting it is useful to consider the unknown $q = \Delta u$ [17, 83]. In the linearization of the 3D Euler equations around Couette flow its second component simply satisfies (recall our convention of denoting functions in the new coordinates defined by (3.2) with a capital letter)

$$\partial_t Q^2 = 0. \tag{3.13}$$

Hence,

$$\hat{U}^2(t,k,\eta,\ell) = -\frac{\hat{Q}^2(t,k,\eta,\ell)}{k^2+(\eta-kt)^2+\ell^2} = -\frac{\hat{q}^2(0,k,\eta,\ell)}{k^2+(\eta-kt)^2+\ell^2}.$$

From the elementary inequality

$$|k,\eta-kt,\ell|^{-1} \lesssim \langle t\rangle^{-1} |k,\eta,\ell| \quad \forall\, k \neq 0$$

we then obtain for any $s \geq 0$ and $s' \in [0,2]$ the standard inviscid damping estimate

$$\|U^2_{\neq}\|_{H^s} \lesssim \langle t\rangle^{-s'} \|q_{in}\|_{H^{s+s'}} \lesssim \langle t\rangle^{-s'} \|u_{in}\|_{H^{s+2+s'}}. \tag{3.14}$$

The important general fact to observe is that when inverting Δ_L we can gain two powers t at the cost of two derivatives. The loss of regularity in (3.14) is physically meaningful and corresponds to the transient unmixing of information from small scales to large scales by the Couette flow. In particular, for $\eta k > 0$ with $|\eta| \gg |k| \approx |\ell|$ the velocity undergoes a large transient amplification on the time interval $[0, \eta/k]$:

$$\left| \frac{\hat{U}^2(t = \eta/k, k, \eta, \ell)}{\hat{U}^2(0, k, \eta, \ell)} \right| = \frac{k^2 + \eta^2 + \ell^2}{k^2 + \ell^2} \gtrsim \frac{\eta^2}{k^2} = t^2. \tag{3.15}$$

The decay (3.14) and the transient growth (3.15) are together known as the *Orr mechanism*, and the times $t = \eta/k$ are referred to as the *Orr critical times*.

3.3.2 Heuristics for the MHD setting

To study inviscid damping in the MHD setting, we follow the ideas of [17, 83] and define the unknowns

$$f^{\pm} = \Delta z^{\pm}. \tag{3.16}$$

Recall here that z was defined in (3.6) and (3.7). When $\nu = 0$, a computation using (3.3) shows that $F^{\pm,2}$ solves

$$\partial_t F^{\pm,2} + \partial_{XY}^L \Delta_L^{-1} F^{\pm,2} = T_{\pm 2\alpha}^t \partial_{XY}^L \Delta_L^{-1} F^{\mp,2}. \tag{3.17}$$

We explain now some heuristics that suggest (3.17) should give a $\langle t \rangle^{-1}$ inviscid damping estimate. First, since the profiles themselves are not oscillating, we expect in a similar spirit to what was found in Section 3.1 that the right-hand side of (3.17) should have a negligible effect over long times. Dropping this term we have

$$\partial_t F^{\pm,2} + \partial_{XY}^L \Delta_L^{-1} F^{\pm,2} \simeq 0,$$

which on the Fourier side reads

$$\frac{d}{dt}\hat{F}^{\pm,2}(t,k,\eta,\ell) \simeq -\frac{k(\eta-kt)}{k^2+(\eta-kt)^2+\ell^2}\hat{F}^{\pm,2}(t,k,\eta,\ell). \tag{3.18}$$

From (3.18), we see that the right-hand side contributes to growth for $t > \eta/k$. For $k \neq 0$ and $t \to \infty$ the ODE behaves like

$$\frac{d}{dt}\hat{F}^{\pm,2}(t) \simeq \frac{1}{t}\hat{F}^{\pm,2}(t). \tag{3.19}$$

We thus expect $\hat{F}^{\pm,2}(t)$ to grow linearly in time as $t \to \infty$. Recovering two powers of time decay by inverting Δ_L we predict then a $\langle t\rangle^{-1}$ inviscid damping estimate for $U^2_{\neq}$ and $B^2_{\neq}$. We can be more precise with the regularity losses by integrating (3.18) directly to obtain

$$\hat{F}^{\pm,2}(t,k,\eta,\ell) \simeq \sqrt{\frac{k^2+(\eta-kt)^2+\ell^2}{k^2+\eta^2+\ell^2}}\hat{F}^{\pm,2}(0,k,\eta,\ell). \tag{3.20}$$

Using $|k,\eta-kt,\ell|^{-1} \lesssim \langle t\rangle^{-1}|k,\eta,\ell|$ as in the Navier-Stokes case then gives the bound

$$\|(U^2_{\neq},B^2_{\neq})\|_{H^s} \lesssim \langle t\rangle^{-1}\|(u^2_{\text{in}},b^2_{\text{in}})\|_{H^{s+2}}. \tag{3.21}$$

The estimate (3.21) is not actually correct. The regularity loss in (3.21) is due to the Orr mechanism, but one needs to pay additional regularity to account for dropping the oscillatory term in (3.17). The correct inviscid damping estimate for (3.17) is of the form

$$\|(U^2_{\neq},B^2_{\neq})\|_{H^s} \lesssim \langle t\rangle^{-1}\|(u^2_{\text{in}},b^2_{\text{in}})\|_{H^{s+2+m}} \tag{3.22}$$

for some $m > 0$ sufficiently large. Justifying (3.22) with integration by parts in time and the results quoted in Section 3.2 is discussed in the next section.

3.4 Integration by parts in time for modes with $k \neq 0$

We will not bother to give a complete proof of (3.22) for the linearized equations because it is somewhat tedious, yet strictly easier than proving the inviscid damping estimate in Theorem 3.3. Instead, we sketch the proof of a linear in time growth estimate for a model equation that captures the same growth and oscillation timescales as (3.17), but removes the frequency dependence in the coefficients. This is enough to motivate how to perform the energy estimates that yield inviscid damping in the nonlinear equations. More generally, it motivates how the integration by parts in time influences the bootstrap hypotheses used in the proof of Theorem 3.3 (see Section 4.4).

We consider the model equation

$$\partial_t F^{\pm} - \frac{1}{t} F^{\pm} = \frac{1}{t} T^{t}_{\pm 2\alpha} F^{\mp} \tag{3.23}$$

for $t \geq 1$ and $k \neq 0$, where for simplicity in this section we drop the second superscript. Noting that the left-hand side of (3.23) can be rewritten as $t\partial_t (t^{-1} F^{\pm})$, we obtain the a priori estimate

$$\frac{1}{2} \| t^{-1} F^{\pm}(t) \|_{L^2}^2 = \frac{1}{2} \| F^{\pm}(1) \|_{L^2}^2 + \int_1^t \int s^{-3} F^{\pm}(s) T^{s}_{\pm 2\alpha} F^{\mp}(s) dV ds. \tag{3.24}$$

Using (2.4) and Plancherel's theorem we integrate by parts in the oscillating term to obtain

$$\Big|\int_1^t \int s^{-3} F^{\pm} T^s_{2\alpha} F^{\mp} dV ds\Big| = \Big|\int_1^t \int s^{-3} F^{\pm} \frac{1}{2\alpha} (\partial_s T^s_{2\alpha}) \partial_\sigma^{-1} F^{\mp} dV ds\Big| \tag{3.25}$$

$$\leq \sum_{k,\ell\in\mathbb{Z}, k\neq 0} \int_{\eta\in\mathbb{R}} \int_1^t \frac{1}{2\alpha|\sigma k+\ell|} \left[s^{-1} |\partial_s (s^{-1}\hat{F}^{\pm})| |s^{-1}\hat{F}^{\mp}| \right] d\eta ds \tag{3.26}$$

$$+ \text{ symmetric terms } + \text{boundary terms},$$

where the symmetric terms correspond to the time derivative landing on the other two factors in the brackets above. Integrating by parts in time gains decay because the time derivatives of the profiles carry an extra factor of t^{-1} in comparison to the profiles themselves. Observe however that this gain costs a $|\sigma k+\ell|^{-1}$ factor, which, even for $\sigma \in \mathbb{R}\setminus\mathbb{Q}$, blows up as the oscillation frequency degenerates for $|k|\to\infty$. As previously mentioned, the idea behind the Diophantine condition in Theorem 1 is that we can absorb such losses by paying regularity. Using

$$\partial_t(t^{-1}\hat{F}^{\pm}) = t^{-2} e^{\pm 2i\alpha(\sigma k+\ell)} \hat{F}^{\mp},$$

(3.12), and the Cauchy-Schwarz inequality it follows that for any $0<\theta\ll 1$ there holds

$$|(3.26)| \lesssim_{\sigma,\theta} \frac{1}{\alpha} \|s^{-1}F^{\mp}\|_{L^\infty(1,t;L^2)} \|s^{-2+\theta}F^{\mp}\|_{L^\infty(1,t;H^n)}. \tag{3.27}$$

By estimating the symmetric and boundary terms above in a similar fashion, one can show using (3.24) and a standard bootstrap argument that the estimate

$$\|F^+(t)\|_{L^2} + \|F^-(t)\|_{L^2} \lesssim t(\|F^+(1)\|_{H^n} + \|F^-(1)\|_{H^n}) \quad \forall\, t \geq 1 \tag{3.28}$$

holds provided both α is sufficiently large and

$$\|F^+(t)\|_{H^n} + \|F^-(t)\|_{H^n} \lesssim t^{2-\theta}(\|F^+(1)\|_{H^n} + \|F^-(1)\|_{H^n}) \quad \forall\, t \geq 1. \tag{3.29}$$

That is, a linear in time growth estimate that loses derivatives holds for (3.23) provided we have an estimate in a sufficiently higher norm that, while losing no regularity, allows for greater time growth. In practice, (3.29) only holds with $\theta = 0$ because the differential inequality that follows for free from (3.23) is

$$\partial_t(|\hat{F}^-|^2 + |\hat{F}^+|^2) \leq \frac{4}{t}(|\hat{F}^-|^2 + |\hat{F}^+|^2),$$

which implies only a quadratic growth bound. This does not pose any issue though since one can start with a quadratic growth bound in H^{2n} and iterate the argument above twice to prove that

$$\|F^+(t)\|_{L^2} + \|F^-(t)\|_{L^2} \lesssim t(\|F^+(1)\|_{H^{2n}} + \|F^-(1)\|_{H^{2n}}) \quad \forall\, t \geq 0. \tag{3.30}$$

We omit the details since this is more of a technical matter than something deep.

Since $n = n(\sigma)$, the computations above suggest that the inviscid damping estimate (3.22) can hold only if we are willing to pay the additional regularity m dependent upon the choice of σ. Moreover, we see that its proof in the nonlinear equations should be based on combining high norm energy estimates that grow in time with lower regularity bounds that allow for less time growth. This general strategy is key for us and used elsewhere in the proof of Theorem 3.3. For example, we also use it when estimating the nonzero mode portion of the lift-up term. We refer to Sections 5.6.1 and 5.8.2 for the main estimates on the nonlinear equations that involve integration by parts in time for the nonzero modes in x.

Remark 4. The loss of regularity from the integration by parts in time is a distinctly 3D phenomenon for the system we are considering. The analogous 2D problem is to consider the stability of the stationary solution

$$u_{s,2D} = (y, 0), \quad b_{s,2D} = (\beta, 0) \tag{3.31}$$

for $\beta > 0$. In this case, for a Fourier mode with $\mathbf{k} = (k, \eta) \in \mathbb{Z} \setminus \{0\} \times \mathbb{R}$ the oscillations induced from the magnetic field occur with frequency $\omega(\mathbf{k}) = |\beta k| \gtrsim 1$. Hence, one can integrate by parts in time without any losses. In spite of this, the 2D MHD problem (3.31) seems to be less well behaved than Couette flow in the 2D Euler (or Navier-Stokes) equations. In particular, when $\mu = \nu = 0$ the 2D MHD system linearized around (3.31) reads

$$\begin{cases} \partial_t \omega + y\partial_x \omega - \beta \partial_x j = 0 \\ \partial_t j + y\partial_x j - \beta \partial_x \omega + 2\partial_{xy}\Delta^{-1} j = 0, \end{cases} \tag{3.32}$$

where $\omega = \partial_x u^2 - \partial_y u^1$ and $j = \partial_x b^2 - \partial_y b^1$ are the vorticity and current perturbations, respectively. Note that below we write J to denote the current in the new coordinate system, not to be confused with J from (2.1). The term $2\partial_{xy}\Delta^{-1}j$ causes growth of the current akin to the growth of $F^{\pm,2}$ just discussed in Sec. 3.3.2. Switching to the unknowns $\omega \pm j$, using the coordinate transform (3.2), and defining profile variables that unwind by the transport along the magnetic field lines results in a Fourier side ODE with the same structure as (2.17). As long as $\beta > 0$ is sufficiently large, studying this ODE with a straightforward application of integration by parts in time (no regularity losses) and then using the Biot-Savart law

$$(U, B) = \begin{pmatrix} -\partial_Y + t\partial_X \\ \partial_X \end{pmatrix} \Delta_L^{-1}(\Omega, J)$$

yields, for any $s \geq 0$, the estimates

$$\|(U, B)\|_{L^\infty H^s} \lesssim \|(u_{in}, b_{in})\|_{H^s},$$
$$\| \langle t \rangle^\delta (U^2_{\neq}, B^2_{\neq})\|_{L^\infty H^s} \lesssim \|(u_{in}, b_{in})\|_{H^{s+\delta}} \quad \forall\, \delta \in [0, 1].$$

This should be contrasted with the linearization of the 2D Euler equations around Couette flow, for which it is well known that *both* velocity components experience inviscid damping. Specifically, using the Biot-Savart law and the fact that the vorticity perturbation simply solves $\partial_t \omega + y\partial_x \omega = 0$, one readily derives the estimate

$$\| \langle t \rangle U^1_{\neq}\|_{L^\infty H^{s-2}} + \| \langle t \rangle^2 U^2_{\neq}\|_{H^{s-3}} \lesssim \|u_{in}\|_{H^s} \quad \forall\, s \geq 3.$$

3.5 Quadratic growth of $F^{\pm,j}$ for $j \in \{1,3\}$

After some calculations with (3.1) we find that when $\nu = 0$, $F^{\pm,1}$ and $F^{\pm,3}$ satisfy

$$\partial_t F^{\pm,j} + 2\partial_{XY}^L \Delta_L^{-1} F^{\pm,j} + \mathbf{1}_{j=1} T_{\pm 2\alpha}^t F^{\mp,2}$$
$$= \partial_j^L \Delta_L^{-1} \partial_X F^{\pm,2} + T_{\pm 2\alpha}^t \partial_j^L \Delta_L^{-1} \partial_X F^{\mp,2}, \quad j \in \{1,3\}.$$

Dropping the oscillatory terms, the leading order behavior as $t \to \infty$ is the ODE

$$\frac{d}{dt}\hat{F}^{\pm,j} \simeq \frac{2}{t}\hat{F}^{\pm,j}, \quad j \in \{1,3\}.$$

The factor of 2 implies that in general $F^{\pm,1}$ and $F^{\pm,3}$ grow quadratically. Thus, inverting Δ_L does not yield any uniform in ν decay for the first or third component of $(U_{\neq}, B_{\neq})$.

3.6 Enhanced dissipation

The modified Laplace operator Δ_L leads to improved dissipation timescales. To see this, consider the model equation

$$\partial_t g = \nu \Delta_L g, \tag{3.33}$$

which on the Fourier side has solution

$$\hat{g}(t,k,\eta,\ell) = \hat{g}(0,t,k,\eta,\ell) e^{-\nu \int_0^t (k^2 + (\eta - ks)^2 + \ell^2) ds}. \tag{3.34}$$

Since $\int_0^t (k^2 + (\eta - ks)^2 + \ell^2) ds \geq k^2 t^3/12$, we obtain the estimate

$$\|g_{\neq}(t)\|_{H^s} \leq e^{-\nu t^3/12} \|g_{\neq}(0)\|_{H^s} \quad \forall\, s \geq 0. \tag{3.35}$$

Hence, the nonzero modes decay on the timescale $t \sim \nu^{-1/3}$, which for $\nu \ll 1$ is a large improvement on the ν^{-1} dissipation timescale of the usual heat equation.

3.7 Linear growth of $(U^1_{\neq}, B^1_{\neq})$

The quadratic growth of $F^{\pm,1}$ and $F^{\pm,3}$ arises naturally from their definition involving Δ_L and is not related to growth of U and B. On the other hand, there is some genuine growth possible for $(U^1_{\neq}, B^1_{\neq})$ from the lift-up term. Specifically, one expects that if the strongest quantitative assumption on the initial data is

$$\|(u_{in}, b_{in})\|_{H^m} \lesssim \epsilon,$$

then in general $\|(U^1_{\neq}, B^1_{\neq})\|_{H^{m'}}$ could reach size $\mathcal{O}(\epsilon \nu^{-1/3})$ when $m' \leq m$ is sufficiently large. To see this, recall that the heuristics in Section 3.4 suggest that when estimating the nonzero modes we can only expect to gain from oscillations if we pay regularity. In the situation at hand, this means that for $m - m'$ sufficiently small we cannot utilize oscillations when estimating the nonzero mode lift-up term in $H^{m'}$. In the worst case scenario then we expect the inviscid problem to satisfy a bound no better than

$$\|U^1_{\neq}(t) \pm B^1_{\neq}(t)\|_{H^{m'}} \lesssim \|U^1_{\neq}(0) \pm B^1_{\neq}(0)\|_{H^{m'}} \\ + \int_0^t \left(\|U^2_{\neq}(\tau)\|_{H^{m'}} + \|B^2_{\neq}(\tau)\|_{H^{m'}}\right) d\tau \lesssim \epsilon \langle t \rangle .$$

Accounting for the enhanced dissipation we predict then the optimal bound

$$\|U^1_{\neq}(t) \pm B^1_{\neq}(t)\|_{H^{m'}} \lesssim \epsilon \langle t \rangle e^{-\nu^{1/3}t} \lesssim \epsilon \nu^{-1/3}.$$

A discussion similar to the one above implies that $(U^2_{\neq}, B^2_{\neq})$ does not decay in $H^{m'}$ (on a timescale independent of ν) for $m - m'$ small enough, so when $\nu = 0$ at best one can hope for

$$\|(U^2_{\neq}, B^2_{\neq})\|_{L^\infty H^{m'}} \lesssim \epsilon.$$

3.8 Summary of linear estimates

We now give a precise statements that summarize the linear estimates for (3.1).

Proposition 3.2. *Let $0 < \nu \ll 1$ and suppose that $\sigma \in \mathbb{R}^+ \setminus \mathbb{Q}$ satisfies (3.12). Let $N \in \mathbb{N}$ be sufficiently large. There exists $m \in \mathbb{N}$ large enough and $c > 0$ sufficiently small so that* *the solution to* (3.1) *satisfies the following estimates:*

$$\|e^{c\nu^{1/3}t}\langle t\rangle(U^2_{\neq},B^2_{\neq})\|_{L^\infty H^{N-2-m}}+\|e^{c\nu^{1/3}t}(U^2_{\neq},B^2_{\neq})\|_{L^\infty H^{N-2}}\lesssim\|(u^2_{in},b^2_{in})\|_{H^N},\tag{3.36}$$

$$\nu^{1/3}\|e^{c\nu^{1/3}t}(U^1_{\neq},B^1_{\neq})\|_{L^\infty H^{N-2}}+\|e^{c\nu^{1/3}t}(U^1_{\neq},B^1_{\neq})\|_{L^\infty H^{N-2-m}}\lesssim\|(u_{in},b_{in})\|_{H^N},\tag{3.37}$$

$$\|e^{c\nu^{1/3}t}(U^3_{\neq},B^3_{\neq})\|_{L^\infty H^{N-2}}\lesssim\|(u_{in},b_{in})\|_{H^N},\tag{3.38}$$

$$\|(u_0,b_0)\|_{L^\infty H^N}\lesssim(u_{0,in},b_{0,in})\|_{H^N}.\tag{3.39}$$

We will not prove Proposition 3.2 in detail. The computations are based essentially entirely on the ideas outlined above and the Fourier multiplier techniques used in the proof of our nonlinear stability theorem stated below. Integration by parts in time is needed to obtain to the estimate on the zero mode and the H^{N-2-m} estimates.

3.9 Statement of main nonlinear stability result

Now that we have discussed in detail the linearization of (2.3) we are ready to give the details of our main result.

Theorem 3.3. *Let $\mu=\nu\in(0,1]$ and suppose that $\sigma\in\mathbb{R}^+\setminus\mathbb{Q}$ is such that*

$$\inf_{p,q\in\mathbb{Z}}|q|^n\,|q\sigma-p|=c>0\tag{3.40}$$

for some $n \geq 1$. Then, there exist universal constants $\delta > 0$ sufficiently small and $c_1 > 0$ sufficiently large such that for any $N \geq 11 + 3n$ there is a constant $c_0(N) > 0$ such that if $\alpha > c_1/c$ and

$$\|(u_{in}, b_{in})\|_{H^{N+2}} = \epsilon \leq c_0 \nu,$$

then the solution to (2.3) is global in time and, denoting $N' = N - 4 - 2n$ and $N'' = N - 9 - 3n$, the profiles $U(t, X, Y, Z) = u(t, X + Yt, Y, Z)$ and $B(t, X, Y, Z) = b(t, X + Yt, Y, Z)$ satisfy the global estimates

$$\|e^{\delta\nu^{1/3}t}\Delta_{X,Z}(U^2_{\neq}, B^2_{\neq})\|_{L^\infty H^N} + \nu^{1/6}\|\Delta_{X,Z}(U^2_{\neq}, B^2_{\neq})\|_{L^2 H^N} \lesssim \epsilon, \quad \text{(3.41a)}$$

$$\|\nabla_{X,Z}(U^2_{\neq}, B^2_{\neq})\|_{L^2 H^{N'}} + \| \langle t \rangle \nabla_{X,Z}(U^2_{\neq}, B^2_{\neq})\|_{L^\infty H^{N'-1}} \lesssim \epsilon, \quad \text{(3.41b)}$$

$$(j \in \{1, 3\}) \quad \|e^{\delta\nu^{1/3}t}\Delta_{X,Z}(U^j_{\neq}, B^j_{\neq})\|_{L^\infty H^{N''}} + \nu^{1/6}\|\Delta_{X,Z}(U^j_{\neq}, B^j_{\neq})\|_{L^2 H^{N''}} \lesssim \epsilon, \quad \text{(3.41c)}$$

$$\|(u_0, b_0)\|_{L^\infty H^N} + \nu^{1/2}\|\nabla(u_0, b_0)\|_{L^2 H^N} \lesssim \epsilon \quad \text{(3.41d)}$$

$$(j \in \{1, 3\}) \quad \|e^{\delta\nu^{1/3}t}\Delta_{X,Z}(U^j_{\neq}, B^j_{\neq})\|_{L^\infty H^N} + \nu^{1/6}\|\Delta_{X,Z}(U^j_{\neq}, B^j_{\neq})\|_{L^2 H^N} \lesssim \epsilon\nu^{-1/3}, \quad \text{(3.41e)}$$

where the implicit constants are independent of ν, N, n, and c.

Remark 5. The enhanced dissipation of the nonzero modes is described by the $e^{\delta\nu^{1/3}t}$ factors and the $\nu^{-1/6}$ scaling of the L^2 in time estimates in (3.41a), (3.41c), and (3.41e). Indeed, for $\nu \ll 1$ the $\nu^{-1/6}$ scaling is an improvement on the $\nu^{-1/2}$ scaling that holds for the heat equation. The inviscid damping is captured by the uniform in ν bounds in (3.41b). Notice in particular that the $\langle t \rangle^{-1}$ decay matches the estimate predicted by the linear theory. The estimate (3.41d) describes the suppression of the lift-up effect. The loss of $\nu^{-1/3}$ for $j = 1$ in the high norm bound (3.41e) is due to the lift-up effect for the nonzero modes, as described in Section 3.7. The loss in the third component is due to various nonlinear interactions with the first component.

Remark 6. The discussion in Remark 3 implies that $n = 1$ is the minimal number satisfying (3.40), and that for almost every $\sigma \in \mathbb{R}$ we may take $n = 1 + r$ for any $r > 0$. Clearly then $n < 2$ is generic, however, the specific value of n does not affect the structure of our proof, and so we take n to be arbitrary to account for possibly exceptional circumstances.

In the case that σ is arbitrary (possibly rational), the methods employed in the proof of Theorem 3.3 yield the following corollary.

Corollary 3.4. *Let $\mu = \nu \in (0, 1]$, $\sigma \in \mathbb{R}^+$, and α_0 be a sufficiently large universal constant. Then, for $\alpha > \alpha_0$ and any $N > 3/2$ we have $\gamma(X_i = H^{N+2}, X_f = X^N) \leq 4/3$. As in Theorem 3.3, X_f measures derivatives on the profiles.*

Notice that $\gamma \leq 4/3$ is still an improvement on the threshold estimate of $\gamma \leq 3/2$ in [17]. The gain is possible because even with rational σ the presence of the magnetic field allows us to eliminate the lift-up effect in the zero mode. The gap between the results in Corollary 3.4 and Theorem 3.3 arises because we lose the inviscid damping when σ does not satisfy (3.40). The proof then does not require a calculation analogous to that in Section 5.6.1. In fact, it only requires integration by parts in time in the zero mode lift-up term, which does not cause a loss of derivatives. We thus only need to perform estimates at a single regularity level, and hence the proof of Corollary 3.4 is much simpler than that of Theorem 3.3.

Chapter 4: Preliminaries and Outline of the Proof

In this chapter carry out preliminary steps in the proof of Theorem 3.3. These preliminaries include defining the Fourier multiplier norm to be used in our energy estimates, reformulating the equation in terms of new dependent variables, and setting up a bootstrap argument.

4.1 Fourier multiplier norm

Inspired by the previous works [16, 17, 19, 23], our proof is based on energy estimates using weighted norms defined through Fourier multipliers. The multipliers that we employ have all, up to small modifications, been previously used in [17].

4.1.1 Quadratic growth multipliers m and $\tilde{m}$

The first class of multipliers we use are concerned with using the dissipation in the optimal way to absorb the quadratic in time growth of $F^{\pm,j}$ ($j \in \{1,3\}$) and the linear growth of $F^{\pm,2}$. Consider the model scalar equation

$$\partial_t g + 2\partial_{XY}^L \Delta_L^{-1} g = \nu \Delta_L g. \tag{4.1}$$

On the Fourier side this equation becomes

$$\partial_t \hat{g} + \frac{2k(\eta - kt)}{k^2 + (\eta - kt)^2 + \ell^2}\hat{g} = -\nu(k^2 + (\eta - kt)^2 + \ell^2)\hat{g}. \tag{4.2}$$

For $k \neq 0$ the term $\frac{2k(\eta-kt)}{k^2+(\eta-kt)^2+\ell^2}\hat{g}$ contributes to growth in $\hat{g}$ for $t \geq \eta/k$. On the other hand, for $k \neq 0$ the term on the right-hand side yields enhanced dissipation, which will overcome the growth for $|t - \eta/k|$ sufficiently large with respect to some inverse power of ν. In fact, one can check that for $k \neq 0$ there holds

$$\left|\frac{2k(\eta - kt)}{k^2 + (\eta - kt)^2 + \ell^2}\right| \leq \frac{\nu}{32}(k^2 + (\eta - kt)^2 + \ell^2)$$

whenever $|t - \eta/k| \geq 4\nu^{-1/3}$. This motivates defining the multiplier m by

$$m(0, k, \eta, \ell) = 1$$

and the ODE

$$\frac{\dot{m}}{m} = \begin{cases} \frac{2k(\eta-kt)}{k^2+(\eta-kt)^2+\ell^2} & \text{if } 0 \leq t - \eta/k \leq 4\nu^{-1/3}, \\ 0 & \text{otherwise.} \end{cases} \tag{4.3}$$

For certain unknowns it will also be useful to use a norm that weakens for each frequency indefinitely after the critical time. We thus define the modified multiplier $\tilde{m}$ by

$$\frac{\dot{\tilde{m}}}{\tilde{m}} = \begin{cases} \frac{2k(\eta-kt)}{k^2+(\eta-kt)^2+\ell^2} & \text{if } t \geq \eta/k, \\ 0 & \text{otherwise.} \end{cases} \tag{4.4}$$

From (4.3) and (4.4) we find that m and $\tilde{m}$ are given by the exact formulas

- $k = 0$: $m(t, 0, \eta, \ell) = \tilde{m}(t, 0, \eta, \ell) = 1$;

- $k \neq 0$, $\eta k < 0$ and $|\eta| \geq 4\nu^{-1/3}|k|$:

$$m(t, k, \eta, \ell) = 1, \quad \tilde{m}(t, k, \eta, \ell) = \frac{k^2 + \eta^2 + \ell^2}{k^2 + (\eta - kt)^2 + \ell^2};$$

- $k \neq 0$, $\eta k < 0$ and $|\eta| \leq 4\nu^{-1/3}|k|$:

$$m(t, k, \eta, \ell) = \begin{cases} \frac{k^2+\eta^2+\ell^2}{k^2+(\eta-kt)^2+\ell^2} & \text{if } t \in [0, \eta/k + 4\nu^{-1/3}), \\ \frac{k^2+\eta^2+\ell^2}{k^2+(4k\nu^{-1/3})^2+\ell^2} & \text{otherwise,} \end{cases}$$

$$\tilde{m}(t, k, \eta, \ell) = \frac{k^2 + \eta^2 + \ell^2}{k^2 + (\eta - kt)^2 + \ell^2};$$

- $k \neq 0$ and $\eta k > 0$:

$$m(t, k, \eta, \ell) = \begin{cases} 1 & \text{if } t \leq \eta/k, \\ \frac{k^2+\ell^2}{k^2+(\eta-kt)^2+\ell^2} & \text{if } t \in (\eta/k, \eta/k + 4\nu^{-1/3}), \\ \frac{k^2+\ell^2}{k^2+(4k\nu^{-1/3})^2+\ell^2} & \text{if } t \geq \eta/k + 4\nu^{-1/3}; \end{cases}$$

$$\tilde{m}(t, k, \eta, \ell) = \begin{cases} 1 & \text{if } t \leq \eta/k, \\ \frac{k^2+\ell^2}{k^2+(\eta-kt)^2+\ell^2} & \text{if } t > \eta/k. \end{cases}$$

The natural multiplier to use in the norm for $F^{\pm,2}$, which is expected to grow linearly in time, is $m^{1/2}$. While it can be obtained from the formulas above, it is useful to know that it satisfies

$$\frac{\dot{m}^{1/2}}{m^{1/2}} = \begin{cases} \frac{k(\eta-kt)}{k^2+(\eta-kt)^2+\ell^2} & \text{if } 0 \le t - \eta/k \le 4\nu^{-1/3}, \\ 0 & \text{otherwise.} \end{cases} \tag{4.5}$$

The fundamental properties of m and $\tilde{m}$ are summarized in the following lemma.

Lemma 4.1. *The multipliers m and $\tilde{m}$ satisfy*

$$\tilde{m}(t,k,\eta,\ell) \le m(t,k,\eta,\ell) \le 1, \tag{4.6a}$$

$$k^2 + \ell^2 \lesssim (k^2 + (\eta - kt)^2 + \ell^2)\tilde{m}, \tag{4.6b}$$

$$\nu^{2/3} \lesssim m(t,k,\eta,\ell), \tag{4.6c}$$

$$\frac{1}{\tilde{m}} + \frac{1}{m} \lesssim \langle t \rangle^2, \tag{4.6d}$$

$$\frac{\tilde{m}(t,k,\eta,\ell)}{\tilde{m}(t,k,\eta',\ell')} + \frac{m(t,k,\eta,\ell)}{m(t,k,\eta',\ell')} \lesssim \langle \eta - \eta' \rangle^2 + \langle \ell - \ell' \rangle^2, \tag{4.6e}$$

$$\tilde{m}(t,k,\eta,\ell) \lesssim \frac{|k,\eta,\ell|^4}{\langle t \rangle^2}. \tag{4.6f}$$

Except for (4.6e), the proof of Lemma 4.1 is essentially immediate from the exact formulas above. Inequality (4.6e) was proven for m in [17] and the proof for $\tilde{m}$ does not require any notable variations. Thus, we omit it for the sake of brevity. In the proof of Theorem 3.3 we will use (4.6a)-(4.6d) so frequently that we will often do so without any remark.

4.1.2 Ghost multiplier M

We also introduce three additional multipliers M_1, M_2, and M_3. These multipliers are defined by $M_j(0, k, \eta, \ell) = 1$, $M_j(t, 0, \eta, \ell) = 1$, and for $k \neq 0$ the differential equations

$$-\frac{\dot{M}_1}{M_1} = \frac{k^2}{k^2 + \ell^2 + (\eta - kt)^2}, \tag{4.7a}$$

$$-\frac{\dot{M}_2}{M_2} = \frac{\langle k\ell \rangle}{k^2 + \ell^2 + (\eta - kt)^2}, \tag{4.7b}$$

$$-\frac{\dot{M}_3}{M_3} = \frac{\nu^{1/3} k^2}{k^2 + \nu^{2/3}(\eta - kt)^2}. \tag{4.7c}$$

We then define $M = M_1 M_2 M_3$ and observe that it satisfies

$$-\frac{\dot{M}}{M} \geq -\frac{\dot{M}_j}{M_j}$$

for each $j \in \{1, 2, 3\}$. It follows readily by direct integration that there exists a universal constant $c_2 > 0$ such that for any j there holds

$$c_2 \leq M_j(t, k, \eta, \ell) \leq 1. \tag{4.8}$$

Hence, M is bounded below by a universal, positive constant. We see then that the multiplier M is essentially a Fourier side analogue of Alinhac's ghost energy method for quasilinear wave equations [5], which is the origin of the terminology "ghost multiplier." The multipliers M_1 and M_2 are used to quantify the inviscid damping with time integrated estimates that do not lose regularity; see for example the first term in (3.41b) and compare with the pointwise estimate (3.22). Moreover, they are useful to control terms arising from the linear pressure. The multiplier M_3 is designed to balance the transient slow down of the enhanced dissipation that occurs near the critical times. This is quantified by the following lemma.

Lemma 4.2. *There exists a universal constant $c_3 > 0$ such that for $k \neq 0$ there holds*

$$c_3\nu^{1/6} \leq \nu^{1/2}|k, \eta - kt, \ell| + \sqrt{-\dot{M}_3 M_3}.$$

Proof. If $|t - \eta/k| \geq \nu^{-1/3}$ then the estimate follows since $|k, \eta - kt, \ell| \geq |t - \eta/k|$ for $k \neq 0$. On the other hand, if $|t - \eta/k| \leq \nu^{-1/3}$ then

$$-\dot{M}_3 M_3 \gtrsim -\frac{\dot{M}_3}{M_3} = \frac{\nu^{1/3}}{1 + \nu^{2/3}\left(t - \frac{\eta}{k}\right)^2} \geq \frac{\nu^{1/3}}{2},$$

as desired. □

Using Lemma 4.2 we can obtain both pointwise and L^2 in time enhanced dissipation estimates that agree with the scaling suggested by the linear theory. See for example the proof of Lemma 4.5 and the treatment of the term L_λ in Section 5.1.4.

4.2 Frequency decompositions

Since we perform estimates at various regularity levels, Fourier space decompositions play an important role in the proof. For our purposes it suffices to define the sharp cutoff function $\chi : \mathbb{R}^6 \to \mathbb{R}$ by

$$\chi(\xi \in \mathbb{R}^3, \xi' \in \mathbb{R}^3) = \begin{cases} 1 & \text{if } |\xi - \xi'| \leq 2|\xi'|, \\ 0 & \text{otherwise.} \end{cases}$$

We then define the paraproduct decomposition

$$\begin{aligned} fg &= \mathcal{F}^{-1} \sum_{k',\ell' \in \mathbb{Z}} \int_{\eta' \in \mathbb{R}} \hat{f}(k', \eta', \ell') \hat{g}(k - k', \eta - \eta', \ell - \ell') \chi(k, \eta, \ell, k', \eta', \ell') d\eta' \\ &+ \mathcal{F}^{-1} \sum_{k',\ell' \in \mathbb{Z}} \int_{\eta' \in \mathbb{R}} \hat{f}(k', \eta', \ell') \hat{g}(k - k', \eta - \eta', \ell - \ell')(1 - \chi(k, \eta, \ell, k', \eta', \ell')) d\eta' \\ &:= f^{\text{Hi}} g^{\text{Lo}} + f^{\text{Lo}} g^{\text{Hi}}. \end{aligned}$$

From the triangle inequality, Young's inequality, and Sobolev embedding we have, for any $s > 0$ and $\kappa > 3/2$,

$$\|f^{\text{Hi}} g^{\text{Lo}}\|_{H^s} \lesssim_{\kappa,s} \|f\|_{H^s} \|g\|_{H^\kappa}. \tag{4.9}$$

4.3 Reformulation of the equations

We work in the coordinate system defined by (3.2) and primarily on the unknowns $F^{\pm,i}$. Recall the definitions (3.6), (3.7), (3.16), the shorthand $T_a^t = e^{at\partial_\sigma}$, and our convention to use capital letters to denote unknowns in the new coordinates. The unknowns $F^{\pm}$ satisfy

$$\begin{aligned}
&\partial_t F^{\pm,1} + T^t_{\pm 2\alpha} Z^{\mp} \cdot \nabla_L F^{\pm,1} + T^t_{\pm 2\alpha} F^{\mp} \cdot \nabla_L Z^{\pm,1} + 2T^t_{\pm 2\alpha} \partial^L_i Z^{\mp,j} \partial^L_{ij} Z^{\pm,1} \\
&\quad + 2\partial^L_{XY} \Delta_L^{-1} F^{\pm,1} + T^t_{\pm 2\alpha} F^{\mp,2} - \partial_{XX} \Delta_L^{-1} (F^{\pm,2} + T^t_{\pm 2\alpha} F^{\mp,2}) \\
&\quad = \partial_X (T^t_{\pm 2\alpha} \partial^L_j Z^{\mp,i} \partial^L_i Z^{\pm,j}) + \nu \Delta_L F^{\pm,1},
\end{aligned} \tag{4.10}$$

$$\begin{aligned}
&\partial_t F^{\pm,2} + T^t_{\pm 2\alpha} Z^{\mp} \cdot \nabla_L F^{\pm,2} + T^t_{\pm 2\alpha} F^{\mp} \cdot \nabla_L Z^{\pm,2} \\
&\quad + 2T^t_{\pm 2\alpha} \partial^L_i Z^{\mp,j} \partial^L_{ij} Z^{\pm,2} + \partial^L_{XY} \Delta_L^{-1} F^{\pm,2} - T^t_{\pm 2\alpha} \partial^L_{XY} \Delta_L^{-1} F^{\mp,2} \\
&\quad = \partial^L_Y (T^t_{\pm 2\alpha} \partial^L_j Z^{\mp,i} \partial^L_i Z^{\pm,j}) + \nu \Delta_L F^{\pm,2},
\end{aligned} \tag{4.11}$$

$$\begin{aligned}
&\partial_t F^{\pm,3} + T^t_{\pm 2\alpha} Z^{\mp} \cdot \nabla_L F^{\pm,3} + T^t_{\pm 2\alpha} F^{\mp} \cdot \nabla_L Z^{\pm,3} + 2T^t_{\pm 2\alpha} \partial^L_i Z^{\mp,j} \partial^L_{ij} Z^{\pm,3} \\
&\quad + 2\partial^L_{XY} \Delta_L^{-1} F^{\pm,3} - \partial_{XZ} \Delta_L^{-1} (F^{\pm,2} + T^t_{\pm 2\alpha} F^{\mp,2}) \\
&\quad = \partial_Z (T^t_{\pm 2\alpha} \partial^L_j Z^{\mp,i} \partial^L_i Z^{\pm,j}) + \nu \Delta_L F^{\pm,3},
\end{aligned} \tag{4.12}$$

where summation over repeated indices is implied, $i, j \in \{1, 2, 3\}$ corresponds to $\{X, Y, Z\}$ in the derivative operators, and we have written $T^t_{\pm 2\alpha} f g$ to mean $(T^t_{\pm 2\alpha} f) g$ in the nonlinear terms. At times we will also work with the unknowns $Q = \Delta_L U$ and $H = \Delta_L B$. In particular, the second components, which satisfy

$$\begin{aligned}
&\partial_t Q^2 + Q \cdot \nabla_L U^2 + U \cdot \nabla_L Q^2 - H \cdot \nabla_L B^2 - B \cdot \nabla_L H^2 + 2\partial^L_i U^j \partial^L_{ij} U^2 \\
&\quad - 2\partial^L_i B^j \partial^L_{ij} B^2 - \alpha \partial_\sigma H^2 = \nu \Delta_L Q^2 + \partial^L_Y (\partial^L_j U^i \partial^L_i U^j - \partial^L_j B^i \partial^L_i B^j)
\end{aligned} \tag{4.13}$$

and

$$\partial_t H^2 + Q\cdot\nabla_L B^2 + U\cdot\nabla_L H^2 - H\cdot\nabla_L U^2 - B\cdot\nabla_L Q^2 \\ + 2\partial_i^L U^j \partial_{ij}^L B^2 - 2\partial_i^L B^j \partial_{ij}^L U^2 + 2\partial_{XY}^L \Delta_L^{-1} H^2 - \alpha\partial_\sigma Q^2 = \nu\Delta_L H^2. \tag{4.14}$$

Lastly, for certain estimates we work directly on $Z_0^{\pm}$, which solves

$$\begin{cases} \partial_t Z_0^{\pm,1} + (T_{\pm 2\alpha}^t Z^{\mp}\cdot\nabla_L Z^{\pm,1})_0 + T_{\pm 2\alpha}^t Z_0^{\mp,2} = \nu\Delta Z_0^{\pm,1}, \\ \partial_t Z_0^{\pm,2} + (T_{\pm 2\alpha}^t Z^{\mp}\cdot\nabla_L Z^{\pm,2})_0 = \nu\Delta Z_0^{\pm,2} + \partial_Y \Delta^{-1}\partial_{ij}(T_{\pm 2\alpha}^t Z^{\mp,i} Z^{\pm,j})_0, \\ \partial_t Z_0^{\pm,3} + (T_{\pm 2\alpha}^t Z^{\mp}\cdot\nabla_L Z^{\pm,3})_0 = \nu\Delta Z_0^{\pm,3} + \partial_Z \Delta^{-1}\partial_{ij}(T_{\pm 2\alpha}^t Z^{\mp,i} Z^{\pm,j})_0. \end{cases} \tag{4.15}$$

Remark 7. Observe the remarkable structure in (4.10)-(4.12) and (4.15) that the "+" variables never interact nonlinearly with the "−" variables. Physically speaking, all nonlinear interactions are between wavepackets transported in opposite directions along the magnetic field lines. On $\mathbb{R}^3$, this amounts to a dispersive effect whereby the waves themselves are not decaying (at least in the ideal case), but nevertheless the nonlinear terms decay as the interacting wavepackets separate in space [74]. In the language of the spacetime resonance method for nonlinear wave equations (see, e.g., [56–58]), this structure means that the nonlinearity is *space non-resonant* uniformly in frequency on $\mathbb{R}^3$. In our setting of $\mathbb{T}\times\mathbb{R}\times\mathbb{T}$, the effect of the relative transport is to provide time oscillations in all nonlinear interactions where the function containing $T_{\pm 2\alpha}^t$ has a nonzero X or Z frequency. For such interactions it is possible to integrate by parts in time, however, we do not know how to use this structure to obtain $\gamma < 1$ because the regularity losses discussed in Section 3.4 limit the possible gain in the high norm estimates.

4.3.1 Shorthands

It will be useful to define some shorthands for the various terms appearing in the equations of the previous section. For concreteness we will only discuss the terms in the form that they appear in (4.10)-(4.12) for the "+" variables. For the linear terms in the equation for $F^{+,2}$ we write

$$\begin{aligned} \mathrm{LS} &= -\partial^L_{XY}\Delta_L^{-1}F^{+,2} && (\text{"linear stretch"}),\\ \mathrm{OLS} &= T^t_{2\alpha}\partial^L_{XY}\Delta_L^{-1}F^{-,2} && (\text{"oscillating linear stretch"}), \end{aligned}$$

and for the linear terms in the equation for $F^{+,\beta}$, $\beta \in \{1, 3\}$, we write

$$\begin{aligned} \mathrm{LU} &= -T^t_{2\alpha}F^{-,2} && (\text{"lift-up"}),\\ \mathrm{LS} &= -2\partial^L_{XY}\Delta_L^{-1}F^{+,\beta} && (\text{"linear stretch"}),\\ \mathrm{LP1} &= \partial^L_{X\beta}\Delta_L^{-1}F^{+,2} && (\text{"linear pressure"}),\\ \mathrm{LP2} &= T^t_{2\alpha}\partial^L_{X\beta}\Delta_L^{-1}F^{-,2} && (\text{"linear pressure"}). \end{aligned}$$

We denote the four types of nonlinear terms in the equation for $F^{+,\beta}$, $\beta \in \{1,2,3\}$ by

$$\text{NLT}(j, s_1, s_2) = -T^t_{2\alpha} Z^{-,j}_{s_1} \partial^L_j F^{+,\beta}_{s_2} \qquad \text{(``nonlinear transport''),}$$

$$\text{NLS1}(j, s_1, s_2) = -T^t_{2\alpha} F^{-,j}_{s_1} \partial^L_j Z^{+,\beta}_{s_2} \qquad \text{(``nonlinear stretch''),}$$

$$\text{NLS2}(i, j, s_1, s_2) = -T^t_{2\alpha} \partial^L_i Z^{-,j}_{s_1} \partial^L_{ij} Z^{+,\beta}_{s_2} \qquad \text{(``nonlinear stretch''),}$$

$$\text{NLP}(i, j, s_1, s_2) = \partial_\beta (T^t_{2\alpha} \partial^L_j Z^{-,i}_{s_1} \partial^L_i Z^{+,j}_{s_2}) \qquad \text{(``nonlinear pressure'').}$$

In the above, $i, j \in \{1,2,3\}$ and s_1, s_2 can be 0 or $\neq$. The generalization of the shorthands above to the other equations is clear, except for perhaps in the equations for Q^2 and H^2 since they have additional nonlinear terms. In this case we simply denote indifferently, for example, $U \cdot \nabla_L H^2$ and $B \cdot \nabla_L Q^2$ as nonlinear transport terms in the equation for H^2. This will not cause any confusion in the proof. We use superscripts HL and LH to denote the two pieces of a term corresponding to the paraproduct decomposition defined in Section 4.2. We will also abuse notation slightly and use the same shorthands above to denote a term's contribution to an energy estimate. For example, in an H^s energy estimate for $F^{+,\beta}$ we write one of the contributions from the nonlinear transport term as

$$\text{NLT}^{\text{HL}}(j, s_1, s_2) = -\int_{t_1}^{t_2} \int \langle \nabla \rangle^s F^{+,\beta} \langle \nabla \rangle^s \left((T^t_{2\alpha} Z^{-,j}_{s_1})^{\text{Hi}} (\partial^L_j F^{+,\beta}_{s_2})^{\text{Lo}} \right) dV dt.$$

When we do not indicate s_1, s_2, or j in the nonlinear shorthands we simply mean the term without any restrictions on the indices or frequency interactions. For example, we write $\text{NLT}(j) = -T^t_{2\alpha} Z^{-,j} \partial^L_j F^{+,\beta}$ and $\text{NLT} = -T^t_{2\alpha} Z^- \cdot \nabla_L F^{+,\beta}$.

4.4 Bootstrap argument

In this section we set up the bootstrap argument that we will use to prove Theorem 3.3.

4.4.1 Local well-posedness

We begin with a statement on the local well-posedness of (2.3).

Lemma 4.3. *Let $s > 7/2$ and fix any $5/2 < s_0 < s$. Suppose that $\mu = \nu > 0$ and that $u_{in}, b_{in} \in H^s$ are divergence free. There exists $T_0(\|(u_{in}, b_{in})\|_{H^{s_0}}) > 0$ (in particular, independent of ν) with $\lim_{x\to 0} T_0(x) = \infty$ and a unique classical solution $(u, b) \in C([0, T_0]; H^s)$ to (2.3). Moreover, there exists $T^* > T_0$ and a unique maximally extended classical solution $(u, b) \in C([0, T^*); H^s)$, and if $T^* < \infty$, then*

$$\limsup_{t\to T^*} \|(u(t), b(t))\|_{H^{s_0}} = \infty.$$

For all $0 < \tau_1 < \tau_2 < T^*$, the maximally extended solution satisfies

$$(u, b) \in C([\tau_1, \tau_2]; H^{s'}) \quad \forall\, s' \geq 0, \tag{4.16a}$$

$$\|\nabla u\|_{L^2(0,\tau_2;H^s)} + \|\nabla b\|_{L^2(0,\tau_2;H^s)} < \infty. \tag{4.16b}$$

Sketch of proof. The (X, Y, Z) coordinates defined at the beginning of Chapter 3 are equivalent to the (x, y, z) coordinates for short times in the sense that for any function

$g(t,x,y,z) = G(t,X,Y,Z)$ there holds

$$\frac{1}{(1+t+t^2)^{s'}}\|G(t)\|^2_{H^{s'}_{X,Y,Z}} \leq \|g(t)\|^2_{H^{s'}_{x,y,z}} \leq (1+t+t^2)^{s'}\|G(t)\|^2_{H^{s'}_{X,Y,Z}} \tag{4.17}$$

for all $s' \geq 0$. Hence, it suffices to prove Lemma 4.3 in the new variables. Switching to the new coordinate system and using the unknowns defined in (3.6) and (3.7), the system (2.3) can be written as

$$\begin{aligned}\partial_t Z^\pm + \mathbb{P}_t\left((T^t_{\pm 2\alpha}Z^\mp)\cdot\nabla_L Z^\pm\right) + \begin{pmatrix} T^t_{\pm 2\alpha}Z^{\mp,2} \\ 0 \\ 0 \end{pmatrix} \\ - \nabla_L\Delta_L^{-1}\partial_X(Z^{\pm,2} + T^t_{\pm 2\alpha}Z^{\mp,2}) = \nu\Delta_L Z^\pm,\end{aligned} \tag{4.18}$$

where $\mathbb{P}_t$ denotes the projection onto ∇_L divergence free vector fields. Since $\mathbb{P}_t$ satisfies the same properties as the standard Leray projector and $T^t_{\pm 2\alpha}$ is bounded on any $H^{s'}$ space and commutes with ∇_L, we see that (4.18) has the same energy structure in the nonlinear term as the Navier-Stokes equations. A calculation involving a commutator estimate then yields, for any $s' \geq s_0$, the a priori bound

$$\frac{d}{dt}\|Z(t)\|^2_{H^{s'}} + \nu\|\nabla_L Z(t)\|^2_{H^{s'}} \lesssim \|Z(t)\|^2_{H^{s'}} + (1+t)\|Z(t)\|_{H^{s_0}}\|Z(t)\|^2_{H^{s'}}, \tag{4.19}$$

where we have defined the $\mathbb{R}^6$ valued function $Z = (Z^+, Z^-)$. Without loss of generality we can suppose that $t \lesssim 1$, and so estimate (4.19) with $s' = s_0$ implies that for some

$C > 0$ there holds

$$\|Z(t)\|_{H^{s_0}} \leq \frac{\|Z(0)\|_{H^{s_0}} e^{Ct}}{1 - \|Z(0)\|_{H^{s_0}}(e^{Ct} - 1)}. \tag{4.20}$$

The existence of a unique classical solution $Z \in C([0, T_0]; H^s)$ to (4.18) for

$$T_0 \gtrsim \log\left(1 + \frac{1}{2\|Z(0)\|_{H^{s_0}}}\right)$$

then follows by the classical energy methods used in [100] to prove local existence in Sobolev spaces for the 3D Navier-Stokes equations. □

A consequence of Lemma 4.3, and in particular (4.19), is that under the assumption on the initial data in Theorem 3.3 there exists $0 < t_1 \ll 1$ independent of ν such that for c_0 sufficiently small there holds

$$\max\left(\|Z^{\pm}\|_{L^\infty(0,2t_1;H^{N+2})},\ \nu^{1/2}\|\nabla_L Z^{\pm}\|_{L^2(0,2t_1;H^{N+2})}\right) \leq 3\epsilon. \tag{4.21}$$

4.4.2 Bootstrap hypotheses and setting up their continuation

Recall the definitions of N, N', and N'' from Theorem 3.3, and let $\tilde{N} = N' + 2 + n$. In what follows we use the shorthand notations

$$\begin{aligned}
\lambda(t) &= e^{\delta\nu^{1/3}t}, \\
A(t,k,\eta,\ell) &= mM\lambda, \\
\tilde{A}(t,k,\eta,\ell) &= \tilde{m}M\lambda, \\
J(t,k,\eta,\ell) &= m^{1/2}M\lambda, \\
\tilde{J}(t,k,\eta,\ell) &= \langle t\rangle^{-1/2} J,
\end{aligned}$$

where $\delta > 0$ is a small number to be fixed later.

Recall the definition of t_1 from (4.21) above. Let $t_2 \geq t_1$ be the maximal time such that the following estimates hold on $[t_1, t_2]$:

- the high norm bounds:

$$\|\tilde{A}F_{\neq}^{\pm,1}\|_{L^\infty H^N} + \nu^{1/2}\|\tilde{A}\nabla_L F_{\neq}^{\pm,1}\|_{L^2H^N} + \|\tilde{m}\lambda\sqrt{-\dot{M}M}F_{\neq}^{\pm,1}\|_{L^2H^N} \leq 30C_0\epsilon\nu^{-1/3}, \tag{4.22a}$$

$$\|AF_{\neq}^{\pm,3}\|_{L^\infty H^N} + \nu^{1/2}\|A\nabla_L F_{\neq}^{\pm,3}\|_{L^2H^N} + \|m\lambda\sqrt{-\dot{M}M}F_{\neq}^{\pm,3}\|_{L^2H^N} \leq 30C_0\epsilon\nu^{-1/3}, \tag{4.22b}$$

$$\|AH_{\neq}^2\|_{L^\infty H^N} + \nu^{1/2}\|A\nabla_L H_{\neq}^2\|_{L^2H^N} + \|m\lambda\sqrt{-\dot{M}M}H_{\neq}^2\|_{L^2H^N} \leq 30\epsilon, \tag{4.22c}$$

$$\|AQ_{\neq}^2\|_{L^\infty H^N} + \nu^{1/2}\|A\nabla_L Q_{\neq}^2\|_{L^2H^N} + \|m\lambda\sqrt{-\dot{M}M}Q_{\neq}^2\|_{L^2H^N} \leq 30\epsilon, \tag{4.22d}$$

$$\|(H_0,Q_0)\|_{L^\infty H^N} + \nu^{1/2}\|\nabla(H_0,Q_0)\|_{L^2H^N} \leq 30\epsilon\nu^{-1/3}; \tag{4.22e}$$

- the intermediate norm bounds:

$$\|\tilde{J}F_{\neq}^{\pm,2}\|_{L^\infty H^{\tilde{N}}} + \nu^{1/2}\|\tilde{J}\nabla_L F_{\neq}^{\pm,2}\|_{L^2H^{\tilde{N}}} + \|\langle t\rangle^{-1/2} m^{1/2}\lambda\sqrt{-\dot{M}M}F_{\neq}^{\pm,2}\|_{L^2H^{\tilde{N}}} \leq 30\epsilon, \tag{4.23a}$$

$$\|JF_{\neq}^{\pm,2}\|_{L^\infty H^{N'}} + \nu^{1/2}\|J\nabla_L F_{\neq}^{\pm,2}\|_{L^2H^{N'}} + \|m^{1/2}\lambda\sqrt{-\dot{M}M}F_{\neq}^{\pm,2}\|_{L^2H^{N'}} \leq 30\epsilon; \tag{4.23b}$$

- the low norm bounds:

$$\|\tilde{A}F_{\neq}^{\pm,1}\|_{L^\infty H^{N''}} + \nu^{1/2}\|\tilde{A}\nabla_L F_{\neq}^{\pm,1}\|_{L^2H^{N''}} + \|\tilde{m}\lambda\sqrt{-\dot{M}M}F_{\neq}^{\pm,1}\|_{L^2H^{N''}} \leq 30C_0\epsilon, \tag{4.24a}$$

$$\|AF_{\neq}^{\pm,3}\|_{L^\infty H^{N''}} + \nu^{1/2}\|A\nabla_L F_{\neq}^{\pm,3}\|_{L^2H^{N''}} + \|m\lambda\sqrt{-\dot{M}M}F_{\neq}^{\pm,3}\|_{L^2H^{N''}} \leq 30C_0\epsilon; \tag{4.24b}$$

- the zero mode bounds on the velocity and magnetic field:

$$\|(u_0,b_0)\|_{L^\infty H^N} + \nu^{1/2}\|\nabla(u_0,b_0)\|_{L^2H^N} \leq 30\epsilon. \tag{4.25}$$

Here, $C_0 \geq 1$ is a constant to be fixed by the proof. We refer to the list of inequalities above as the bootstrap hypotheses. Henceforth, all norms will be taken on $[t_1, t_2]$.

We claim that $t_2 \geq 2t_1$ if t_1 is chosen sufficiently small (still uniformly in ν). Indeed, this follows from (4.21), $|\Delta_L| \leq (1+t+t^2)|\Delta|$, $\max(m, \tilde{m}, M) \leq 1$, $\sqrt{-\dot{M}M} \leq 2$, and the fact that $\lambda(t)$ is continuous and equal to unity at $t = 0$. The plan is then to prove that $t_2 = \infty$ under the assumptions of Theorem 3.3. Since all of the norms in (4.22)-(4.25) take values continuously in time, it suffices to prove the following proposition.

Proposition 4.4. *Suppose that $\mu = \nu \in (0, 1]$, $\sigma \in \mathbb{R}^+ \setminus \mathbb{Q}$ satisfies (3.40),*

$$\|(u_{in}, b_{in})\|_{H^{N+2}} = \epsilon \leq c_0\nu,$$

and the estimates in (4.22)-(4.25) hold on $[t_1, t_2]$ for some $t_2 > t_1 > 0$. Suppose further that $C_0 \geq 1$ is fixed sufficiently large, $0 < \delta \leq 1$ is fixed sufficiently small, and $t_1 < t_0$ for a sufficiently small universal constant t_0. Then, for $\alpha > 0$ sufficiently large and $c_0 = c_0(N) > 0$ sufficiently small, the estimates in (4.22)-(4.25) hold on $[t_1, t_2]$ with all the occurrences of "30" replaced by "20."

The proof of Proposition 4.4 is carried out in Chapter 5 and the fact that Proposition 4.4 implies Theorem 3.3 is proven below in Lemma 4.5.

Remark 8. The purpose of defining the bootstrap hypotheses on $[t_1, t_2]$ instead of $[0, t_2]$ is to ensure that the classical solution we perform our calculations with satisfies $Z^\pm \in C([t_1, t_2]; H^{s'})$ for every $s' \geq 0$, which follows from (4.16a).

Remark 9. In light of the discussion just before Remark 4, the general structure of the bootstrap hypotheses should be expected. Perhaps the most subtle aspect is the inclusion of $\tilde{m}$ in the norm for $F^1_{\neq}$. Physically, this represents allowing the frequencies of $F^1_{\neq}$ to grow indefinitely after the critical time. This enables us to use integration by parts in time to control the lift-up effect in the low norm with no losses. The key inequality here is (4.6f). It is also worth pointing out that the use of two intermediate norms is more of a technical detail than something deep, and arises essentially from the same scaling that forces one to take $\theta > 0$ in (3.27).

4.4.3 Choice of constants

Recall the definitions of c and c_0 from the statement of Theorem 3.3. In the proof the various constants will be fixed as follows. We first fix $C_0 \geq 1$ to be a sufficiently large universal constant and $0 < \delta \leq 1$ to be sufficiently small. Then, α and c_0 are chosen to satisfy $\alpha \gg 1 + C_0/c$ and $c_0 \ll (\delta/C_0)^p$ for p sufficiently large. We pick t_0 in Proposition 4.4 such that $e^{2t_0}(1 + t_0 + t_0^2)^2 \leq 2$.

4.4.4 Estimates following from the bootstrap hypotheses

Now we prove a lemma that details the enhanced dissipation and inviscid damping estimates that follow immediately from the bootstrap hypotheses.

Lemma 4.5. *Let G denote either Q or H, and V denote either U or B. Under the bootstrap hypotheses the following estimates hold on $[t_1, t_2]$:*

- *the enhanced dissipation of $Q_{\neq}$ and $H_{\neq}$:*

$$\nu^{1/3}\|\tilde{A}G^1_{\neq}\|_{L^2H^N} + \|\tilde{A}G^1_{\neq}\|_{L^2H^{N''}} \lesssim \epsilon\nu^{-1/6}, \tag{4.26a}$$

$$\|AG^2_{\neq}\|_{L^2H^N} + \|JG^2_{\neq}\|_{L^2H^{N'}} \lesssim \epsilon\nu^{-1/6}, \tag{4.26b}$$

$$\nu^{1/3}\|AG^3_{\neq}\|_{L^2H^N} + \|AG^3_{\neq}\|_{L^2H^{N''}} \lesssim \epsilon\nu^{-1/6}; \tag{4.26c}$$

- *the bounds on $U_{\neq}$ and $B_{\neq}$, denoting $j \in \{1,3\}$:*

$$\begin{aligned} \nu^{1/3}\|e^{\delta\nu^{1/3}t}\Delta_{X,Z}V^j_{\neq}\|_{L^\infty H^N} + \nu^{5/6}\|e^{\delta\nu^{1/3}t}\nabla_L\Delta_{X,Z}V^j_{\neq}\|_{L^2H^N} \\ + \nu^{1/2}\|e^{\delta\nu^{1/3}t}\Delta_{X,Z}V^j_{\neq}\|_{L^2H^N} \lesssim \epsilon, \end{aligned} \tag{4.27a}$$

$$\begin{aligned} \|e^{\delta\nu^{1/3}t}\Delta_{X,Z}V^j_{\neq}\|_{L^\infty H^{N''}} + \nu^{1/2}\|e^{\delta\nu^{1/3}t}\nabla_L\Delta_{X,Z}V^j_{\neq}\|_{L^2H^{N''}} \\ + \nu^{1/6}\|e^{\delta\nu^{1/3}t}\Delta_{X,Z}V^j_{\neq}\|_{L^2H^{N''}} \lesssim \epsilon, \end{aligned} \tag{4.27b}$$

$$\begin{aligned} \|e^{\delta\nu^{1/3}t}\Delta_{X,Z}V^2_{\neq}\|_{L^\infty H^N} + \nu^{1/2}\|e^{\delta\nu^{1/3}t}\nabla_L\Delta_{X,Z}V^2_{\neq}\|_{L^2H^N} \\ + \nu^{1/6}\|e^{\delta\nu^{1/3}t}\Delta_{X,Z}V^2_{\neq}\|_{L^2H^N} \lesssim \epsilon; \end{aligned} \tag{4.27c}$$

- *the inviscid damping of $U^2_{\neq}$ and $B^2_{\neq}$:*

$$\|e^{\delta\nu^{1/3}t}\nabla_{X,Z}V^2_{\neq}\|_{L^2H^{N'}} + \|e^{\delta\nu^{1/3}t}\langle t\rangle\nabla_{X,Z}V^2_{\neq}\|_{L^\infty H^{N'-1}} \lesssim \epsilon. \tag{4.28}$$

Proof. First consider the estimates in (4.26). Observe that for any $s \geq 0$ and $G \in H^{s+1}$ we have, by Lemma 4.2,

$$\nu^{1/6}\|G_{\neq}\|_{H^s} \lesssim \nu^{1/2}\|\nabla_L G_{\neq}\|_{H^s} + \|\sqrt{-\dot{M}M}G_{\neq}\|_{H^s}.$$

The inequalities in (4.26) then follow immediately from the bootstrap hypotheses. The estimates in (4.27) follow similarly after employing also (4.6b). Now we turn to the inviscid damping estimates. For the first term in (4.28) we use that $|\nabla_{X,Z}| \lesssim m^{1/2}|\nabla_L| \lesssim$

$m^{1/2}\sqrt{-\dot{M}M}|\Delta_L|$ (in the sense of their symbols as Fourier multipliers) to obtain

$$\|e^{\delta\nu^{1/3}t}\nabla_{X,Z}V^2_{\neq}\|_{L^2H^{N'}} \lesssim \|m^{1/2}\lambda\sqrt{-\dot{M}M}G^2_{\neq}\|_{L^2H^{N'}},$$

and hence the desired inequality follows from the bootstrap hypothesis (4.23b). For the other term in (4.28) we use $|\nabla_{X,Z}| \lesssim m^{1/2}|\nabla_L|^{-1}|\Delta_L|$ along with $|\nabla_L|^{-1} \lesssim \langle t\rangle^{-1}\langle\nabla\rangle$ to derive

$$\|e^{\delta\nu^{1/3}t}\nabla_{X,Z}V^2_{\neq}\|_{H^{N'-1}} \lesssim \|\lambda m^{1/2}|\nabla_L|^{-1}G^2_{\neq}\|_{H^{N'-1}} \lesssim \langle t\rangle^{-1}\|JG^2_{\neq}\|_{H^{N'}},$$

and so the result follows again from (4.23b). □

We will use the enhanced dissipation estimates in Lemma 4.5 so frequently throughout the proof that we will typically do so without any remark.

Chapter 5: Energy Estimates

In this chapter we carry out the energy estimates needed to prove Proposition 4.4. This involves continuing (more precisely, *improving*) each of the bootstrap estimates from Section 4.4.2. The high norm estimates are carried out in Sections 5.1-5.5, the intermediate norm estimates in Sections 5.6 and 5.7, the low norm estimates in Section 5.8, and the zero mode velocity estimates in Section 5.9.

Before proceeding to the estimates we establish some simplifying notation to keep the formulas as concise as possible. As noted in Remark 7, our proof does not rely on the non-resonance structure of the nonlinearity. We will thus, beyond writing out the initial energy estimate, systematically drop the transport operator $T^t_{\pm 2\alpha}$ in the nonlinear terms. This is inconsequential because T^t_a commutes with derivatives and preserves norms on H^s spaces. Similarly, it is irrelevant in the nonlinearity which variables are "$+$" type and which are "$-$" type, and so in the nonlinear terms we will simply drop this superscript. Lastly, by the symmetry of (4.10)-(4.12) and (4.15) it clearly suffices to estimate only the "$+$" variables.

Remark 10. The weighted energy estimates in the following sections are best understood as being performed on the Fourier side. Note however that the multiplier m is not C^1 in time and the a priori bounds on the solution are not enough to ensure that its

Fourier transform is continuous. To make the estimates rigorous we mollify m in time, approximate the solution by using a smooth cutoff in the Y variable, and then pass to the limit. This procedure yields the same estimates as one would obtain from a formal calculation because the weak derivative of m is uniformly bounded in time and frequency. For simplicity we omit these steps in the computations.

5.1 High norm estimate of $F_{\neq}^{\pm,1}$

In this section we improve (4.22a). Recall the shorthands defined in Sec. 4.3. An energy estimate gives

$$
\begin{aligned}
&\frac{1}{2}\|\tilde{A}F_{\neq}^{+,1}(t_2)\|_{H^N}^2 + \nu\|\tilde{A}\nabla_L F_{\neq}^{+,1}\|_{L^2H^N}^2 + \|\tilde{m}\lambda\sqrt{-\dot{M}M}F_{\neq}^{+,1}\|_{L^2H^N}^2 \\
&+ \|M\lambda\sqrt{-\dot{\tilde{m}}\tilde{m}}F_{\neq}^{+,1}\|_{L^2H^N}^2 = \frac{1}{2}\|\tilde{A}F_{\neq}^{+,1}(t_1)\|_{H^N}^2 - \int_{t_1}^{t_2}\left\langle \tilde{A}F_{\neq}^{+,1}, T_{2\alpha}^t\tilde{A}F_{\neq}^{-,2}\right\rangle_{H^N} dt \\
&- 2\int_{t_1}^{t_2}\left\langle \tilde{A}F_{\neq}^{+,1}, \tilde{A}\partial_{XY}^L\Delta_L^{-1}F_{\neq}^{+,1}\right\rangle_{H^N} dt + \int_{t_1}^{t_2}\left\langle \tilde{A}F_{\neq}^{+,1}, \tilde{A}\partial_{XX}\Delta_L^{-1}F_{\neq}^{+,2}\right\rangle_{H^N} dt \\
&+ \int_{t_1}^{t_2}\left\langle \tilde{A}F_{\neq}^{+,1}, \tilde{A}\partial_{XX}T_{2\alpha}^t\Delta_L^{-1}F_{\neq}^{-,2}\right\rangle_{H^N} dt + \delta\nu^{1/3}\int_{t_1}^{t_2}\|\tilde{A}F_{\neq}^{+,1}\|_{H^N}^2 dt \\
&+ \int_{t_1}^{t_2}\left\langle \tilde{A}F_{\neq}^{+,1}, \partial_X\tilde{A}(\partial_j^L T_{2\alpha}^t Z^{-,i}\partial_i^L Z^{+,j})\right\rangle_{H^N} dt \\
&- \int_{t_1}^{t_2}\left\langle \tilde{A}F_{\neq}^{+,1}, \tilde{A}(T_{2\alpha}^t Z^{-}\cdot\nabla_L F^{+,1})\right\rangle_{H^N} dt \\
&- \int_{t_1}^{t_2}\left\langle \tilde{A}F_{\neq}^{+,1}, \tilde{A}(T_{2\alpha}^t F^{-}\cdot\nabla_L Z^{+,1})\right\rangle_{H^N} dt \\
&- 2\int_{t_1}^{t_2}\left\langle \tilde{A}F_{\neq}^{+,1}, \tilde{A}(T_{2\alpha}^t\partial_i^L Z^{-,j}\partial_{ij}^L Z^{+,1})\right\rangle_{H^N} dt \\
&= \frac{1}{2}\|\tilde{A}F_{\neq}^{+,1}(t_1)\|_{H^N}^2 + \mathrm{LU} + \mathrm{LS} + \mathrm{LP1} + \mathrm{LP2} + \mathrm{L}_\lambda \\
&\qquad + \mathrm{NLP} + \mathrm{NLT} + \mathrm{NLS1} + \mathrm{NLS2},
\end{aligned}
$$

where we have introduced the additional shorthand

$$\mathrm{L}_\lambda = \delta \nu^{1/3} \int_{t_1}^{t_2} \|\tilde{A} F_{\neq}^{+,1}\|_{H^N}^2 dt$$

for the term where the time derivative lands on the exponentially growing multiplier λ. We will continue to use this shorthand throughout for the analogous term in future estimates. Note that each abbreviation in the main energy estimate stands for only one integral sign. The choice of t_0 in Section 4.4.3, along with $\|Z^{\pm}(t_1)\|_{H^{N+2}}^2 \leq 9\epsilon^2$, guarantees that $\|\tilde{A} F_{\neq}^{+,1}(t_1)\|_{H^N}^2 \leq 18\epsilon^2$, which is consistent with Proposition 4.4.

5.1.1 Lift-up term

By Cauchy-Schwarz and $|\tilde{A}| \leq |A|$ we have

$$|\mathrm{LU}| \leq \|\tilde{A} F_{\neq}^{+,1}\|_{L^2 H^N} \|A F_{\neq}^{-,2}\|_{L^2 H^N} \lesssim \epsilon^2 C_0 \nu^{-1/2} \nu^{-1/6} = C_0^{-1} (\epsilon C_0 \nu^{-1/3})^2,$$

which suffices for C_0 chosen sufficiently large.

5.1.2 Linear stretching term

It follows by the definition of $\tilde{m}$ that

$$\mathrm{LS} \leq \|\lambda M \sqrt{-\dot{\tilde{m}} \tilde{m}} F_{\neq}^{+,1}\|_{L^2 H^N}^2,$$

and so this term is absorbed into the left-hand side of the energy estimate.

5.1.3 Linear pressure terms

Both linear pressure terms are treated similarly, and so we only consider LP1. By (4.7a) and $\nu \in (0,1]$ we have

$$
\begin{aligned}
|\mathrm{LP1}| &\lesssim \|\tilde{m}\lambda\sqrt{-\dot{M}M}F_{\neq}^{+,1}\|_{L^2H^N}\|m\lambda\sqrt{-\dot{M}M}F_{\neq}^{+,2}\|_{L^2H^N} \\
&\lesssim \epsilon^2 C_0 \nu^{-1/3} \leq C_0^{-1}(\epsilon C_0 \nu^{-1/3})^2,
\end{aligned}
$$

which is consistent for C_0 sufficiently large.

5.1.4 The term L_λ

By Lemma 4.2 it follows that for $0 < \delta < 1$ sufficiently small there holds

$$
\delta \nu^{1/3} \leq \frac{\nu}{2}(k^2 + (\eta - kt)^2 + \ell^2) - \frac{1}{2}\frac{\dot{M}}{M},
$$

from which we obtain

$$
L_\lambda \leq \frac{\nu}{2}\|\tilde{A}\nabla_L F_{\neq}^{+,1}\|_{L^2H^N}^2 + \frac{1}{2}\|\tilde{m}\lambda\sqrt{-\dot{M}M}F_{\neq}^{+,1}\|_{L^2H^N}^2.
$$

Therefore, L_λ can be absorbed into the left-hand side of the energy estimate in a way consistent with Proposition 4.4.

5.1.5 Nonlinear terms

Recall the energy estimate shorthands defined at the end of Section 4.3. We begin with the transport term

$$\mathrm{NLT} = -\int_{t_1}^{t_2}\int \tilde{A}\langle\nabla\rangle^N F^1_{\neq}\tilde{A}\langle\nabla\rangle^N (Z\cdot\nabla_L F^1)_{\neq},$$

where as described above we have dropped the $\pm$ superscripts and the relative transport between the interacting profiles, as they will not be relevant in the nonlinear terms. We will first control the interaction between the nonzero modes. Using $N'' \geq 2 > 3/2$, (4.6b), and the paraproduct decomposition defined in Section 4.2, we have, for $j \in \{1,3\}$,

$$\begin{aligned}
|\mathrm{NLT}(j,\neq,\neq)| &\leq |\mathrm{NLT}^{\mathrm{LH}}(j,\neq,\neq)| + |\mathrm{NLT}^{\mathrm{HL}}(j,\neq,\neq)| \\
&\lesssim \int_{t_1}^{t_2} \|\tilde{A}F^1_{\neq}\|_{H^N}\|\lambda Z^j_{\neq}\|_{H^{N''}}\|\tilde{m}^{1/2}\nabla_L F^1_{\neq}\|_{H^N} dt \\
&\quad + \int_{t_1}^{t_2} \|\tilde{A}F^1_{\neq}\|_{H^N}\|\lambda Z^j_{\neq}\|_{H^N}\|\tilde{m}^{1/2}\nabla_L F^1_{\neq}\|_{H^{N''}})dt \\
&\lesssim \delta^{-1}\nu^{-1/3}\|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}\|\tilde{A}F^j_{\neq}\|_{L^2H^{N''}}\|\tilde{A}\nabla_L F^1_{\neq}\|_{L^2H^N} \\
&\quad + \delta^{-1}\nu^{-1/3}\|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}\|\tilde{A}F^j_{\neq}\|_{L^2H^N}\|\tilde{A}\nabla_L F^1_{\neq}\|_{L^2H^{N''}} \\
&\lesssim \epsilon^3\delta^{-1}C_0^3\nu^{-1/3}(\nu^{-1/3}\nu^{-1/6}\nu^{-5/6} + \nu^{-1/3}\nu^{-1/2}\nu^{-1/2}) \\
&\lesssim (\epsilon C_0\nu^{-1/3})^2\epsilon\nu^{-1}C_0\delta^{-1},
\end{aligned}$$

which suffices for $\epsilon\nu^{-1} \leq c_0 \ll \delta C_0^{-1}$. In the third line above we have used (4.6d) and the fact that $t^s e^{-at} \lesssim_s a^{-s}$ for $a \geq 0$ to deduce that for any $s \geq 0$ there holds

$$1 = \tilde{m}^{-s}\lambda^{-1}\tilde{m}^s\lambda \lesssim \langle t\rangle^{2s}\lambda^{-1}\tilde{m}^s\lambda \lesssim \delta^{-2s}\nu^{-2s/3}\tilde{m}^s\lambda. \tag{5.1}$$

Using (5.1) as we have done above to compensate for the fact that $\tilde{m}$ does not satisfy (4.6c) will be done frequently throughout the proof and is always possible when estimating a term where two nonzero modes interact in the nonlinearity. When we appeal to (5.1) in what follows we will typically do so without any remark, and moreover we will not indicate that it causes the underlying constant to depend on an inverse power of δ. We will also no longer show the factors of C_0 that appear when estimating the nonlinear terms, as they are not relevant. In the case $j = 2$ we use $N' - 1 > 3/2$ and the proof of (4.28) to obtain

$$\begin{aligned}
|\text{NLT}^{\text{LH}}(2, \neq, \neq)| &\lesssim \int_{t_1}^{t_2} \|\tilde{A}F^1_{\neq}\|_{H^N}\| \langle t\rangle \lambda Z^2_{\neq}\|_{H^{N'-1}}\| \langle t\rangle^{-1} \nabla_L F^1_{\neq}\|_{H^N} dt \\
&\lesssim \int_{t_1}^{t_2} \|\tilde{A}F^1_{\neq}\|_{H^N}\|JF^2_{\neq}\|_{H^{N'}}\nu^{-1/3}\|\tilde{A}\nabla_L F^1_{\neq}\|_{H^N} dt \\
&\lesssim \|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}\|JF^2_{\neq}\|_{L^2 H^{N'}}\nu^{-1/3}\|\tilde{A}\nabla_L F^1_{\neq}\|_{L^2 H^N} \\
&\lesssim \epsilon^3\nu^{-1/3}\nu^{-1/6}\nu^{-1/3}\nu^{-5/6} = (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1}.
\end{aligned}$$

For the other term in the decomposition, we use (5.1) and the enhanced dissipation bounds in Lemma 4.5 to deduce

$$|\text{NLT}^{\text{HL}}(2,\neq,\neq)| \lesssim \|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}\|\lambda Z^2_{\neq}\|_{L^2H^N}\nu^{-2/3}\|\tilde{A}\nabla_L F^1_{\neq}\|_{L^2H^{N''}}$$
$$\lesssim \epsilon^3\nu^{-1/3}\nu^{-1/6}\nu^{-2/3}\nu^{-1/2} = (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1},$$

which suffices. Now we turn to the interaction between the zero and nonzero modes. For $\text{NLT}(\neq,0)$ we have, using (4.6b),

$$|\text{NLT}(\neq,0)| \lesssim \|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}\|\tilde{A}F_{\neq}\|_{L^2H^N}\|\nabla F^1_0\|_{L^2H^N}$$
$$\lesssim \epsilon^3\nu^{-1/3}\nu^{-1/2}\nu^{-5/6} = (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1},$$

while for $\text{NLT}(0,\neq)$ we apply (4.6e) to obtain

$$|\text{NLT}(0,\neq)| \lesssim \|\tilde{A}F^1_{\neq}\|_{L^2H^N}\|Z_0\|_{L^\infty H^{N+2}}\|\tilde{A}\nabla_L F^1_{\neq}\|_{L^2H^N}$$
$$\lesssim \epsilon^3\nu^{-1/2}\nu^{-1/3}\nu^{-5/6} = (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1}.$$

Note that we combined bootstrap hypotheses (4.22e) and (4.25) to deduce the bound on $\|Z_0\|_{H^{N+2}}$. This completes the nonlinear transport estimate.

Next we consider

$$\text{NLS1}(j) = -\int_{t_1}^{t_2}\int \tilde{A}\langle\nabla\rangle^N F^1_{\neq}\tilde{A}\langle\nabla\rangle^N (F^j\partial_j^L Z^1)_{\neq}dV\,dt$$

and

$$\text{NLS2}(i,j) = -2\int_{t_1}^{t_2}\int \tilde{A}\left\langle\nabla\right\rangle^N F^1_{\neq}\left\langle\nabla\right\rangle^N \tilde{A}(\partial_i^L Z^j \partial_{ij}^L Z^1)_{\neq} dV\,dt.$$

We start with NLS1, and as before we begin with the interaction between the nonzero modes. When $j \in \{1,3\}$, we use (5.1) and (4.6b) to obtain

$$\begin{aligned}
|\text{NLS1}(j,\neq,\neq)| &\leq |\text{NLS1}^{\text{LH}}(j,\neq,\neq)| + |\text{NLS1}^{\text{HL}}(j,\neq,\neq)| \\
&\lesssim \nu^{-2/3}\|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}\|\tilde{A}F^j_{\neq}\|_{L^2H^{N''}}\|\tilde{A}F^1_{\neq}\|_{L^2H^N} \\
&\quad + \nu^{-2/3}\|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}\|\tilde{A}F^j_{\neq}\|_{L^2H^N}\|\tilde{A}F^1_{\neq}\|_{L^2H^{N''}} \\
&\lesssim \epsilon^3\nu^{-1/3}\nu^{-2/3}\nu^{-1/6}\nu^{-1/2} = (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1},
\end{aligned}$$

which suffices. When $j=2$ we have, employing now also (4.6c),

$$\begin{aligned}
|\text{NLS1}(2,\neq,\neq)| &\leq |\text{NLS1}^{\text{LH}}(2,\neq,\neq)| + |\text{NLS1}^{\text{HL}}(2,\neq,\neq)| \\
&\lesssim \|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}(\|\lambda F^2_{\neq}\|_{L^2H^{N'}}\|\partial_Y^L Z^1_{\neq}\|_{L^2H^N} + \|\lambda F^2_{\neq}\|_{L^2H^N}\|\partial_Y^L Z^1_{\neq}\|_{L^2H^{N''}}) \\
&\lesssim \|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}\nu^{-1/3}\|JF^2_{\neq}\|_{L^2H^{N'}}\nu^{-1/3}\|\tilde{A}F^1_{\neq}\|_{L^2H^N} \\
&\quad + \|\tilde{A}F^1_{\neq}\|_{L^\infty H^N}\nu^{-2/3}\|AF^2_{\neq}\|_{L^2H^N}\nu^{-1/3}\|\tilde{A}F^1_{\neq}\|_{L^2H^{N''}} \\
&\lesssim \epsilon^3\nu^{-1/3}(\nu^{-1/3}\nu^{-1/6}\nu^{-1/3}\nu^{-1/2} + \nu^{-2/3}\nu^{-1/6}\nu^{-1/3}\nu^{-1/6}) \lesssim (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1}.
\end{aligned}$$

Now we turn to $\text{NLS1}(j,\neq,0)$. First, notice that it is only nonzero when $j\in\{2,3\}$. This is because the nonlinear part of $\text{NLS1}(1,\neq,0)$ is $F^1_{\neq}\partial_X Z^1_0$, which is zero since Z^1_0 does not depend on X. For $j \in \{2,3\}$ we control the term with (4.6e), $\tilde{m} \le m$, and (4.6c):

$$\begin{aligned}
|\text{NLS1}(j, \neq, 0)| &\lesssim \|\tilde{A}F^1_{\neq}\|_{L^2H^N} \|\lambda m^{1/2} F^j_{\neq}\|_{L^2H^N} \|Z^1_0\|_{L^\infty H^{N+2}} \\
&\lesssim \|\tilde{A}F^1_{\neq}\|_{L^2H^N} \nu^{-1/3} \|AF^j_{\neq}\|_{L^2H^N} \|Z^1_0\|_{L^\infty H^{N+2}} \\
&\lesssim \epsilon^3 \nu^{-1/2}\nu^{-1/3}\nu^{-1/2}\nu^{-1/3} = (\epsilon\nu^{-1/3})^2 \epsilon\nu^{-1}.
\end{aligned}$$

Lastly, using (4.6b) we have

$$\begin{aligned}
|\text{NLS1}(0, \neq)| &\lesssim \|\tilde{A}F^1_{\neq}\|_{L^2H^N} \|F_0\|_{L^\infty H^N} \|\tilde{A}\nabla_L F^1_{\neq}\|_{L^2H^N} \\
&\lesssim \epsilon^3 \nu^{-1/2}\nu^{-1/3}\nu^{-5/6} = (\epsilon\nu^{-1/3})^2 \epsilon\nu^{-1}.
\end{aligned}$$

This completes the estimate of NLS1. For NLS2, the interactions between the nonzero modes can be treated in the same manner as they were for NLS1 by using (5.1), (4.6b), and paraproduct decompositions. We thus skip these terms for the sake of brevity. Turning then to the interaction between the zero and the nonzero modes, we have

$$\begin{aligned}
|\text{NLS2}(i, j, \neq, 0)| &\lesssim \|\tilde{A}F^1_{\neq}\|_{L^2H^N} \nu^{-1/3} \|AF^j_{\neq}\|_{L^2H^N} \|Z^1_0\|_{L^\infty H^{N+2}} \\
&\lesssim \epsilon^3 \nu^{-1/2}\nu^{-1/3}\nu^{-1/2}\nu^{-1/3} = (\epsilon\nu^{-1/3})^2 \epsilon\nu^{-1},
\end{aligned}$$

where we noted that we can apply (4.6c) since the term is only nonzero for $j \neq 1$. Lastly, we have, using (4.6e) and (4.6b),

$$\begin{aligned}|\text{NLS2}(i,j,0,\neq)| &\lesssim \|\tilde{A}F^1_{\neq}\|_{L^2H^N}\|Z_0\|_{L^\infty H^{N+2}}\|\tilde{m}^{1/2}\lambda F^1_{\neq}\|_{L^2H^N}\\ &\lesssim \|\tilde{A}F^1_{\neq}\|_{L^2H^N}\|Z_0\|_{L^\infty H^{N+2}}\|\tilde{A}\nabla_L F^1_{\neq}\|_{L^2H^N}\\ &\lesssim \epsilon^3\nu^{-1/2}\nu^{-1/3}\nu^{-5/6} = (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1},\end{aligned}$$

which completes the treatment of the nonlinear stretching terms.

Now we consider the nonlinear pressure, which by an integration by parts can be written

$$\begin{aligned}\text{NLP}(i,j) &= \int_{t_1}^{t_2}\int \langle\nabla\rangle^N \tilde{A}F^1_{\neq}\langle\nabla\rangle^N \tilde{A}\partial_X(\partial^L_j Z^i\partial^L_i Z^j)_{\neq}dV\,dt\\ &= -\int_{t_1}^{t_2}\int \partial_X\langle\nabla\rangle^N \tilde{A}F^1_{\neq}\langle\nabla\rangle^N \tilde{A}(\partial^L_j Z^i\partial^L_i Z^j)_{\neq}dV\,dt.\end{aligned}$$

There are three distinct cases to consider: $i,j\in\{1,3\}$, $i=2$ and $j\neq 2$, and $i=j=2$. For the first case, we notice that due to the symmetry in the bootstrap hypotheses for F^1 and F^3 we can assume without loss of generality that Z^j has a nonzero X-frequency. Then, we have the estimate

$$\begin{aligned}|\text{NLP}(i\in\{1,3\},j\in\{1,3\})| &\lesssim \|\nabla_L\tilde{A}F^1_{\neq}\|_{L^2H^N}\|\partial^L_j Z^i\|_{L^\infty H^N}\|\tilde{A}F^j_{\neq}\|_{L^2H^N}\\ &\lesssim \|\nabla_L\tilde{A}F^1_{\neq}\|_{L^2H^N}\left(\|\tilde{A}F^i_{\neq}\|_{L^\infty H^N}+\|Z^i_0\|_{L^\infty H^{N+2}}\right)\|\tilde{A}F^j_{\neq}\|_{L^2H^N}\\ &\lesssim \epsilon^3\nu^{-5/6}\nu^{-1/3}\nu^{-1/2} = (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1}.\end{aligned}$$

Turning now to the case $i=2$ and $j\neq 2$, we first consider when $j=3$. Splitting the term between the different frequency interactions gives

$$\begin{aligned}|\mathrm{NLP}(2,3)| &\leq |\mathrm{NLP}(2,3,0,\neq)| + |\mathrm{NLP}(2,3,\neq,0)| + |\mathrm{NLP}(2,3,\neq,\neq)| \\ &:= |\mathrm{NLP}(2,3,0,\neq)| + \mathrm{NLP}(2,3,\neq,\cdot).\end{aligned}$$

For NLP$(2,3,0,\neq)$ we use (4.6e) and (4.6b) to obtain

$$\begin{aligned}|\mathrm{NLP}(2,3,0,\neq)| &\lesssim \|\nabla_L \tilde{A} F^1_{\neq}\|_{L^2 H^N} \|\partial_Z Z^2_0\|_{L^\infty H^{N+1}} \|\lambda \tilde{m}^{1/2} \partial^L_Y Z^3_{\neq}\|_{L^2 H^N} \\ &\lesssim \|\nabla_L \tilde{A} F^1_{\neq}\|_{L^2 H^N} \|Z^2_0\|_{L^\infty H^{N+2}} \|A F^3_{\neq}\|_{L^2 H^N} \\ &\lesssim \epsilon^3 \nu^{-5/6} \nu^{-1/3} \nu^{-1/2} = \epsilon \nu^{-1} (\epsilon \nu^{-1/3})^2.\end{aligned}$$

We then estimate the other two pieces using (4.6c):

$$\begin{aligned}\mathrm{NLP}(2,3,\neq,\cdot) &\lesssim \|\nabla_L \tilde{A} F^1_{\neq}\|_{L^2 H^N} \|A F^2_{\neq}\|_{L^2 H^N} \|Z^3_0\|_{L^\infty H^{N+2}} \\ &\quad + \nu^{-1/3} \|\nabla_L \tilde{A} F^1_{\neq}\|_{L^2 H^N} \|A F^3_{\neq}\|_{L^\infty H^N} \\ &\lesssim \epsilon^3 \nu^{-5/6} \nu^{-1/6} \nu^{-2/3} = (\epsilon \nu^{-1/3})^2 \epsilon \nu^{-1}.\end{aligned}$$

A similar estimate holds for NLP$(2,1)$ due to the fact that NLP$(2,1,0,\neq) = 0$. The only variation is that we must use (5.1) for the interaction between the nonzero modes because $\tilde{m}$ does not satisfy (4.6c). We omit the details. Lastly, we consider the case $i = j = 2$, for which we have the estimate

$$\begin{aligned}|\mathrm{NLP}(2,2)| &\lesssim \|\nabla_L \tilde{A} F^1_{\neq}\|_{L^2 H^N} \|\lambda \partial_Y^L Z^2_{\neq}\|_{L^2 H^N} \|\partial_Y^L Z^2\|_{L^\infty H^N} \\ &\lesssim \|\nabla_L \tilde{A} F^1_{\neq}\|_{L^2 H^N} \nu^{-1/3} \|A F^2_{\neq}\|_{L^2 H^N} (\|Z^2_0\|_{L^\infty H^{N+2}} + \nu^{-1/3} \|A F^2_{\neq}\|_{L^\infty H^N}) \\ &\lesssim \epsilon^3 \nu^{-5/6} \nu^{-1/3} \nu^{-1/6} \nu^{-1/3} = (\epsilon \nu^{-1/3})^2 \epsilon \nu^{-1},\end{aligned}$$

which suffices and completes the estimate of $F^{\pm,1}_{\neq}$ in the high norm.

5.2 High norm estimate of $Q^2_{\neq}$ and $H^2_{\neq}$

In this section we improve (4.22c) and (4.22d). In particular, we need to verify that with just $\epsilon \ll \nu$ the high norm controls on $H^2_{\neq}$ and $Q^2_{\neq}$, unlike (4.22a) and (4.22b), do not need to lose the factor $\nu^{-1/3}$. This will be possible because in the low norm $H^2_{\neq}$ and $Q^2_{\neq}$ only grow linearly in time, as opposed to quadratically like $F^1_{\neq}$ and $F^3_{\neq}$. Exploiting the gain of one power of $\langle t \rangle$ does not require us to dramatically alter the structure of the estimates carried out in Section 5.1.5. For the sake of brevity the only nonlinear terms we will estimate in detail are NLS1 and NLP.

We derive an energy estimate by considering

$$\frac{1}{2}\frac{d}{dt}\|A Q^2_{\neq}\|^2_{H^N} + \frac{1}{2}\frac{d}{dt}\|A H^2_{\neq}\|^2_{H^N}.$$

Using (4.13), (4.14), and absorbing the term arising from the time derivative landing on

λ as in Section 5.1.4 we obtain

$$\begin{aligned}
&\frac{1}{2}\|AQ^2_{\neq}(t_2)\|^2_{H^N} + \frac{1}{2}\|AH^2_{\neq}(t_2)\|^2_{H^N} + \frac{\nu}{2}\|A\nabla_L Q^2_{\neq}\|^2_{L^2H^N} + \frac{\nu}{2}\|A\nabla_L H^2_{\neq}\|^2_{L^2H^N}\\
&+ \|M\lambda\sqrt{-\dot{m}m}H^2_{\neq}\|^2_{L^2H^N} + \|M\lambda\sqrt{-\dot{m}m}Q^2_{\neq}\|^2_{L^2H^N} + \frac{1}{2}\|m\lambda\sqrt{-\dot{M}M}Q^2_{\neq}\|^2_{L^2H^N}\\
&+ \frac{1}{2}\|m\lambda\sqrt{-\dot{M}M}H^2_{\neq}\|^2_{L^2H^N} - \alpha\int_{t_1}^{t_2}\int A\langle\nabla\rangle^N Q^2_{\neq}\partial_\sigma(A\langle\nabla\rangle^N H^2_{\neq})dV\,dt\\
&+ \int_{t_1}^{t_2}\int A\langle\nabla\rangle^N H^2_{\neq}\partial_\sigma(A\langle\nabla\rangle^N Q^2_{\neq})dV\,dt\\
&\le \frac{1}{2}\|AQ^2_{\neq}(t_1)\|^2_{H^N} + \frac{1}{2}\|AH^2_{\neq}(t_1)\|^2_{H^N} - 2\int_{t_1}^{t_2}\left\langle AH^2_{\neq}, A\partial^L_{XY}\Delta_L^{-1}H^2_{\neq}\right\rangle_{H^N} dt\\
&- \int_{t_1}^{t_2}\left\langle AH^2_{\neq}, A(U\cdot\nabla_L H^2 - B\cdot\nabla_L Q^2)\right\rangle_{H^N} dt\\
&- \int_{t_1}^{t_2}\left\langle AH^2_{\neq}, A(Q\cdot\nabla_L B^2 - H\cdot\nabla_L U^2)\right\rangle_{H^N} dt\\
&- 2\int_{t_1}^{t_2}\left\langle AH^2_{\neq}, A(\partial_i^L U^j\partial_{ij}^L B^2 - \partial_i^L B^j\partial_{ij}^L U^2)\right\rangle_{H^N} dt\\
&+ \int_{t_1}^{t_2}\left\langle AQ^2_{\neq}, A\partial_Y^L(\partial_j^L U^i\partial_i^L U^j - \partial_j^L B^i\partial_i^L B^j)\right\rangle_{H^N} dt\\
&- \int_{t_1}^{t_2}\left\langle AQ^2_{\neq}, A(Q\cdot\nabla_L U^2 + U\cdot\nabla_L Q^2 - H\cdot\nabla_L B^2)\right\rangle dt\\
&- \int_{t_1}^{t_2}\left\langle AQ^2_{\neq}, B\cdot\nabla_L H^2 - 2\partial_i^L U^j\partial_{ij}^L U^2 + 2\partial_i^L B^j\partial_{ij}^L B^2)\right\rangle_{H^N} dt\\
&= \frac{1}{2}\|AQ^2_{\neq}(t_1)\|^2_{H^N} + \frac{1}{2}\|AH^2_{\neq}(t_1)\|^2_{H^N} + \mathrm{LS} + \mathrm{NLT} + \mathrm{NLS1} + \mathrm{NLS2} + \mathrm{NLP}\\
&\qquad + \mathrm{NLQ1} + \mathrm{NLQ2}.
\end{aligned}$$

where for the nonlinear stretching and transport terms we have written, for example, NLT to refer to both $-U\cdot\nabla_L H^2$ and $B\cdot\nabla_L Q^2$ (each abbreviation above is given its own integral sign). The terms NLQ1 and NLQ2 in the last line of the expression above are the nonlinear terms that arise in $\int_{t_1}^{t_2}\left\langle AQ^2_{\neq}, A\partial_t Q^2\right\rangle_{H^N} dt$, excluding the nonlinear pressure term NLP. Since H^2 and Q^2 satisfy the same estimates, these terms can all be controlled in the same way as NLT, NLS1, and NLS2. We omit the details for the sake of brevity. Notice also that from integration by parts we have

$$\int A \langle\nabla\rangle^N Q^2_{\neq} \partial_\sigma (A \langle\nabla\rangle^N H^2_{\neq}) dV + \int A \langle\nabla\rangle^N H^2_{\neq} \partial_\sigma (A \langle\nabla\rangle^N Q^2_{\neq}) dV = 0,$$

so the left-hand side of our energy estimate above contains only nonnegative terms.

5.2.1 Linear stretching term

From the definition of m we have

$$\text{LS} \leq \|\sqrt{-\dot{m}m} M\lambda H^2_{\neq}\|^2_{L^2H^N} + \frac{\nu}{32}\|\nabla_L A H^2_{\neq}\|_{L^2H^N},$$

and hence the linear stretching term can be absorbed into the left-hand side of the energy estimate at the cost of changing the factor of $1/2$ multiplying the dissipation term $\nu\|A\nabla_L H^2_{\neq}\|^2_{L^2H^N}$ into a $15/32$. Noting that $\|AQ^2_{\neq}(t_1)\|^2_{H^N}$ and $\|AH^2_{\neq}(t_1)\|^2_{H^N}$ are both bounded by $18\epsilon^2$, it is easy to check that this is consistent with Proposition 4.4.

5.2.2 Nonlinear terms

As described above, the only nonlinear terms that we estimate in detail are NLS1 and NLP. We begin with the stretching term and observe that since Q and H satisfy the same estimates it suffices to simply write

$$\text{NLS1} = -\int_{t_1}^{t_2} \int \langle\nabla\rangle^N A H^2_{\neq} \langle\nabla\rangle^N A(Q \cdot \nabla_L B^2) dV dt.$$

We first consider the interaction between the nonzero modes. When $j \in \{1,3\}$ we use $|\partial_X|, |\partial_Z| \lesssim |m\Delta_L|$, (4.6c), and (5.1) to obtain

$$\begin{aligned} |\mathrm{NLS1}^{\mathrm{LH}}(j,\neq,\neq)| &\lesssim \|AH^2_{\neq}\|_{L^\infty H^N} \nu^{-2/3} \|\tilde{A}Q^j_{\neq}\|_{L^2 H^{N''}} \|AH^2_{\neq}\|_{L^2 H^N} \\ &\lesssim \epsilon^3 \nu^{-2/3} \nu^{-1/6} \nu^{-1/6} = \epsilon^3 \nu^{-1}. \end{aligned}$$

For the other piece of the paraproduct, first notice that from (4.6b) we have the elementary inequality

$$| \langle t \rangle \nabla_{X,Z}| = \frac{\langle t \rangle}{|\nabla_L|} |\nabla_{X,Z}||\nabla_L| \lesssim \frac{\langle t \rangle}{|\nabla_L|} m^{1/2} |\Delta_L| \lesssim \langle \nabla \rangle m^{1/2} |\Delta_L|. \tag{5.2}$$

From $N' - 1 > 3/2$, (4.6d), (5.1) with $s = 1/2$, and (5.2) we then get

$$\begin{aligned} |\mathrm{NLS1}^{\mathrm{HL}}(j,\neq,\neq)| &\lesssim \int_{t_1}^{t_2} \|AH^2_{\neq}\|_{H^N} \| \langle t \rangle^{-1} Q^j_{\neq}\|_{H^N} \|\lambda \langle t \rangle \nabla_{X,Z} B^2_{\neq}\|_{H^{N'-1}} dt \\ &\lesssim \int_{t_1}^{t_2} \|AH^2_{\neq}\|_{H^N} \|\tilde{m}^{1/2} Q^j_{\neq}\|_{H^N} \|JH^2_{\neq}\|_{H^{N'}} dt \\ &\lesssim \|AH^2_{\neq}\|_{L^\infty H^N} \nu^{-1/3} \|\tilde{A}Q^j_{\neq}\|_{L^2 H^N} \|JH^2_{\neq}\|_{L^2 H^{N'}} \\ &\lesssim \epsilon^3 \nu^{-1/3} \nu^{-1/2} \nu^{-1/6} = \epsilon^3 \nu^{-1}. \end{aligned}$$

For $j = 2$ we use (4.6c) and (4.6b) to obtain

$$
\begin{aligned}
&|\mathrm{NLS1}(2,\neq,\neq)| \le |\mathrm{NLS1}^{\mathrm{LH}}(2,\neq,\neq)| + |\mathrm{NLS1}^{\mathrm{HL}}(2,\neq,\neq)| \\
&\lesssim \|AH^2_{\neq}\|_{L^\infty H^N} \left(\|Q^2_{\neq}\|_{L^2 H^{N'}} \|\lambda \partial_Y^L B^2_{\neq}\|_{L^2 H^N} + \|Q^2_{\neq}\|_{L^2 H^N} \|\lambda \partial_Y^L B^2_{\neq}\|_{L^2 H^{N'}} \right) \\
&\lesssim \nu^{-2/3} \|AH^2_{\neq}\|_{L^\infty H^N} \nu^{-2/3} \|m^{1/2} Q^2_{\neq}\|_{L^2 H^{N'}} \|\lambda \partial_Y^L m^{1/2} B^2_{\neq}\|_{L^2 H^N} \\
&\quad + \nu^{-2/3} \|AH^2_{\neq}\|_{L^\infty H^N} \|m Q^2_{\neq}\|_{L^2 H^N} \|\lambda \partial_Y^L B^2_{\neq}\|_{L^2 H^{N'}} \\
&\lesssim \|AH^2_{\neq}\|_{L^\infty H^N} \nu^{-2/3} \left(\|JQ^2_{\neq}\|_{L^2 H^{N'}} \|AH^2_{\neq}\|_{L^2 H^N} + \|AQ^2_{\neq}\|_{L^2 H^N} \|JH^2_{\neq}\|_{L^2 H^{N'}} \right) \\
&\lesssim \epsilon^3 \nu^{-2/3} \nu^{-1/6} \nu^{-1/6} = \epsilon^3 \nu^{-1}.
\end{aligned}
$$

Now we turn to $\mathrm{NLS1}(j,\neq,0)$, which for the same reason as in Section 5.1.5 is only nonzero when $j \in \{2,3\}$. In either of these cases we have from (4.6e), (4.6c), $N'' > 3/2$, and $N \ge 5$ the bound

$$
\begin{aligned}
&|\mathrm{NLS1}(j,\neq,0)| \le |\mathrm{NLS1}^{\mathrm{HL}}(j,\neq,0)| + |\mathrm{NLS1}^{\mathrm{LH}}(j,\neq,0)| \\
&\lesssim \|AH^2_{\neq}\|_{L^2 H^N} \left(\|AQ^j_{\neq}\|_{L^2 H^N} \|\nabla B^2_0\|_{L^\infty H^4} + \|m^{1/2} \lambda Q^j_{\neq}\|_{L^2 H^{N''}} \|\nabla B^2_0\|_{L^\infty H^{N+1}} \right) \\
&\lesssim \|AH^2_{\neq}\|_{L^2 H^N} \nu^{-1/3} \left(\|AQ^j_{\neq}\|_{L^2 H^N} \|B^2_0\|_{L^\infty H^N} + \|AQ^j_{\neq}\|_{L^2 H^{N''}} \|B^2_0\|_{L^\infty H^{N+2}} \right) \\
&\lesssim \epsilon^3 \nu^{-1/6} \nu^{-1/3} (\nu^{-1/2} + \nu^{-1/6} \nu^{-1/3}) = \epsilon^3 \nu^{-1}.
\end{aligned}
$$

Here, as in Section 5.1.5, we have used the bootstrap hypotheses (4.22e) and (4.25) to bound the zero modes.

Remark 11. Since our estimates in this section are weighted by the multiplier A (and not $\tilde{A}$) the cancellation $\mathrm{NLS}(1,\neq,0) = 0$ is absolutely crucial for the estimate to close.

Indeed, the bootstrap hypotheses only give us control on $\tilde{m}Q^1_{\neq}$, and so, due to $m \geq \tilde{m}$, when $j = 1$ we would not be able to use (4.6e) as in the $j \in \{2,3\}$ estimate just performed. Instead, we would need to find a way to introduce $\tilde{m}$ to close the estimate. The key point is that unlike when dealing with the $(\neq, \neq)$ interactions, this cannot be accomplished using (5.1) because λB_0 is not controlled. Since we have control on $\nabla_L Q^1_{\neq}$ from the dissipation, it would be possible to introduce $\tilde{m}^{1/2}$ by using $1 \lesssim \tilde{m}^{1/2}|\nabla_L|$. This however would not be sufficient because we need to introduce a full power of $\tilde{m}$. Thus, it is not clear how one could proceed without the cancellation. The important structure that causes NLS$(1, \neq, 0)$ to vanish is not unique to the term. Rather, it comes from the general $Z^j\partial^L_j$ structure that prevents Z^1 from interacting nonlinearly with an X average because it is always paired with an X derivative. This cancellation is used in many of the $(\neq, 0)$ estimates that we omit for $F^2_{\neq}$ and $F^3_{\neq}$, and is important for the proof to work.

Lastly, for NLS1$(0, \neq)$ we obtain from (4.6b) and (4.6c) that

$$
\begin{aligned}
|\text{NLS1}(0,\neq)| &\lesssim \|AH^2_{\neq}\|_{L^2H^N}\|Q_0\|_{L^\infty H^N}\nu^{-1/3}\|AH^2_{\neq}\|_{L^2H^N} \\
&\lesssim \epsilon^3\nu^{-1/6}\nu^{-1/3}\nu^{-1/3}\nu^{-1/6} = \epsilon^3\nu^{-1}.
\end{aligned}
$$

This completes the estimate of NLS1.

We turn now to the nonlinear pressure term. It reads

$$
\begin{aligned}
\text{NLP}(i,j) &= \int_{t_1}^{t_2}\int A\langle\nabla\rangle^N H^2_{\neq}A\langle\nabla\rangle^N \partial^L_Y(\partial^L_j U^i\partial^L_i U^j)_{\neq}dV\,dt \\
&= -\int_{t_1}^{t_2}\int \partial^L_Y A\langle\nabla\rangle^N H^2_{\neq}A\langle\nabla\rangle^N (\partial^L_j U^i\partial^L_i U^j)dV\,dt.
\end{aligned}
$$

We begin with the interaction between the zero and nonzero modes. By symmetry, it suffices to consider NLP$(i, j, \neq, 0)$, which we note is only nonzero for $i \neq 1$. Using (4.6e) and (4.6b), we have

$$\begin{aligned}
|\text{NLP}(i, j, \neq, 0)| &\leq |\text{NLP}^{\text{HL}}(i, j, \neq, 0)| + |\text{NLP}^{\text{LH}}(i, j, \neq, 0)| \\
&\lesssim \|A\nabla_L H^2_{\neq}\|_{L^2 H^N} \left(\|AQ^i_{\neq}\|_{L^2 H^N} \|U^j_0\|_{L^\infty H^N} + \|AQ^i_{\neq}\|_{L^2 H^{N''}} \|U^j_0\|_{L^\infty H^{N+2}}\right) \\
&\lesssim \epsilon^3 \nu^{-1/2}(\nu^{-1/2} + \nu^{-1/6}\nu^{-1/3}) \lesssim \epsilon^3 \nu^{-1},
\end{aligned}$$

which is consistent.

Remark 12. In the related papers [17, 124] the term NLP$(3, 1, \neq, 0)$ is the leading order piece of the nonlinearity, but for us this term is relatively easy due to the suppression of the lift-up effect for Q^1_0.

For the interaction between the nonzero modes there are, as before, three distinct cases to consider: $i, j \in \{1, 3\}$, $i = 2$ and $j \neq 2$, and $i = j = 2$. We first consider when $i, j \in \{1, 3\}$. By symmetry, we only need to control the HL interaction. Employing (4.6b) we have

$$\begin{aligned}
&|\text{NLP}^{\text{HL}}(i \in \{1, 3\}, j \in \{1, 3\}, \neq, \neq)| \\
&\qquad\lesssim \|\nabla_L A H^2_{\neq}\|_{L^2 H^N} \|\tilde{A} Q^i_{\neq}\|_{L^\infty H^N} \|\tilde{A} Q^j_{\neq}\|_{L^2 H^{N''}} \\
&\qquad\lesssim \epsilon^3 \nu^{-1/2} \nu^{-1/3} \nu^{-1/6} = \epsilon^3 \nu^{-1}.
\end{aligned}$$

Turning to the second case, we treat the LH interaction using Lemma 4.5, (4.6d), and

(4.6b):

$$|\mathrm{NLP}^{\mathrm{LH}}(i=2,j\neq 2,\neq,\neq)|$$
$$\lesssim \|\nabla_L A H^2_{\neq}\|_{L^2H^N}\| \langle t\rangle \nabla_{X,Z} U^2_{\neq}\|_{L^\infty H^{N'-1}}\|\tilde{A} Q^j_{\neq}\|_{L^2H^N}$$
$$\lesssim \epsilon^3\nu^{-1/2}\nu^{-1/2} = \epsilon^3\nu^{-1}.$$

For the HL term we have, using (5.1) and (4.6b),

$$|\mathrm{NLP}^{\mathrm{HL}}(i=2,j\neq 2,\neq,\neq)| \lesssim \|\nabla_L A H^2_{\neq}\|_{L^2H^N}\|AQ^2_{\neq}\|_{L^\infty H^N}\nu^{-1/3}\|\tilde{A} Q^j_{\neq}\|_{L^2H^{N''}}$$
$$\lesssim \epsilon^3\nu^{-1/2}\nu^{-1/3}\nu^{-1/6} = \epsilon^3\nu^{-1}.$$

Lastly we consider $i=j=2$, for which it suffices to consider only the HL term. Using (4.6b) and (4.6c), we have the estimate

$$|\mathrm{NLP}^{\mathrm{HL}}(2,2,\neq,\neq)| \lesssim \|A\nabla_L H^2_{\neq}\|_{L^2H^N}\nu^{-1/3}\|AQ^2_{\neq}\|_{L^2H^N}\|JQ^2_{\neq}\|_{L^\infty H^{N'}} \lesssim \epsilon^3\nu^{-1}.$$

This completes the estimate of $H^2_{\neq}$ and $Q^2_{\neq}$ in the high norm.

5.3 High norm estimate of $F^3_{\neq}$

Improving (4.22b) follows from essentially the same methods used in Section 5.1. To see this, first note that, disregarding the lift-up term, F^3 satisfies the same equation as F^1 except for the presence of ∂_Z instead of ∂_X in the pressure terms. This is inconsequential in the estimates. For example, the linear pressure is simply controlled with (4.7b) instead of (4.7a). The only other variations are due to the use of m instead of $\tilde{m}$ in the norm and were already encountered in the estimate of $F^2_{\neq}$. In particular, we treat the linear stretching term as in Section 5.2.1, and we rely on the crucial nonlinear structure noted above in Remark 11.

5.4 Summary of high norm nonzero mode interactions

For the sake of clarity in the remainder of the paper it is useful to gather the above calculations into a general lemma. Let $(\partial_t F^j)_{\mathcal{NL}}$ denote the nonlinear terms in $\partial_t F^j$. We then also define $(\partial_t F^j)^{\neq\neq}_{\mathcal{NL}}$ and $(\partial_t F^j)^{00}_{\mathcal{NL}}$ to denote $(\partial_t F^j)_{\mathcal{NL}}$ restricted to either the interaction between the nonzero modes or the interaction between the zero modes.

Our estimates of the $(\neq, \neq)$ nonlinear interactions in Sections 5.1-5.3 only relied on the enhanced dissipation of the functions in the particular nonlinear term (i.e., we did not use the enhanced dissipation of the function that plays the role of G in Lemma 5.1 below). Moreover, we did not employ any commutator type estimates for m and $\tilde{m}$ other than (4.6b), (4.6d), and (4.6c). In particular, we did not use (4.6e). Due to these observations, our calculations above yield the following lemma.

Lemma 5.1. *Let $\beta \geq 0$ and suppose that $G : [t_1, t_2] \times \mathbb{T} \times \mathbb{R} \times \mathbb{T} \to \mathbb{R}$ satisfies $G \in C([t_1, t_2]; H^{s'})$ for every $s' \geq 0$. If*

$$\|G\|_{L^\infty H^s} + \nu^{1/2} \|\nabla_L G\|_{L^2 H^s} \lesssim \epsilon \nu^{-\beta}$$

for some $s \leq N$, then for any bounded and measurable Fourier multiplier $\mathcal{M}$ there holds

$$\left| \int_{t_1}^{t_2} \left\langle \lambda \mathcal{M} (\partial_t F^j)^{\neq\neq}_{\mathcal{NL}}, G \right\rangle_{H^s} dt \right| \lesssim \epsilon^3 \nu^{-4/3-\beta} \quad (j \in \{1, 3\}), \tag{5.3}$$

$$\left| \int_{t_1}^{t_2} \left\langle \lambda \mathcal{M} (\partial_t F^2)^{\neq\neq}_{\mathcal{NL}}, G \right\rangle_{H^s} dt \right| \lesssim \epsilon^3 \nu^{-1-\beta}. \tag{5.4}$$

5.5 High norm estimate of Q_0 and H_0

In this section we improve (4.22e). For any $r \in \{1, 2, 3\}$, an energy estimate gives

$$\begin{aligned}
&\frac{1}{2}\|F_0^{+,r}(t_2)\|^2_{H^N} + \nu \|\nabla F_0^{+,r}\|^2_{L^2 H^N} = \frac{1}{2}\|F_0^{+,r}(t_1)\|^2_{H^N} \\
&\quad - \mathbf{1}_{r=1} \int_{t_1}^{t_2} \left\langle F_0^{+,1}, T^t_{2\alpha} F_0^{-,2} \right\rangle_{H^N} dt - \int_{t_1}^{t_2} \left\langle F_0^{+,r}, T^t_{2\alpha} Z^- \cdot \nabla_L F^{+,r} \right\rangle_{H^N} dt \\
&\quad - \int_{t_1}^{t_2} \left\langle F_0^{+,r}, T^t_{2\alpha} F^- \cdot \nabla_L Z^{+,r} \right\rangle_{H^N} dt - \int_{t_1}^{t_2} \left\langle F_0^{+,r}, T^t_{2\alpha} \partial_i^L Z^{-,j} \partial_{ij}^L Z^{+,r} \right\rangle_{H^N} dt \\
&\quad + \mathbf{1}_{r\neq 1} \int_{t_1}^{t_2} \left\langle F_0^{+,r}, \partial_r (T^t_{2\alpha} \partial_j^L Z^{-,i} \partial_i^L Z^{+,j}) \right\rangle_{H^N} dt \\
&\quad = \frac{1}{2}\|F_0^{+,r}(t_1)\|^2_{H^N} + \mathrm{LU} + \mathrm{NLT} + \mathrm{NLS1} + \mathrm{NLS2} + \mathrm{NLP}.
\end{aligned}$$

5.5.1 Nonlinear terms

For the interaction between the nonzero modes we have, by Lemma 5.1 and (4.22e),

$$\left| \int_{t_1}^{t_2} \left\langle F_0^r, (\partial_t F^r)^{\neq\neq}_{\mathcal{NL}} \right\rangle_{H^N} dt \right| \lesssim (\epsilon \nu^{-1/3})^2 \epsilon \nu^{-1},$$

which is consistent. It then only remains to consider the interaction between the zero modes. We begin with the transport term

$$\mathrm{NLT}(j,0,0) = -\int_{t_1}^{t_2}\int \langle\nabla\rangle^N F_0^{+,r}\,\langle\nabla\rangle^N (Z_0^{-,j}\partial_j F_0^{+,r})dV dt.$$

When $j = 2$ we use that incompressibility implies that Z_0^2 always has a nonzero Z-frequency to obtain

$$\begin{aligned}|\mathrm{NLT}(2,0,0)| &\lesssim \|F_0^r\|_{L^\infty H^N}\|\nabla Z_0^2\|_{L^2H^N}\|\nabla F_0^r\|_{L^2H^N}\\ &\lesssim \epsilon^3\nu^{-1/3}\nu^{-1/2}\nu^{-5/6} = (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1}.\end{aligned}$$

For $j = 3$ we observe that the term vanishes unless at least one of Z_0^3 or F_0^r has a nonzero Z-frequency. Hence,

$$\begin{aligned}|\mathrm{NLT}(3,0,0)| &\lesssim \|F_0^r\|_{L^\infty H^N}\|\nabla Z_0^3\|_{L^2H^N}\|\partial_Z F_0^r\|_{L^2H^N}\\ &\quad + \|\nabla F_0^r\|_{L^2H^N}\|Z_0^3\|_{L^\infty H^N}\|\partial_Z F_0^r\|_{L^2H^N}\\ &\lesssim \epsilon^3(\nu^{-1/3}\nu^{-1/2}\nu^{-5/6} + \nu^{-5/6}\nu^{-5/6}) \lesssim (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1},\end{aligned}$$

which suffices and completes the estimate of the transport nonlinearity. Using incompressibility and $\partial_Z = \partial_Z P_{\ell\neq 0}$ in a similar fashion as above, both of the stretching terms are treated in essentially the same way as the transport term. We thus skip them and turn to the pressure, for which, after an integration by parts, we have the estimate

$$
\begin{aligned}
|\mathrm{NLP}(i,j,0,0)| &\lesssim \|\nabla F_0^r\|_{L^2H^N}\|\nabla Z_0^j\|_{L^2H^N}\|Z_0^i\|_{L^\infty H^{N+2}} \\
&\lesssim \epsilon^3\nu^{-5/6}\nu^{-1/2}\nu^{-1/3} = (\epsilon\nu^{-1/3})^2\epsilon\nu^{-1}.
\end{aligned}
$$

This completes the high norm estimate of the nonlinear terms for F_0^r, $r \in \{1,2,3\}$.

The computations in this section were not sensitive to the component of F_0 being estimated. In fact, the estimates of NLT, NLS1, and NLS2, which are each quadratic in the r component, do not even rely on any structures that would cause the same methods to fail if the two occurrences of r were replaced by r and some r'. This generality gives us the following lemma, which will be useful when considering the nonlinear terms that arise in our treatment of the lift-up effect with integration by parts in time.

Lemma 5.2. *Let $r \in \{1,2,3\}$ and suppose that $G : [t_1,t_2]\times\mathbb{T}\times\mathbb{R}\times\mathbb{T}\to\mathbb{R}$ satisfies $G \in C([t_1,t_2];H^{s'})$ for every $s' \geq 0$. If*

$$
\|G\|_{L^\infty H^N} + \nu^{1/2}\|\nabla G\|_{L^2H^N} \lesssim \epsilon\nu^{-1/3},
$$

then

$$
\left|\int_{t_1}^{t_2}\left\langle G, (\partial_t F^r)^{00}_{\mathcal{NL}}\right\rangle_{H^N} dt\right| \lesssim \epsilon^3\nu^{-5/3}.
$$

5.5.2 Suppression of the lift-up effect

As discussed above, the main stabilizing effect in our work is that the magnetic field induces oscillations that suppress the lift-up effect. In this section, we show how to estimate LU with no losses by exploiting these oscillations using integration by parts in time. Of crucial importance is that incompressibility implies that $F_0^{-,2}$ always has a nonzero Z-frequency, which ensures that there is no component of the lift-up term that does not oscillate. Noting that $T_\alpha^t g_0 = e^{\alpha t \partial_Z} g_0$ for any function g, we integrate by parts in time to obtain

$$
\begin{aligned}
-\mathrm{LU} &= \frac{1}{2\alpha}\int_{t_1}^{t_2}\int \langle\nabla\rangle^N P_{\ell\neq 0}F_0^{+,1}\,\langle\nabla\rangle^N \partial_Z^{-1}P_{\ell\neq 0}\partial_t T_{2\alpha}^t F_0^{-,2}\,dV\,dt \\
&= \frac{1}{2\alpha}\left\langle P_{\ell\neq 0}F_0^{+,1}(t_2), \partial_Z^{-1}T_{2\alpha}^{t_2}P_{\ell\neq 0}F_0^{-,2}(t_2)\right\rangle_{H^N} \qquad (5.5)\\
&\quad - \frac{1}{2\alpha}\left\langle P_{\ell\neq 0}F_0^{+,1}(t_1), \partial_Z^{-1}T_{2\alpha}^{t_1}P_{\ell\neq 0}F_0^{-,2}(t_1)\right\rangle_{H^N} \qquad (5.6)\\
&\quad - \frac{1}{2\alpha}\int_{t_1}^{t_2}\left\langle P_{\ell\neq 0}\partial_t F_0^{+,1}, \partial_Z^{-1}T_{2\alpha}^{t}P_{\ell\neq 0}F_0^{-,2}\right\rangle_{H^N} dt \qquad (5.7)\\
&\quad - \frac{1}{2\alpha}\int_{t_1}^{t_2}\left\langle P_{\ell\neq 0}F_0^{+,1}, \partial_Z^{-1}T_{2\alpha}^{t}P_{\ell\neq 0}\partial_t F_0^{-,2}\right\rangle_{H^N} dt. \qquad (5.8)
\end{aligned}
$$

The boundary terms (5.5) and (5.6) are both treated similarly. For example, by Cauchy-Schwarz and the fact that $\partial_Z^{-1}P_{\ell\neq 0}$ is bounded on H^N we have

$$
\begin{aligned}
\left|\frac{1}{2\alpha}\left\langle P_{\ell\neq 0}F_0^{+,1}(t_2), \partial_Z^{-1}T_{2\alpha}^{t_2}P_{\ell\neq 0}F_0^{-,2}(t_2)\right\rangle_{H^N}\right| &\lesssim \frac{1}{\alpha}\|F_0^{+,1}\|_{L^\infty H^N}\|F_0^{-,2}\|_{L^\infty H^N} \\
&\lesssim \frac{1}{\alpha}(\epsilon\nu^{-1/3})^2,
\end{aligned}
$$

which is consistent for α sufficiently large. Now we turn to (5.7). Expanding it out gives

$$
\begin{aligned}
&\frac{1}{2\alpha}\int_{t_1}^{t_2} \left\langle P_{\ell\neq 0}\partial_t F_0^{+,1}, \partial_Z^{-1} T_{2\alpha}^t P_{\ell\neq 0} F_0^{-,2}\right\rangle_{H^N} dt \\
&\quad = \frac{1}{2\alpha}\int_{t_1}^{t_2} \left\langle \nu\Delta P_{\ell\neq 0} F_0^{+,1}, T_{2\alpha}^t \partial_Z^{-1} P_{\ell\neq 0} F_0^{-,2}\right\rangle_{H^N} dt \\
&\qquad - \frac{1}{2\alpha}\int_{t_1}^{t_2} \left\langle T_{2\alpha}^t F_0^{-,2}, \partial_Z^{-1} T_{2\alpha}^t P_{\ell\neq 0} F_0^{-,2}\right\rangle_{H^N} dt \\
&\qquad + \frac{1}{2\alpha}\int_{t_1}^{t_2} \left\langle P_{\ell\neq 0}(\partial_t F_0^{+,1})_{\mathcal{NL}}, \partial_Z^{-1} T_{2\alpha}^t P_{\ell\neq 0} F_0^{-,2}\right\rangle_{H^N} dt \\
&\quad = \mathrm{LU1} + \mathrm{LU2} + \mathcal{NL}.
\end{aligned}
$$

The linear term LU1 arising from the dissipation is treated naturally with an integration by parts:

$$
\begin{aligned}
|\mathrm{LU1}| &= \left| \frac{1}{2\alpha}\int_{t_1}^{t_2} \left\langle \nu\nabla P_{\ell\neq 0} F_0^{+,1}, T_{2\alpha}^t \partial_Z^{-1} \nabla P_{\ell\neq 0} F_0^{-,2}\right\rangle_{H^N} dt \right| \\
&\lesssim \frac{\nu}{\alpha} \|\nabla F_0^{+,1}\|_{L^2 H^N} \|\nabla F_0^{-,2}\|_{L^2 H^N} \\
&\lesssim \frac{\nu}{\alpha} \epsilon^2 \nu^{-5/6} \nu^{-5/6} = \frac{1}{\alpha}(\epsilon \nu^{-1/3})^2,
\end{aligned}
$$

which is consistent for α sufficiently large. Crucially, the other linear term LU2 vanishes. Indeed, the inner product under the time integral can be rewritten as

$$
\begin{aligned}
&\int P_{\ell\neq 0} \langle\nabla\rangle^N T_{2\alpha}^t F_0^{-,2} \langle\nabla\rangle^N \partial_Z^{-1} T_{2\alpha}^t P_{\ell\neq 0} F_0^{-,2} dV \\
&\qquad = \frac{1}{2}\int \partial_Z \left(\partial_Z^{-1} P_{\ell\neq 0} \langle\nabla\rangle^N T_{2\alpha}^t F_0^{-,2}\right)^2 dV = 0.
\end{aligned}
$$

For the nonlinear contribution $\mathcal{NL}$ we have, by (4.22e) and Lemmas 5.1 and 5.2,

$$\begin{aligned}|\mathcal{NL}| &\lesssim \left|\int_{t_1}^{t_2} \left\langle \partial_Z^{-1} T_{2\alpha}^t P_{\ell\neq 0} F_0^{-,2}, (\partial_t F^{+,1})_{\mathcal{NL}}^{\neq\neq} \right\rangle_{H^N} dt\right| \\ &\quad + \left|\int_{t_1}^{t_2} \langle \partial_Z^{-1} T_{2\alpha}^t P_{\ell\neq 0} F_0^{-,2}, (\partial_t F^{+,1})_{\mathcal{NL}}^{00} \rangle_{H^N} dt\right| \\ &\lesssim \epsilon^3 \nu^{-5/3} = (\epsilon\nu^{-1/3})^2 \epsilon\nu^{-1}.\end{aligned}$$

This complete the estimate of (5.7). The second term (5.8) is treated similarly, and so we omit the details. In fact, this term is simpler because the only linear contribution comes from the dissipation.

5.6 Intermediate norm estimate of $F_{\neq}^2$ in $H^{N'+2+n}$

Our focus in this section is how to use (3.12) along with the the high norm control on $F_{\neq}^2$ to improve (4.23). We will provide the details for improving (4.23a) and then briefly discuss how the same techniques carry over to the other estimate.

Recall the notations $\tilde{N} = N' + 2 + n$ and $\tilde{J} = \langle t \rangle^{-1/2} J$. Dropping the negative term from the time derivative landing on the decaying time weight and absorbing L_λ as before into the left-hand side, we obtain the energy estimate

$$
\begin{aligned}
&\frac{1}{2}\|\tilde{J}F_{\neq}^{+,2}(t_2)\|_{H^{\tilde{N}}}^2+\frac{\nu}{2}\|\tilde{J}\nabla_L F_{\neq}^{+,2}\|_{L^2H^{\tilde{N}}}^2+\frac{1}{2}\|\left\langle t\right\rangle^{-1/2}m^{1/2}\lambda\sqrt{-\dot{M}M}F_{\neq}^{+,2}\|_{L^2H^{\tilde{N}}}^2\\
&+\|\left\langle t\right\rangle^{-1/2}M\lambda\sqrt{-\dot{m}^{1/2}m^{1/2}}F_{\neq}^{+,2}\|_{L^2H^{\tilde{N}}}^2\leq\frac{1}{2}\|\tilde{J}F_{\neq}^{+,2}(t_1)\|_{H^N}^2\\
&-\int_{t_1}^{t_2}\left\langle\tilde{J}F_{\neq}^{+,2},\partial_{XY}^L\Delta_L^{-1}\tilde{J}F_{\neq}^{+,2}\right\rangle_{H^{\tilde{N}}}dt+\int_{t_1}^{t_2}\left\langle\tilde{J}F_{\neq}^{+,2},\partial_{XY}^L\Delta_L^{-1}T_{2\alpha}^t\tilde{J}F_{\neq}^{-,2}\right\rangle_{H^{\tilde{N}}}dt\\
&-\int_{t_1}^{t_2}\left\langle\tilde{J}F_{\neq}^{+,2},\tilde{J}(T_{2\alpha}^t Z^-\cdot\nabla_L F^{+,2})\right\rangle_{H^{\tilde{N}}}dt\\
&-\int_{t_1}^{t_2}\left\langle\tilde{J}F_{\neq}^{+,2},\tilde{J}(T_{2\alpha}^t F^-\cdot\nabla_L Z^{+,2})\right\rangle_{H^{\tilde{N}}}dt\\
&-2\int_{t_1}^{t_2}\left\langle\tilde{J}F_{\neq}^{+,2},\tilde{J}(T_{2\alpha}^t\partial_i^L Z^{-,j}\partial_{ij}^L Z^{+,2})\right\rangle_{H^{\tilde{N}}}dt\\
&+\int_{t_1}^{t_2}\left\langle\tilde{J}F_{\neq}^{+,2},\partial_Y^L\tilde{J}(T_{2\alpha}^t\partial_j^L Z^{-,i}\partial_i^L Z^{+,j})\right\rangle_{H^{\tilde{N}}}dt\\
&=\frac{1}{2}\|\tilde{J}F_{\neq}^{+,2}(t_1)\|_{H^N}^2+\mathrm{LS}+\mathrm{OLS}+\mathrm{NLT}+\mathrm{NLS1}+\mathrm{NLS2}+\mathrm{NLP}.
\end{aligned}
$$

Below we consider only OLS and the nonlinear terms, since the term LS can absorbed into the left-hand side of the estimate by using (4.5) in the same way that we used (4.3) in Section 5.2.1.

5.6.1 Oscillating linear stretching term

We now use integration by parts in time to control the oscillating linear stretching term with no losses, which is key to the proof and the fundamental difference between the results in Theorem 3.3 and Corollary 3.4. We begin by introducing the shorthand notation

$S = S(t, \nabla) = \partial_{XY}^L \Delta_L^{-1}$. That is, S is the Fourier multiplier with symbol

$$S(t, k, \eta, \ell) = \frac{k(\eta - kt)}{k^2 + \ell^2 + (\eta - kt)^2}.$$

Note that we have the inequality

$$|S(t)| \lesssim \frac{1}{\langle t \rangle} |k, \ell, \eta|. \tag{5.9}$$

Integrating by parts in time we write the term as

$$\begin{aligned}
\text{OLS} &= \frac{1}{2\alpha} \int_{t_1}^{t_2} \left\langle \tilde{J} F_{\neq}^{+,2}, \partial_t T_{2\alpha}^t \partial_\sigma^{-1} S \tilde{J} F_{\neq}^{-,2} \right\rangle_{H^{\tilde{N}}} dt \\
&= \frac{1}{2\alpha} \left\langle \tilde{J} F_{\neq}^{+,2}(t_2), T_{2\alpha}^{t_2} \partial_\sigma^{-1} S \tilde{J} F_{\neq}^{-,2}(t_2) \right\rangle_{H^{\tilde{N}}} & (5.10) \\
&\quad - \frac{1}{2\alpha} \left\langle \tilde{J} F_{\neq}^{+,2}(t_1), T_{2\alpha}^{t_1} \partial_\sigma^{-1} S \tilde{J} F_{\neq}^{-,2}(t_1) \right\rangle_{H^{\tilde{N}}} & (5.11) \\
&\quad - \frac{1}{2\alpha} \int_{t_1}^{t_2} \left\langle \tilde{J} F_{\neq}^{+,2}, \left(2 \frac{\dot{\tilde{J}}}{\tilde{J}} S + \dot{S} \right) T_{2\alpha}^t \partial_\sigma^{-1} \tilde{J} F_{\neq}^{-,2} \right\rangle_{H^{\tilde{N}}} dt & (5.12) \\
&\quad - \frac{1}{2\alpha} \int_{t_1}^{t_2} \left\langle \tilde{J} \partial_t F_{\neq}^{+,2}, T_{2\alpha}^t \partial_\sigma^{-1} S \tilde{J} F_{\neq}^{-,2} \right\rangle_{H^{\tilde{N}}} dt & (5.13) \\
&\quad - \frac{1}{2\alpha} \int_{t_1}^{t_2} \left\langle \tilde{J} F_{\neq}^{+,2}, T_{2\alpha}^t \partial_\sigma^{-1} \tilde{J} S \partial_t F_{\neq}^{-,2} \right\rangle_{H^{\tilde{N}}} dt. & (5.14)
\end{aligned}$$

The boundary terms (5.10) and (5.11) are both treated similarly, and so we will only estimate (5.10). Recalling the definitions of c and n from Theorem 3.3, we have, by

Cauchy-Schwarz, (3.12), (4.6d), and (5.9),

$$
\begin{aligned}
\frac{1}{2\alpha}&\left|\left\langle \tilde{J}F_{\neq}^{+,2}(t_2), T_{2\alpha}^{t_2}\partial_\sigma^{-1}S\tilde{J}F_{\neq}^{-,2}(t_2)\right\rangle_{H^{\tilde{N}}}\right| \\
&\lesssim \frac{1}{c\alpha}\|\tilde{J}F_{\neq}^{+,2}\|_{L^\infty H^{\tilde{N}}}\|\left\langle t\right\rangle^{-1/2} m^{-1/2}SAF_{\neq}^{-,2}\|_{L^\infty H^{\tilde{N}+n}} \\
&\lesssim \frac{1}{c\alpha}\|\tilde{J}F_{\neq}^{+,2}\|_{L^\infty H^{\tilde{N}}}\|\left\langle t\right\rangle^{1/2} SAF_{\neq}^{-,2}\|_{L^\infty H^{\tilde{N}+n}} \\
&\lesssim \frac{1}{c\alpha}\|\tilde{J}F_{\neq}^{+,2}\|_{L^\infty H^{\tilde{N}}}\|AF_{\neq}^{-,2}\|_{L^\infty H^{\tilde{N}+1+n}} \\
&\lesssim \frac{\epsilon^2}{c\alpha},
\end{aligned}
$$

which suffices for $c\alpha$ chosen sufficiently large. In the last line above we have used that $\tilde{N}+1+n = N'+3+2n \leq N$. Next consider (5.12), which splits into five terms since

$$
\frac{\dot{\tilde{J}}}{\tilde{J}}S + \dot{S} = \left(\frac{\dot{m}^{1/2}}{m^{1/2}} + \delta\nu^{1/3} + \frac{\dot{M}}{M} - \frac{1}{2}t\left\langle t\right\rangle^{-2}\right)S + \dot{S}. \tag{5.15}
$$

In the order listed in (5.15) we label these five terms as

$$
(5.12) = T_1 + T_2 + T_3 + T_4 + T_5.
$$

For T_1 we recall (4.5), which states that $|\dot{m}^{1/2}/m^{1/2}| = |S(t)|$ on it's support. Hence, again using (5.9), (4.6d), and (3.12), there holds

$$
\begin{aligned}
|T_1| &\lesssim \frac{1}{c\alpha}\int_{t_1}^{t_2}\|\tilde{J}F_{\neq}^{+,2}\|_{H^{\tilde{N}}}\frac{1}{\left\langle t\right\rangle^{3/2}}\|AF_{\neq}^{-,2}\|_{H^{\tilde{N}+2+n}}dt \\
&\lesssim \frac{1}{c\alpha}\|\tilde{J}F_{\neq}^{+,2}\|_{L^\infty H^{\tilde{N}}}\|AF_{\neq}^{-,2}\|_{L^\infty H^{\tilde{N}+2+n}} \lesssim \frac{\epsilon^2}{c\alpha},
\end{aligned}
$$

which is consistent. Since $|\dot{S}| \lesssim |S(t)|^2$ and $t \langle t \rangle^{-2} \lesssim \langle t \rangle^{-1}$, it follows that both T_4 and T_5 can be estimated in exactly the same manner as T_1. We thus skip these terms and turn to T_2 and T_3, for which we have the slight variations

$$
\begin{aligned}
|T_2| &\lesssim \frac{\nu^{1/3}}{c\alpha} \|\tilde{J} F_{\neq}^{+,2}\|_{L^2 H^{\tilde{N}}} \|A F_{\neq}^{-,2}\|_{L^2 H^{\tilde{N}+1+n}} \lesssim \frac{\nu^{1/3}}{c\alpha} \epsilon^2 \nu^{-1/6} \nu^{-1/6} = \frac{\epsilon^2}{c\alpha}, \\
|T_3| &\lesssim \frac{1}{c\alpha} \|\sqrt{-\dot{M} M} \tilde{J} F_{\neq}^{+,2}\|_{L^2 H^{\tilde{N}}} \|\sqrt{-\dot{M} M} A F_{\neq}^{-,2}\|_{L^2 H^{\tilde{N}+1+n}} \lesssim \frac{\epsilon^2}{c\alpha}.
\end{aligned}
$$

Now we consider (5.13) and (5.14). Expanding out (5.13) gives

$$
\begin{aligned}
(5.13) = &\frac{1}{2\alpha} \int_{t_1}^{t_2} \left\langle \tilde{J} S F_{\neq}^{+,2}, T_{2\alpha}^{t} \partial_\sigma^{-1} S \tilde{J} F_{\neq}^{-,2} \right\rangle_{H^{\tilde{N}}} dt && (5.16) \\
&- \frac{1}{2\alpha} \int_{t_1}^{t_2} \left\langle T_{2\alpha}^{t} \tilde{J} S F_{\neq}^{-,2}, T_{2\alpha}^{t} \partial_\sigma^{-1} S \tilde{J} F_{\neq}^{-,2} \right\rangle_{H^{\tilde{N}}} dt && (5.17) \\
&- \frac{\nu}{2\alpha} \int_{t_1}^{t_2} \left\langle \tilde{J} \Delta_L F_{\neq}^{+,2}, T_{2\alpha}^{t} \partial_\sigma^{-1} S \tilde{J} F_{\neq}^{-,2} \right\rangle_{H^{\tilde{N}}} dt - \mathcal{NL}, && (5.18)
\end{aligned}
$$

where

$$
\mathcal{NL} = \frac{1}{2\alpha} \int_{t_1}^{t_2} \left\langle \tilde{J} (\partial_t F^{+,2})_{\mathcal{NL}}, T_{2\alpha}^{t} \partial_\sigma^{-1} S \tilde{J} F_{\neq}^{-,2} \right\rangle_{H^{\tilde{N}}} dt. \tag{5.19}
$$

Using that S is self-adjoint, (5.16) and (5.17) above are bounded as was T_1. For the linear part of (5.18) we integrate by parts:

$$\begin{aligned}
&\left|\frac{1}{2\alpha}\int_{t_1}^{t_2}\left\langle \nu\Delta_L \tilde{J}F_{\neq}^{+,2}, T_{2\alpha}^{t}\partial_\sigma^{-1}\tilde{J}SF_{\neq}^{-,2}\right\rangle_{H^{\tilde{N}}} dt\right| \\
&\qquad = \left|\frac{1}{2\alpha}\int_{t_1}^{t_2}\left\langle \nu\nabla_L \tilde{J}F_{\neq}^{+,2}, T_{2\alpha}^{t}\partial_\sigma^{-1}\tilde{J}S\nabla_L F_{\neq}^{-,2}\right\rangle_{H^{\tilde{N}}} dt\right| \\
&\qquad \lesssim \frac{\nu}{c\alpha}\|\tilde{J}\nabla_L F_{\neq}^{+,2}\|_{L^2H^{\tilde{N}}}\|A\nabla_L F_{\neq}^{-,2}\|_{L^2H^{\tilde{N}+1+n}} \\
&\qquad \lesssim \frac{\nu}{c\alpha}\epsilon^2\nu^{-1/2}\nu^{-1/2} = \frac{\epsilon^2}{c\alpha}.
\end{aligned}$$

The contribution to (5.13) from $\mathcal{NL}$ will be considered below in Section 5.6.2 along with the natural nonlinear terms that arise in the energy estimate. Regarding (5.14), we observe that it can be rewritten as

$$(5.14) = \frac{1}{2\alpha}\int_{t_1}^{t_2}\left\langle T_{-2\alpha}^{t}\partial_\sigma^{-1}S\tilde{J}F_{\neq}^{+,2}, \tilde{J}\partial_t F_{\neq}^{-,2}\right\rangle_{H^{\tilde{N}}} dt,$$

and hence it is essentially symmetric to (5.13). It can thus be estimated in the same way. This completes the estimate of OLS.

5.6.2 Nonlinear terms

In this section we treat the nonlinear terms from the energy estimate in Section 5.6 as well as the term $\mathcal{NL}$ defined in (5.19) above. Recall the notations $(\partial_t F^j)_{\mathcal{NL}}$ and $(\partial_t F^j)_{\mathcal{NL}}^{\neq\neq}$ from the beginning of Section 5.4. We split $\mathcal{NL}$ as

$$\mathcal{NL} = \mathcal{NL}^{0\neq} + \mathcal{NL}^{\neq\neq}$$

corresponding to writing $(\partial_t F^{+,2})_{\mathcal{NL}} = \left((\partial_t F^{+,2})_{\mathcal{NL}} - (\partial_t F^{+,2})^{\neq\neq}_{\mathcal{NL}}\right) + (\partial_t F^{+,2})^{\neq\neq}_{\mathcal{NL}}$ in (5.19).

We begin with the nonlinear terms that correspond to interactions between the nonzero modes. For the terms arising in the initial energy estimate from Section 5.6 we have, by (4.23a) and Lemma 5.1,

$$\left| \int_{t_1}^{t_2} \left\langle \tilde{J} F^2_{\neq}, \tilde{J}(\partial_t F^2)^{\neq\neq}_{\mathcal{NL}} \right\rangle_{H^{\tilde{N}}} dt \right| \lesssim \epsilon^3 \nu^{-1}.$$

The term $\mathcal{NL}^{\neq\neq}$ is controlled similarly. Indeed, using (5.9), (3.12), and $\tilde{N} \ll N$, we have

$$\|T^t_{2\alpha} \partial_\sigma^{-1} S \tilde{J} F^2_{\neq}\|_{H^{\tilde{N}}} \lesssim \|A F^2_{\neq}\|_{H^N}, \tag{5.20}$$

and hence $|\mathcal{NL}^{\neq\neq}| \lesssim \epsilon^3 \nu^{-1}$ by (4.22c), (4.22d), and Lemma 5.1.

Now we turn to the interactions between the zero and nonzero modes. Unlike the $(\neq, \neq)$ terms, they do not follow directly from previous calculations. This is because we used (4.6e) when controlling these interactions in Section 5.2.2, and m weakens more than $m^{1/2}$ near the critical times ($m^{1/2}/m$ can become size $\nu^{-1/3}$). It turns out however that due to (4.25) and $\tilde{N} + 3 \leq N$ these terms are not difficult to control, as we now demonstrate. We begin with the terms in the energy estimate written in Section 5.6. By

(4.6e) we obtain

$$|\text{NLT}(0,\neq)| = \left|\int_{t_1}^{t_2}\int \langle\nabla\rangle^{\tilde{N}}\tilde{J}F^2_{\neq}\langle\nabla\rangle^{\tilde{N}}\tilde{J}(Z_0\cdot\nabla_L F^2_{\neq})dV\,dt\right|$$
$$\lesssim \|\tilde{J}F^2_{\neq}\|_{L^2H^{\tilde{N}}}\|Z_0\|_{L^\infty H^{\tilde{N}+1}}\|\tilde{J}\nabla_L F^2_{\neq}\|_{L^2H^{\tilde{N}}}$$
$$\lesssim \epsilon^3\nu^{-1/6}\nu^{-1/2} = \epsilon^3\nu^{-2/3}$$

and

$$|\text{NLT}(\neq,0)| = \left|\int_{t_1}^{t_2}\int \langle\nabla\rangle^{\tilde{N}}\tilde{J}F^2_{\neq}\langle\nabla\rangle^{\tilde{N}}\tilde{J}(Z_{\neq}\cdot\nabla F^2_0)dV\,dt\right|$$
$$\lesssim \|\tilde{J}F^2_{\neq}\|_{L^2H^{\tilde{N}}}\|Z_{\neq}\|_{L^2H^{\tilde{N}}}\|Z^2_0\|_{L^\infty H^{\tilde{N}+3}}$$
$$\lesssim \epsilon^3\nu^{-1/6}\nu^{-1/2} = \epsilon^3\nu^{-2/3},$$

which both suffice. For NLS2 we have

$$|\text{NLS2}(i,j,0,\neq)| = \left|\int_{t_1}^{t_2}\int \langle\nabla\rangle^{\tilde{N}}\tilde{J}F^2_{\neq}\langle\nabla\rangle^{\tilde{N}}\tilde{J}(\partial_i Z^j_0\partial^L_{ij}Z^2_{\neq})dV\,dt\right|$$
$$\lesssim \|\tilde{J}F^2_{\neq}\|_{L^2H^{\tilde{N}}}\|Z_0\|_{L^\infty H^{\tilde{N}+2}}\|\tilde{J}F^2_{\neq}\|_{L^2H^{\tilde{N}}}$$
$$\lesssim \epsilon^3\nu^{-1/6}\nu^{-1/6} = \epsilon^3\nu^{-1/3}$$

and

$$
\begin{aligned}
|\text{NLS2}(i,j,\neq,0)| &= \left|\int_{t_1}^{t_2}\int \langle\nabla\rangle^{\tilde{N}}\,\tilde{J}F^2_{\neq}\,\langle\nabla\rangle^{\tilde{N}}\,\tilde{J}(\partial_i^L Z^j_{\neq}\partial_{ij}Z^2_0)dV\,dt\right| \\
&\lesssim \|\tilde{J}F^2_{\neq}\|_{L^2H^{\tilde{N}}}\|AF^j_{\neq}\|_{L^2H^{\tilde{N}}}\|Z^2_0\|_{L^\infty H^{\tilde{N}+3}} \\
&\lesssim \epsilon^3\nu^{-1/6}\nu^{-1/2} = \epsilon^3\nu^{-2/3}.
\end{aligned}
$$

Next, one can check that our methods in Section 5.2.2 only employed (4.6e) in the form

$$
\sqrt{\frac{m(t,k,\eta',\ell')}{m(t,k,\eta,\ell)}} \lesssim \langle\eta-\eta'\rangle + \langle\ell-\ell'\rangle\,,
$$

and hence the $(0,\neq)$ and $(\neq,0)$ interactions for NLS1 and NLP follow immediately from the estimates in Section 5.2.2. For $\mathcal{NL}^{0\neq}$ we notice that by (5.20) all of the inequalities above hold with $\|\tilde{J}F^2_{\neq}\|_{L^2H^{\tilde{N}}}$ replaced with $\|AF^2_{\neq}\|_{L^2H^N}$, which is inconsequential in the final inequality because both quantities are controlled by $\epsilon\nu^{-1/6}$.

5.7 Intermediate norm estimate of $F^2_{\neq}$ in $H^{N'}$

To improve estimate (4.23b) we use the same strategy as in Section 5.6, except now the $H^{\tilde{N}}$ bounds in (4.23a) play the role of the high norm control that absorbs the loss of derivatives arising from integration by parts in time. In particular, in Section 5.6.1 we observed that by using (4.6d) the gap between A and $\tilde{J}$ could be compensated for by paying $\langle t\rangle^{1/2}$ and then using (5.9). Since $J/\tilde{J} = \langle t\rangle^{1/2}$, the same structure applies to the $H^{N'}$ estimate. For example, we estimate the boundary term at $t = t_2$ analogous to (5.10) as follows:

$$\begin{aligned}
\frac{1}{2\alpha}&\left|\left\langle JF_{\neq}^{+,2}(t_2), T_{2\alpha}^{t_2}\partial_\sigma^{-1}SJF_{\neq}^{-,2}(t_2)\right\rangle_{H^{N'}}\right| \\
&\lesssim \frac{1}{c\alpha}\|JF_{\neq}^{+,2}\|_{L^\infty H^{N'}}\| \left\langle t\right\rangle^{1/2} S\tilde{J}F_{\neq}^{-,2}\|_{L^\infty H^{N'+n}} \\
&\lesssim \frac{1}{c\alpha}\|JF_{\neq}^{+,2}\|_{L^\infty H^{N'}}\|\tilde{J}F_{\neq}^{-,2}\|_{L^\infty H^{N'+1+n}} \lesssim \frac{\epsilon^2}{c\alpha}.
\end{aligned}$$

In the last line above we have used (5.9) and the assumption that $N' + 1 + n \leq \tilde{N}$. The treatment of all the other terms encountered in Section 5.6 generalizes similarly.

5.8 Low norm energy estimates

In this section we improve (4.24a) and (4.24b). We provide the details only for (4.24a), as the estimate of $F^3_{\neq}$ follows similarly. An energy estimate gives

$$\begin{aligned}
&\frac{1}{2}\|\tilde{A}F^{+,1}_{\neq}(t_2)\|^2_{H^{N''}} + \frac{\nu}{2}\|\tilde{A}\nabla_L F^{+,1}_{\neq}\|^2_{L^2H^{N''}} + \frac{1}{2}\|\tilde{m}\lambda\sqrt{-\dot{M}M}F^{+,1}_{\neq}\|^2_{L^2H^{N''}} \\
&\leq \frac{1}{2}\|\tilde{A}F^{+,1}_{\neq}(t_1)\|^2_{H^{N''}} - \int_{t_1}^{t_2}\left\langle \tilde{A}F^{+,1}_{\neq}, T^t_{2\alpha}\tilde{A}F^{-,2}_{\neq}\right\rangle_{H^{N''}} dt \\
&+ \int_{t_1}^{t_2}\left\langle \tilde{A}F^{+,1}_{\neq}, \tilde{A}\partial_{XX}\Delta_L^{-1}F^{+,2}_{\neq}\right\rangle_{H^{N''}} dt \\
&+ \int_{t_1}^{t_2}\left\langle \tilde{A}F^{+,1}_{\neq}, \tilde{A}\partial_{XX}\Delta_L^{-1}T^t_{2\alpha}F^{-,2}_{\neq}\right\rangle_{H^{N''}} dt \\
&+ \int_{t_1}^{t_2}\left\langle \tilde{A}F^{+,1}_{\neq}, \partial_X\tilde{A}(T^t_{2\alpha}\partial^L_j Z^{-,i}\partial^L_i Z^{+,j})\right\rangle_{H^{N''}} dt \\
&- \int_{t_1}^{t_2}\left\langle \tilde{A}F^{+,1}_{\neq}, \tilde{A}(T^t_{2\alpha}Z^{-}\cdot\nabla_L F^{+,1})\right\rangle_{H^{N''}} dt \\
&- \int_{t_1}^{t_2}\left\langle \tilde{A}F^{+,1}_{\neq}, \tilde{A}(T^t_{2\alpha}F^{-}\cdot\nabla_L Z^{+,1})\right\rangle_{H^{N''}} dt \\
&- 2\int_{t_1}^{t_2}\left\langle \tilde{A}F^{+,1}_{\neq}, \tilde{A}(T^t_{2\alpha}\partial^L_i Z^{-,j}\partial^L_{ij}Z^{+,1})\right\rangle_{H^{N''}} dt \\
&= \frac{1}{2}\|\tilde{A}F^{+,1}_{\neq}(t_1)\|^2_{H^{N''}} + \mathrm{LU} + \mathrm{LP1} + \mathrm{LP2} + \mathrm{NLP} + \mathrm{NLT} + \mathrm{NLS1} + \mathrm{NLS2},
\end{aligned}$$

where we have skipped the step of absorbing LS and L_λ into the left-hand side since it is done in the same manner as in previous estimates. Moreover, the linear stretch and linear pressure terms are dealt with exactly as in Section 5.1, and so we skip them below.

5.8.1 Nonlinear terms

Since N'' is chosen sufficiently smaller than N', we see that NLT, NLS1, and NLS2 can each be controlled in the same manner as the LH interaction of the associated term in Section 5.1. For example, for NLT$(2, \neq, \neq)$ we have the estimate

$$\begin{aligned}
|\text{NLT}(2, \neq, \neq)| &= \left| \int_{t_1}^{t_2} \int \langle \nabla \rangle^{N''} \tilde{A} F^1_{\neq} \langle \nabla \rangle^{N''} \tilde{A}(Z^2_{\neq} \partial_Y^L F^1_{\neq}) dV \, dt \right| \\
&\lesssim \int_{t_1}^{t_2} \|\tilde{A} F^1_{\neq}\|_{H^{N''}} \| \langle t \rangle^{-1} J F^2_{\neq}\|_{H^{N''+1}} \|\nabla_L F^1_{\neq}\|_{H^{N''}} dt \\
&\lesssim \nu^{-1/3} \|\tilde{A} F^1_{\neq}\|_{L^\infty H^{N''}} \|J F^2_{\neq}\|_{L^2 H^{N''+1}} \|\tilde{A} \nabla_L F^1_{\neq}\|_{L^2 H^{N''}} \\
&\lesssim \epsilon^3 \nu^{-1/3} \nu^{-1/6} \nu^{-1/2} = \epsilon^3 \nu^{-1}.
\end{aligned}$$

The nonlinear pressure terms also follow relatively easily due to the low regularity. Carrying out the calculations as just described completes the estimate of the nonlinear terms and moreover yields the following lemma, which will be useful in the controlling the lift-up term using integration by parts in time.

Lemma 5.3. *Let $j \in \{1, 3\}$ and suppose that $G : [t_1, t_2] \times \mathbb{T} \times \mathbb{R} \times \mathbb{T} \to \mathbb{R}$ satisfies $G \in C([t_1, t_2]; H^{s'})$ for every $s' \geq 0$. If*

$$\|G_{\neq}\|_{L^\infty H^{N''}} + \nu^{1/6} \|G_{\neq}\|_{L^2 H^{N''}} + \nu^{1/2} \|\nabla_L G_{\neq}\|_{L^2 H^{N''}} \lesssim \epsilon,$$

then

$$\left| \int_{t_1}^{t_2} \langle A_j (\partial_t F^j)_{\mathcal{NL}}, G_{\neq} \rangle_{H^{N''}} \, dt \right| \lesssim \epsilon^3 \nu^{-1}, \tag{5.21}$$

where $A_1 = \tilde{A}$ *and* $A_3 = A$.

5.8.2 Lift-up term

Now we verify that by allowing the modes of $F^1_{\neq}$ to grow indefinitely after the critical time (quantified by the use of $\tilde{m}$ as opposed to m in the norm for $F^1_{\neq}$) we can treat the lift-up term with no losses. As in the previous sections, we integrate by parts in time to rewrite

$$-\mathrm{LU} = \frac{1}{2\alpha}\left\langle \tilde{A}F^{+,1}_{\neq}(t_2), \partial_\sigma^{-1}T^{t_2}_{2\alpha}\tilde{A}F^{-,2}_{\neq}(t_2)\right\rangle_{H^{N''}} \tag{5.22}$$

$$-\frac{1}{2\alpha}\left\langle \tilde{A}F^{+,1}_{\neq}(t_1), \partial_\sigma^{-1}T^{t_1}_{2\alpha}\tilde{A}F^{-,2}_{\neq}(t_1)\right\rangle_{H^{N''}} \tag{5.23}$$

$$-\frac{1}{\alpha}\int_{t_1}^{t_2}\left\langle \tilde{A}F^{+,1}_{\neq}, \frac{\dot{\tilde{A}}}{\tilde{A}}\partial_\sigma^{-1}T^{t}_{2\alpha}\tilde{A}F^{-,2}_{\neq}\right\rangle_{H^{N''}} dt \tag{5.24}$$

$$-\frac{1}{2\alpha}\int_{t_1}^{t_2}\left\langle \tilde{A}\partial_t F^{+,1}_{\neq}, \partial_\sigma^{-1}T^{t}_{2\alpha}\tilde{A}F^{-,2}_{\neq}\right\rangle_{H^{N''}} dt \tag{5.25}$$

$$-\frac{1}{2\alpha}\int_{t_1}^{t_2}\left\langle \tilde{A}F^{+,1}_{\neq}, \partial_\sigma^{-1}T^{t}_{2\alpha}\tilde{A}\partial_t F^{-,2}_{\neq}\right\rangle_{H^{N''}} dt. \tag{5.26}$$

Due to $\tilde{m} \leq m$, the estimate of the boundary terms is the same as in Section 5.6.1, and so we skip it and move on to (5.24). We split (5.24) into

$$(5.24) = \tilde{T}_1 + \tilde{T}_2 + \tilde{T}_3,$$

corresponding to

$$\frac{\dot{\tilde{A}}}{\tilde{A}} = \frac{\dot{M}}{M} + \delta\nu^{1/3} + \frac{\dot{\tilde{m}}}{\tilde{m}}.$$

By (3.12), (4.6f), and $\tilde{m} \le m$ we have

$$|\tilde{T}_1| \lesssim \frac{1}{c\alpha}\int_{t_1}^{t_2} \|\tilde{A}F_{\neq}^{+,1}\|_{H^{N''}} \frac{1}{\langle t\rangle}\|\sqrt{-\dot{M}M}JF_{\neq}^{-,2}\|_{H^{N''+n+2}}dt$$
$$\lesssim \frac{1}{c\alpha}\|\tilde{A}F_{\neq}^{+,1}\|_{L^\infty H^{N''}}\|\sqrt{-\dot{M}M}JF_{\neq}^{-,2}\|_{L^2 H^{N'}} \lesssim C_0\frac{\epsilon^2}{c\alpha},$$

which suffices for $\alpha \gg C_0/c$. The estimate of $\tilde{T}_3$ follows from similar techniques and the fact that $|\dot{\tilde{m}}/\tilde{m}| \lesssim \langle t\rangle^{-1}|k,\eta,\ell|$ on its support:

$$|\tilde{T}_3| \lesssim \frac{1}{c\alpha}\int_{t_1}^{t_2} \|\tilde{A}F_{\neq}^{+,1}\|_{H^{N''}} \frac{1}{\langle t\rangle^2}\|JF_{\neq}^{-,2}\|_{H^{N''+n+3}}dt$$
$$\lesssim \frac{1}{c\alpha}\|\tilde{A}F_{\neq}^{+,1}\|_{L^\infty H^{N''}}\|JF_{\neq}^{-,2}\|_{L^\infty H^{N'}} \lesssim C_0\frac{\epsilon^2}{c\alpha}.$$

Controlling $\tilde{T}_2$ is the same as the analogous term in Section 5.6.1, and so we omit the details. This completes the estimate of (5.24).

Next we consider (5.26). The linear terms do not require any methods beyond what we have employed thus far, and so we will only sketch how to deal with them. The dissipation terms are easily estimated using integration by parts as in Section 5.6.1. Next, both $\partial_t F_{\neq}^{+,1}$ and $\partial_t F_{\neq}^{-,2}$ contain a LS (or OLS) term. Each of these terms carries a factor of $S(t)$ (recall the notation defined in Section 5.6.1), and so by using (5.9) we bound these terms as we did $\tilde{T}_3$. There are also linear contributions from LP1 and LP2 in $\partial_t F_{\neq}^{+,1}$. Using that $|\partial_{XX}\Delta_L^{-1}| \lesssim -\dot{M}M$, these are terms bounded like $\tilde{T}_1$ above. Lastly, there is a linear term that arises from the lift-up term in $\partial_t F_{\neq}^{+,1}$, but this term vanishes like the analogous term did in Section 5.5.2.

Now we turn to the nonlinear terms created in (5.26), which we write as (dropping the irrelevant minus signs and factors of α^{-1})

$$\begin{aligned}\int_{t_1}^{t_2} \left\langle \tilde{A}(\partial_t F^{+,1})_{\mathcal{NL}}, \partial_\sigma^{-1} T^t_{2\alpha} \tilde{A} F^{-,2}_{\neq} \right\rangle_{H^{N''}} dt \\ + \int_{t_1}^{t_2} \left\langle \tilde{A} F^{+,1}_{\neq}, \partial_\sigma^{-1} T^t_{2\alpha} \tilde{A}(\partial_t F^{-,2})_{\mathcal{NL}} \right\rangle_{H^{N''}} dt \\ = \mathcal{NL}_1 + \mathcal{NL}_2.\end{aligned}$$

Since $\|\partial_\sigma^{-1} T^t_{2\alpha} \tilde{A} F^2_{\neq}\|_{H^{N''}} \lesssim \|A F^2_{\neq}\|_{H^N}$ it follows by (4.22c), (4.22d), and Lemma 5.3 that $|\mathcal{NL}_1| \lesssim \epsilon^3 \nu^{-1}$. The term $\mathcal{NL}_2$ is less immediate since the loss of regularity caused by ∂_σ^{-1} implies that we must appeal to bootstrap hypotheses (4.22a) and (4.22b). We first notice that by an integration by parts, (3.12), and (4.6f) there holds

$$\begin{aligned}|\mathcal{NL}_2| &\lesssim \int_{t_1}^{t_2} \|\nabla_L \tilde{A} F^1_{\neq}\|_{H^{N''}} \langle t \rangle^{-1} \|\lambda \langle t \rangle^{-1} \nabla_L (Z \cdot \nabla_L Z^2)_{\neq}\|_{H^{N''+n+4}} dt \\ &\quad + \int_{t_1}^{t_2} \|\tilde{A} F^1_{\neq}\|_{H^{N''}} \langle t \rangle^{-1} \|\lambda \langle t \rangle^{-1} \partial_Y^L (\partial_j^L Z^i \partial_i^L Z^j)_{\neq}\|_{H^{N''+n+4}} dt \\ &\lesssim \epsilon \nu^{-1/2} \|\lambda \langle t \rangle^{-1} \nabla_L (Z \cdot \nabla_L Z^2)_{\neq}\|_{L^\infty H^{N''+n+4}} \\ &\quad + \epsilon \nu^{-1/6} \|\lambda \langle t \rangle^{-1} \partial_Y^L (\partial_j^L Z^i \partial_i^L Z^j)_{\neq}\|_{L^\infty H^{N''+n+4}},\end{aligned}$$

and hence to complete the desired estimate under the assumptions of Theorem 3.3 it suffices to prove that

$$\|\lambda \langle t \rangle^{-1} \nabla_L (Z \cdot \nabla_L Z^2)_{\neq}\|_{L^\infty H^{N''+n+4}} \lesssim \epsilon^2 \nu^{-1/2}, \tag{5.27}$$

$$\|\lambda \langle t \rangle^{-1} \partial_Y^L (\partial_j Z^i \partial_i Z^j)_{\neq}\|_{L^\infty H^{N''+n+4}} \lesssim \epsilon^2 \nu^{-5/6}. \tag{5.28}$$

To prove (5.27) we use $N'' + n + 5 \le N' \le N$ to obtain

$$\begin{aligned}
\|\lambda \langle t\rangle^{-1} \nabla_L (Z{\cdot}\nabla_L Z^2)_{\neq}\|_{H^{N''+n+4}} &\lesssim \|\lambda(Z\cdot\nabla_L Z^2)_{\neq}\|_{H^{N''+n+5}} \\
&\lesssim \|\lambda Z_{\neq}\|_{H^N}\|\nabla_L Z^2\|_{H^{N'}} + \|Z_0\|_{H^N}\|\lambda\nabla_L Z^2_{\neq}\|_{H^{N'}} \\
&\lesssim (\|\lambda Z_{\neq}\|_{H^N} + \|Z_0\|_{H^N})(\|JF^2_{\neq}\|_{H^{N'}} + \|Z_0\|_{H^N}) \\
&\lesssim \epsilon^2\nu^{-1/3},
\end{aligned}$$

which suffices. For (5.28) we have

$$\begin{aligned}
\|\lambda \langle t\rangle^{-1} \partial_Y^L(\partial_j^L Z^i \partial_i^L Z^j)_{\neq}\|_{H^{N''+n+4}} &\lesssim \|\lambda(\partial_j^L Z^i\partial_i^L Z^j)_{\neq}\|_{H^{N''+n+5}} \\
&\lesssim \epsilon^2\nu^{-2/3},
\end{aligned}$$

where the last line follows by, as in previous estimates, using (4.6b), (4.6c), and considering separately the cases $i,j\in\{1,3\}$, $i=2$ and $j\neq 2$, and $i=j=2$.

5.9 Zero mode velocity estimates

In this section we improve (4.25). For any $r \in \{1, 2, 3\}$, an energy estimate gives

$$\begin{aligned}
&\frac{1}{2}\|Z_0^{+,r}(t_2)\|_{H^N}^2 + \nu\|\nabla Z_0^{+,r}\|_{L^2H^N}^2 = \frac{1}{2}\|Z_0^{+,r}(t_1)\|_{H^N}^2 \\
&- \mathbf{1}_{r=1}\int_{t_1}^{t_2}\int \langle\nabla\rangle^N Z_0^{+,1} \langle\nabla\rangle^N T_{2\alpha}^t Z_0^{-,2} dV dt \\
&- \int_{t_1}^{t_2}\int \langle\nabla\rangle^N Z_0^{+,r} \langle\nabla\rangle^N \left(T_{2\alpha}^t Z^- \cdot \nabla_L Z^{+,r} dV dt\right) \\
&+ \mathbf{1}_{r\neq 1}\int_{t_1}^{t_2}\int \langle\nabla\rangle^N Z_0^{+,r} \langle\nabla\rangle^N \partial_r\Delta^{-1}(T_{2\alpha}^t\partial_j Z^{-,i}\partial_i Z^{+,j})_0 dV dt \\
&= \frac{1}{2}\|Z_0^{+}(t_1)\|_{H^N}^2 + \mathrm{LU} + \mathrm{NLT} + \mathrm{NLP}.
\end{aligned}$$

The lift-up term can be dealt with using integration by parts in time as in Section 5.5.2. We skip it and turn to the nonlinear terms, beginning with the transport term

$$\mathrm{NLT}(j) = \int_{t_1}^{t_2}\int \langle\nabla\rangle^N Z_0^r \langle\nabla\rangle^N (Z^j\partial_j Z^r)_0 dV dt.$$

First we treat the interaction between the nonzero modes. When $j = 1, 3$ we have

$$\begin{aligned}
|\mathrm{NLT}(j,\neq,\neq)| &\lesssim \|Z_0^r\|_{L^\infty H^N}\|\tilde{A}F_{\neq}^j\|_{L^2H^N}\|\tilde{A}F_{\neq}^r\|_{L^2H^N} \\
&\lesssim \epsilon^3\nu^{-1/2}\nu^{-1/2} = \epsilon^3\nu^{-1},
\end{aligned}$$

while for $j = 2$ there holds

$$|\text{NLT}(2, \neq, \neq)| \lesssim \|Z_0^r\|_{L^\infty H^N} \|AF_{\neq}^2\|_{L^2 H^N} \nu^{-1/3} \|\tilde{A} F_{\neq}^r\|_{L^2 H^N}$$
$$\lesssim \epsilon^3 \nu^{-1/6} \nu^{-1/3} \nu^{-1/2} = \epsilon^3 \nu^{-1}.$$

For the interaction between the zero modes, we use the divergence free condition to obtain

$$|\text{NLT}(2, 0, 0)| \lesssim \|Z_0^r\|_{L^\infty H^N} \|\nabla Z_0^2\|_{L^2 H^N} \|\nabla Z_0^r\|_{L^2 H^N}$$
$$\lesssim \epsilon^3 \nu^{-1/2} \nu^{-1/2} = \epsilon^3 \nu^{-1}.$$

The NLT(3,0,0) term is bounded similarly by employing the method used to treat the associated term in Sec. 5.5.1. We omit the details. Turning now to the pressure, we observe that by using incompressibility and integration by parts it can be written as

$$\text{NLP}(i, j) = -\int_{t_1}^{t_2} \int \langle \nabla \rangle^N \partial_r Z_0^r \langle \nabla \rangle^N \partial_{ij} \Delta^{-1} (Z^i Z^j)_0 dV dt,$$

where the term is nonzero only for $i, j \in \{2, 3\}$. For the interaction between the nonzero modes we use that $\partial_{ij}\Delta^{-1}$ is bounded on H^N along with a paraproduct decomposition to obtain

$$|\text{NLP}(i, j, \neq, \neq)| \lesssim \|\nabla Z_0^r\|_{L^2 H^N} \|Z_{\neq}^j\|_{L^2 H^{N''}} \|Z_{\neq}^i\|_{L^\infty H^N} + \text{ symmetric term}$$
$$\lesssim \epsilon^3 \nu^{-1/2} \nu^{-1/6} \nu^{-1/3} = \epsilon^3 \nu^{-1},$$

which is consistent. For NLP$(i, j, 0, 0)$ we use that $i, j \in \{2, 3\}$ along with incompressibility implies that at least one of Z_0^i or Z_0^j has a nonzero Z-frequency. Thus, there holds

$$\begin{aligned}|\text{NLP}(i, j, 0, 0)| &\lesssim \|\nabla Z_0^r\|_{L^2 H^N} \|\nabla Z_0^i\|_{L^2 H^N} \|Z_0^j\|_{L^\infty H^N} + \text{ symmetric term} \\ &\lesssim \epsilon^3 \nu^{-1/2} \nu^{-1/2} = \epsilon^3 \nu^{-1}.\end{aligned}$$

Chapter 6: Introduction to Part II

In this part of the dissertation, we study the quantitative (with respect to ϵ) smoothing properties and exponential convergence to equilibrium for a class of hypoelliptic PDEs on $\mathbb{R}^d$ of the form

$$\partial_t \mu_t = L_\epsilon^* \mu_t := \epsilon \sum_{j=1}^{r} (Z_j \cdot \nabla)^2 \mu_t + \epsilon \nabla \cdot (A x \mu_t) + \epsilon^\alpha B x \cdot \nabla \mu_t + N \cdot \nabla \mu_t \tag{6.1}$$

for parameters $0 < \epsilon \ll 1$ and $\alpha \geq 0$. Here, $\{Z_j\}_{j=1}^r$ is a collection of constant vector fields on $\mathbb{R}^d$ ($r \ll d$ in general), $A \in \mathbb{M}^{d\times d}$ (the vector space of $d \times d$ matrices with real entries) is a positive definite matrix that plays the role of dissipation, $B \in \mathbb{M}^{d\times d}$ is skew-symmetric (possibly zero), and $N : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth, nonlinear drift such that $N(x) := N(x, x, \ldots, x)$ for a multlinear function $N(x_1, x_2, \ldots, x_p)$ of $p \geq 2$ arguments. We assume moreover that $N(x)$ is divergence free and obeys the energy conservation property

$$N(x) \cdot x = 0 \quad \forall x \in \mathbb{R}^d. \tag{6.2}$$

The skew-symmetry of B implies that Bx satisfies (6.2) and $\nabla \cdot Bx = \mathrm{Tr}(B) = 0$, so that the term $\epsilon^\alpha Bx$ plays the role of a linear (and lower order when $\alpha > 0$) conservative drift.

Equation (6.1) arises naturally as the forward Kolmogorov for the diffusion process on $\mathbb{R}^d$

$$dx_t^\epsilon = -\epsilon A x_t^\epsilon dt - \epsilon^\alpha B x_t^\epsilon dt - N(x_t^\epsilon)dt + \sum_{j=1}^{r} \sqrt{2\epsilon} Z_j dW_t^{(j)}, \tag{6.3}$$

where $\left\{W_t^{(j)}\right\}_{j=1}^r$ are independent one-dimensional Wiener processes on a common filtered probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The form of (6.3) is fairly general and captures a number of fundamental dynamical systems driven by a (possibly degenerate) white-in-time forcing, such as Lorenz-96 [96] and Galerkin truncations of the Navier-Stokes equations (see [101] and Section 6.2 below). Notice that we have chosen the scaling for which the noise and dissipation are balanced in (6.3), so that can hope to prove bounds on the equilibrium density that do not depend on ϵ. Due to the homogeneity of N, by rescaling time and x_t^ϵ, treating this scaling also implies corresponding statements for both the large forcing and the small dissipation cases. For instance, when $B = 0$ and N is quadratic (which is the case for instance in the Navier-Stokes equations), rescaling $x_t \to \epsilon^{-1/2} x_{\sqrt{\epsilon} t}$ and then redefining $\epsilon^{3/2} \to \epsilon$ shows that (6.3) is equivalent to

$$dx_t^\epsilon = -\epsilon A x_t^\epsilon dt - N(x_t^\epsilon)dt + \sum_{j=1}^{r} \sqrt{2} Z_j dW_t^{(j)}. \tag{6.4}$$

In this way, for many important examples, studying the limit $\epsilon \to 0$ in (6.3) is equivalent to the limit commonly considered for chaotic/turbulent systems where the noise and conservative drift are fixed $\mathcal{O}(1)$ and the dissipation strength is sent to zero.

By *hypoelliptic*, we mean that while $\{Z_j\}_{j=1}^r$ may not span $\mathbb{R}^d$, we assume that the set of vector fields $\{\epsilon A x + \epsilon^\alpha B x + N, Z_1, ..., Z_r\}$ satisfies the classical *parabolic Hörmander condition* (see discussions in e.g. [68] and the references therein). For an open set $\Omega \subseteq \mathbb{R}^d$, we write $T(\Omega)$ for the collection of all smooth vector fields defined on Ω. In the remainder of this part of the dissertation, we identity vector fields on $\mathbb{R}^d$ with first-order differential operators. That is, for a function $f : \Omega \to \mathbb{R}$ and a vector field $X \in T(\Omega)$ we write Xf to mean $X \cdot \nabla f$. For $X, Y \in T(\Omega)$ we denote by $[X, Y] \in T(\Omega)$ the vector field

$$[X, Y] = XY - YX.$$

Definition 6.1 (Locally uniform parabolic Hörmander)**.** For an open set $\Omega \subseteq \mathbb{R}^d$ and $\{X_0, X_1, \ldots, X_k\} \subseteq T(\Omega)$, let $V_0 = \{X_1, \ldots, X_k\}$ and

$$V_n = V_{n-1} \cup \{[X_j, Y] : 0 \leq j \leq k, Y \in V_{n-1}\} \quad \forall n \geq 1. \tag{6.5}$$

We say that the family $\{X_0, \ldots, X_k\}$ satisfies the *parabolic Hörmander condition* on Ω if $\cup_{n=0}^{\infty} V_n$ spans $\mathbb{R}^d$ at every point $x \in \Omega$. We say that $\{X_0, \ldots, X_k\}$ satisfies the *uniform parabolic Hörmander condition on* Ω *with constants* $(N_0, C_0) \in \mathbb{N} \times (0, \infty)$ if for every $x \in \Omega$ there exists a set $\{Y_i\}_{i=1}^d \subseteq V_{N_0}$ such that

$$|\det(Y_1(x), Y_2(x), \ldots, Y_d(x))|^{-1} \leq C_0. \tag{6.6}$$

In many settings, especially time-independent ones, it is natural to allow X_0 in the definition of V_0 above. In this case, if $\cup_{n=0}^{\infty} V_n$ spans $\mathbb{R}^d$ at every point $x \in \Omega$, then $\{X_0, \ldots, X_k\}$ is said to satisfy *Hörmander's condition* on Ω. The notion of uniformity extends in the obvious way.

Remark 13. Our interest in the case there $r \ll d$ is motivated primarily by hydrodynamic turbulence, where it is typical in experiments for fluctuations to be injected only at a few fixed characteristic length scales (see e.g. [55]). Physically, Hörmander bracket conditions ensure that the randomness injected into the system by the stochastic forcing can spread to all degrees of freedom through the action of the drift.

Under our assumptions on the vector fields in L_ϵ^*, for any initial condition $x_0^\epsilon = x \in \mathbb{R}^d$ the SDE (6.3) admits a unique, global solution $(x_t^\epsilon)_{t\geq 0}$ that defines a Markov process with generator

$$L_\epsilon := \epsilon \sum_{j=1}^{r} (Z_j \cdot \nabla)^2 - Ax \cdot \nabla - \epsilon^\alpha Bx \cdot \nabla - N \cdot \nabla. \tag{6.7}$$

We denote the transition probabilities for x_t^ϵ as $\mathcal{P}_t^\epsilon(x, A) = \mathbf{P}(x_t^\epsilon \in A | x_0^\epsilon = x)$ for all $A \in \mathcal{B}(\mathbb{R}^d)$ (the set of Borel sets) and $x \in \mathbb{R}^d$. The family of measures $\{\mathcal{P}_t^\epsilon(x, \cdot)\}_{x\in\mathbb{R}^d}$ defines a *Markov transition kernel* (for a precise definition, see the bullets preceding Lemma 7.3). The Markov semigroup $\mathcal{P}_t^\epsilon$ is then defined to act on bounded measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ by

$$\mathcal{P}_t^\epsilon f(x) := \int_{\mathbb{R}^d} f(y) \mathcal{P}_t(x, dy), \tag{6.8}$$

and a measure $\mu \in \mathcal{M}(\mathbb{R}^d)$ (the space of Borel probability measures on $\mathbb{R}^d$) is called *stationary* (or *invariant*) for $\mathcal{P}_t^\epsilon$ if for every $A \in \mathcal{B}(\mathbb{R}^d)$ one has

$$(\mathcal{P}_t^\epsilon)^* \mu(A) := \int_{\mathbb{R}^d} \mathcal{P}_t^\epsilon(x, A) \mu(dx) = \mu(A). \tag{6.9}$$

It is well known that if $\{\epsilon Ax + \epsilon^\alpha Bx + N, Z_1, \ldots, Z_r\}$ satisfies the parabolic Hörmander condition on $\mathbb{R}^d$, then the semigroup $(\mathcal{P}_t^\epsilon)^*$ generated by L_ϵ^* is instantly smoothing, despite the fact that L_ϵ^* is not elliptic. In this case, for every $\epsilon > 0$ there is a unique probability measure μ_ϵ solving $L_\epsilon^* \mu_\epsilon = 0$, and moreover μ_ϵ is absolutely continuous with respect to Lebesgue measure with a smooth density f_ϵ. The measure μ_ϵ is also the unique stationary measure for the Markov semigroup $\mathcal{P}_t^\epsilon$ generated by (6.3). For precise statements and proof sketches of the claims in this paragraph, see the lemmas in Chapter 7.

Finally, it is known that $\forall \epsilon > 0$ the semigroup $\mathcal{P}_t^\epsilon$ converges exponentially on suitable weighted spaces. In particular, let $V : \mathbb{R}^d \to [0, \infty)$ be continuous and for measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ define

$$\|f\|_V = \sup_{x \in \mathbb{R}^d} \frac{|f(x)|}{1 + V(x)}. \tag{6.10}$$

Then, it is known that if $V \in C^2(\mathbb{R}^d)$ has compact level sets and satisfies

$$L_\epsilon V \leq -\theta V + \lambda$$

for some constants $\lambda, \theta > 0$, i.e., V satisfies a drift condition, then there exists $C_\epsilon, \gamma_\epsilon > 0$ (both depending on ϵ) such that for all measurable $f : \mathbb{R}^d \to \mathbb{R}$ with $\|f\|_V < \infty$ there holds

$$||\mathcal{P}_t^\epsilon f - \mu_\epsilon(f)||_V \leq C_\epsilon e^{-\gamma_\epsilon t} \, ||f - \mu_\epsilon(f)||_V \,, \tag{6.11}$$

where we have written $\mu_\epsilon(f) = \int f d\mu_\epsilon$. Even for infinite-dimensional analogs of (6.3) (for e.g. complex Ginzburg-Landau in $\mathbb{T}^d$, Navier-Stokes in $\mathbb{T}^2$, or hyper-viscous Navier-Stokes in $\mathbb{T}^3$), the existence and uniqueness of a stationary measure μ_ϵ (see e.g. [53, 69, 71, 86–88]) with smooth finite-dimensional projections [69, 103] is known for $\epsilon > 0$, as are exponential convergence results similar to (6.11) (see e.g. [62, 63, 72, 86] and the references therein).

In general, it is a very important question to understand the limit $\epsilon \to 0$, both to characterize properties of μ_ϵ and to quantify $C_\epsilon, \gamma_\epsilon$ as functions of ϵ. In the case of e.g. the (infinite dimensional) 3D stochastic Navier-Stokes equations, characterizing μ_ϵ as $\epsilon \to 0$ is equivalent to understanding many properties of turbulence in the statistically stationary regime, whereas quantifying γ_ϵ amounts to estimating the convergence rate to statistical equilibrium with respect to the fluid viscosity in the inviscid limit, also a question of fundamental importance to the theory of turbulence. In spite of its importance, little is known about quantifying γ_ϵ, C_ϵ. In finite dimensions, it is not difficult to deduce (see Theorem 6.7 below) that $\gamma_\epsilon \lesssim \epsilon$, but lower bounds are much harder to come by. In infinite dimensions, the methods of e.g. [72] yield a lower bound on γ_ϵ that is exponentially bad in ϵ, even if one takes non-degenerate stochastic forcing. The situation in finite dimensions is not significantly better, as standard proofs of convergence to equilibrium for (6.3) similarly yield a spectral gap that is not even polynomial in ϵ.

A key reason that previously proven lower bounds on γ_ϵ (in either finite or infinite dimensions) are far from optimal is a lack of quantitative irreducibility results. It is well-known that unique ergodicity and the convergence rate to equilibrium for a Markov semigroup is in part determined by its irreducibility properties [73, 104, 105], in particular the extent to which the support of transition probabilities arising from distinct points either overlap (see e.g. [Assumption 2, [73]]) or become arbitrarily close to each other (see e.g. [70] and [Assumption 6, [72]]). The former is common for finite-dimensional diffusions, while in degenerate, infinite-dimensional settings one often must resort to the latter. In most of the previous works, the irreducibility is obtained by taking advantage of rare events in the forcing. Previous works, such as e.g. [72, 126], use that for any initial condition there is a small probability that the energy input from the noise is low enough that the dissipation causes the process to drift back to any neighborhood of the origin. Along with some regularity of transition probabilities and a suitable Lyapunov structure, this is a strong enough irreducibility statement to prove exponential convergence statements such as (6.11). However, rare excursions to the origin are clearly not the actual mechanism for irreducibility in high-dimensional, chaotic systems, and as such the estimates on the mixing time one obtains from such an analysis are sub-optimal [89]. More sophisticated approaches to irreducibility rely on Hörmander's condition and optimal control theory (see e.g. [2, 61] and the references therein). However, these arguments similarly rely on rare events where the diffusion completely dominates the drift and hence still do not yield any type of uniform-in-ϵ irreducibility for the transition probabilities of (6.3), nor do they capture true mechanisms behind mixing in the fluctuation dissipation limit.

While improving estimates on γ_ϵ in infinite dimensions seems to be an extremely challenging problem and currently out of reach, in this portion of the dissertation we rectify the above issue in finite dimensions and obtain the optimal estimates $\gamma_\epsilon \approx \epsilon$ and C_ϵ independent of ϵ. As a necessary step, we also obtain a uniform-in-ϵ, pointwise Gaussian upper bound on f_ϵ (the density of the stationary measure) and moreover obtain a uniform-in-ϵ, pointwise lower bound on every compact set.

One example of a quantitative mixing result in a vanishing noise limit similar to what we consider can be found in [48]. This paper studies a system of weakly coupled harmonic oscillators, each subject to weak energy injection and damping described by a parameter $\epsilon > 0$. As a relatively small part of their analysis, the author proves (see Appendix A) uniform-in-ϵ convergence to equilibrium on the timescale ϵ^{-1} in a total variation space. This result is in the same spirit as our main result Theorem 6.5 below, but the techniques used in the proof are not applicable to our situation. In particular, the noise in [48] is sufficiently nondegenerate to allow for the use of Girsanov's theorem. Moreover, after rescaling $t \to \epsilon^{-1}t$ the nonlinearity remains bounded as $\epsilon \to 0$ (due to the assumption of weak coupling), which is not the case for (6.3).

Part II of this dissertation is organized as follows. In the remainder of Chapter 6 we state precisely and discuss our main theorems, and provide some examples of systems to which they apply. In Chapter 7 we give precise statements and proof sketches of some basic facts concerning (6.3), namely global well-posedness, the regularizing properties of the semigroup, and the existence and uniqueness of μ_ϵ. In Chapter 8, outline the proof of Theorem 6.5 assuming Theorem 6.2 and stating intermediate steps as lemmas to be proven in Chapter 11. Chapters 9-11 are dedicated to establishing the results needed to prove Theorem 6.2 and carry out the program described in Chapter 8. This includes a detailed discussion of uniform Hörmander inequalities in Chapter 9.

6.1 Main results and discussion

In the statements of our results below, and throughout the remainder of this part of the dissertation, we write $a \lesssim b$ to mean that $a \le Cb$ for a constant C depending possibly on A, B, N, $\{Z_j\}_{j=1}^r$, and the dimension d, but not on ϵ. Any ϵ dependence in estimates or constants we define will always be made explicit. Also, throughout the entire paper we write B_R for the open ball of radius R centered at the origin.

6.1.1 Statement of main assumptions

We now state precisely our main assumptions. They consist of the dissipative and conservative structures of A, B, and N, a uniform-in-ϵ nondegeneracy condition, and strict (but *qualitative*) positivity of the stationary density f_ϵ for every $\epsilon > 0$.

Assumption 1. *The matrix $A \in \mathbb{M}^{d\times d}$ is positive definite, $B \in \mathbb{M}^{d\times d}$ is skew-symmetric, and $N(x) := N(x, \ldots, x)$ for a smooth, multilinear function of $p \geq 2$ arguments satisfying the conservation properties $\nabla \cdot N(x) = 0$ and (6.2).*

Assumption 2. *For every $R > 0$ there exists $M \in \mathbb{N}$ and $C > 0$ (depending possibly on R) so that for every $\epsilon_1, \epsilon_2 \in [0, 1]$ the collection of vector fields*

$$\{N + \epsilon_1 Ax + \epsilon_2 Bx, Z_1, \ldots, Z_r\}$$

satisfies the uniform parabolic Hörmander condition on B_R with constants (M, C).

Assumption 3. *For every $\epsilon > 0$, the density f_ϵ of the unique invariant measure μ_ϵ is strictly positive.*

We discuss specific examples that satisfy Assumptions 1-3 in Section 6.2. For now, we remark that the uniform-in-ϵ spanning condition of Assumption 2 is quite natural for (6.3). Indeed, one usually verifies the parabolic Hörmander condition by showing that the collection

$$\{Z_1, \ldots, Z_r\} \cup \{Y : Y = [\ldots[N, Z_{i_1}], Z_{i_2}], \ldots], Z_{i_{p-1}}], Z_{i_p}], 1 \leq i_j \leq r\}$$

satisfies Hörmander's condition on $\mathbb{R}^d$ (recall that p denotes the degree of N); see for example [126] and Section 6.2. Since $p > 1$, in this situation the linear drift terms $\epsilon_1 Ax$ and $\epsilon_2 Bx$ do not change the bracket structure and so Assumption 2 is satisfied. Regarding Assumption 3, strict positivity of the stationary density is typically proven by verifying suitable hypoelliptic and control theoretic assumptions (see e.g. [61,76] and the references therein).

Remark 14. Beyond stating our main results in Sections 6.1.2 and 6.1.3 below, throughout the remainder of the dissertation we will always assume that Assumptions 1 and 2 both hold unless remarked otherwise. On the other hand, Assumption 3 is only needed in select locations, and so we will always indicate explicitly when it is required.

6.1.2 Uniform-in-ϵ hypoelliptic estimates

In this section we state quantitative hypoelliptic estimates, in particular uniform-in-ϵ pointwise bounds on the equilibrium density f_ϵ, and a long-time $L^2 \to L^\infty$ regularization estimate for $\mathcal{P}_t^\epsilon$. We are motivated to obtain such bounds mostly to use as lemmas in the proof that $\gamma_\epsilon \approx \epsilon$. However, they are of independent interest, as estimates on hypoelliptic equations that are uniform in a small parameter are a delicate matter. In fact, except for the recent works [15] and [4], the latter appearing after [20] (the paper recounted in this part of the disseration) and the former involving the author's advisor and being completed at the same time as [20], we are not aware of any quantitative hypoelliptic estimates from the literature that are in the same spirit as the results below.

Theorem 6.2 (Uniform-in-ϵ estimates on f_ϵ)**.** *Under Assumptions 1 and 2 there exists $\lambda > 0$ so that the smooth density f_ϵ of the unique stationary measure μ_ϵ for (6.3) satisfies the pointwise estimate*

$$\sup_{\epsilon \in (0,1)} f_\epsilon(x) \lesssim e^{-\lambda |x|^2}. \tag{6.12}$$

If in addition Assumption 3 is satisfied, then we also have the lower bound

$$\inf_{\epsilon \in (0,1)} \inf_{|x| \leq R} f_\epsilon(x) \gtrsim_R 1. \tag{6.13}$$

Remark 15. Without Assumption 3, our proof of (6.13) shows that for every $R > 0$ there exists $\epsilon_*(R) \in (0,1)$ such that

$$\inf_{\epsilon \in (0,\epsilon_*)} \inf_{|x| \leq R} f_\epsilon(x) \gtrsim_R 1.$$

Remark 16. A consequence of the proof of Theorem 6.2 is that we have uniform-in-ϵ control on $\|f_\epsilon\|_{H^s_{\text{loc}}}$ for some s sufficiently small (independent of ϵ). We do not know how to obtain uniform-in-ϵ bounds in higher regularity, nor do we even necessarily expect such bounds to be true. In particular, we do not have continuity in any sense that is uniform as $\epsilon \to 0$.

Remark 17. The uniform-in-ϵ H^s_{loc} control on f_ϵ and Sobolev embedding imply that f_ϵ converges strongly in L^p_{loc} for some $p > 2$ as $\epsilon \to 0$ to a limit f_0 solving $N \cdot \nabla f_0 = 0$ in the sense of distributions. Combined with a uniform-in-ϵ exponential moment bound (see (7.9)), this implies that f_ϵ converges to f_0 in the Wasserstein-1 norm.

Lemma 6.3 (Quantitative $L^2_{\mu_\epsilon} \to L^\infty$ regularization)**.** *Write*

$$\|f\|^2_{L^2_\mu} = \int_{\mathbb{R}^d} |f(x)|^2 \mu(dx)$$

for a measure $\mu \in \mathcal{M}(\mathbb{R}^d)$ and measurable function f. Under Assumptions 1-3, for every bounded and measurable observable $f : \mathbb{R}^d \to \mathbb{R}$ and $R \geq 1$ there holds, uniformly in $\epsilon \in (0, 1)$,

$$\|\mathcal{P}^\epsilon_{\epsilon^{-1}} f\|_{L^\infty(B_R)} \lesssim_R \|f\|_{L^2_{\mu_\epsilon}}. \tag{6.14}$$

Remark 18. One should view Lemma 6.3 as saying that for long times the semigroup $\mathcal{P}^\epsilon_t$ satisfies a smoothing estimate that scales in the same way as the basic $L^2 \to L^\infty$ regularization estimate for the heat equation $\partial_t u = \epsilon \Delta u$.

There has been a lot of work on the regularity of solutions to hypoelliptic and degenerate parabolic equations, especially Harnack inequalities [1, 36, 64, 84], pointwise bounds on fundamental solutions [90, 114], and Hölder/L^∞_{loc} estimates [6, 64, 113, 122, 127] inspired by the classical De Giorgi-Nash-Moser theory for elliptic PDEs [59]. Most of the work just mentioned has been concerned however with reducing regularity requirements on the coefficients, which is a signficantly different from the goal of our present study, that is, understanding quantitative dependence on a small parameter in a setting with smooth coefficients. Nevertheless, our work makes important use of some ideas from recent work in kinetic theory [64, 107] (see also the earlier preprints [65, 80]). These works are focused on adapting De Giorgi-Nash-Moser methods to kinetic equations, with the key starting point being a local gain of integrability available from velocity averaging lemmas [27].

Besides adapting ideas from [64], our proof of Theorem 6.2 and Lemma 6.3 is based on a careful study of functional inequalities derived by Hörmander in the original proof of his celebrated hypoellipticity theorem [79]. These estimates, now typically referred to as *Hörmander inequalities*, allow one to control a fractional Sobolev norm with the a priori bounds available for second order equations such as (6.1) (see e.g. Lemma 9.4). We obtain the upper bounds (6.12) and (6.14) with hypoelliptic Moser iterations. In both cases, the main difficulty is to obtain a local gain of integrability that does not depend on ϵ. For this, we derive suitable uniform Hörmander inequalities (see Chapter 9) and make careful use of the structural assumptions $\nabla \cdot N = 0$ and $N(x) \cdot x = 0$. The parabolic version (Lemma 9.5) needed for Lemma 6.3 is the most delicate, and while the proof does not require deep modifications to Hörmander's original methods, to our knowledge nothing quite analogous can be found in the literature. An additional challenge for (6.12) as compared to previous works such as [64,65,80], is that to close the iteration scheme we must deduce a uniform-in-ϵ upper bound on $\|f_\epsilon\|_{L^2}$. This requires using moment bounds available from the energy conservation property $N(x)\cdot x = 0$ and a Hörmander inequality that is quantitative in the diameter of the set to which it is localized (Lemma 9.4).

The main challenge in the proof of the lower bound (6.13) is adapting to our setting a compactness-rigidity argument used in [65, 80] to prove a De Giorgi "intermediate value lemma" for subsolutions of a kinetic Fokker-Planck equation. In particular, we prove a uniform-in-ϵ intermediate value inequality for solutions to $L_\epsilon^* w \geq 0$ obeying some additional conditions (see Lemma 10.3). This is then combined with a quantitative hypoelliptic $L^2 \to L^\infty$ estimate for subsolutions (Lemma 10.2) and classical ideas from the De Giorgi elliptic theory [119]. It is an interesting problem to determine if solutions to (6.1) have uniform-in-ϵ Hölder regularity (or any modulus of continuity, for that matter). A positive answer to this question remains out of reach with our current techniques and a direction of future interest, since our present methods rely crucially on $N(x) \cdot x = 0$ and hence are not invariant under translations.

6.1.3 Quantitative geometric ergodicity and consequences

In this section, we give precise statements of our results on the geometric ergodicity of (6.3). First, we need to define an appropriate notion of a uniform drift condition.

Definition 6.4. We say that a nonnegative function $V \in C^2(\mathbb{R}^d)$ with

$$\lim_{|x|\to\infty} V(x) = \infty$$

is a *uniform Lyapunov function for* $\{\mathcal{P}_t^\epsilon\}_{\epsilon\in(0,1)}$ if there exists $\kappa, b > 0$ so that for all $\epsilon \in (0,1)$ and $\delta \in [0,1]$ there holds

$$\epsilon\delta\Delta V + L_\epsilon V \leq -\epsilon\kappa V + \epsilon b. \tag{6.15}$$

We include the term $\epsilon\delta\Delta V$ on the left-hand side since at times it will be convenient to work with the regularized operator $\epsilon\delta\Delta + L_\epsilon^*$. It is easy to check that $V(x) = e^{\gamma x^2}$ is a uniform Lyapunov function provided that γ is chosen sufficiently small. Along with the notation defined in (6.10) we then have the following.

Theorem 6.5 (Quantitative geometric ergodicity)**.** *Let V be a uniform Lyapunov function for $\{\mathcal{P}_t^\epsilon\}_{\epsilon\in(0,1)}$. Under Assumptions 1-3 there exists $K, \delta > 0$ that do not depend on ϵ such that for all $\epsilon \in (0,1)$, $t > 0$, and measurable $f : \mathbb{R}^d \to \mathbb{R}$ satisfying $\|f\|_V < \infty$ there holds*

$$\|\mathcal{P}_t^\epsilon f - \mu_\epsilon(f)\|_V \le K e^{-\epsilon\delta t}\|f - \mu_\epsilon(f)\|_V. \tag{6.16}$$

Remark 19. Theorem 6.5 has implications for the spectrum of $\mathcal{P}_t^\epsilon$. To make this precise, let V be any uniform Lyapunov function and let C_V denote the closure of C_0^∞ under the norm $\sup_{x\in\mathbb{R}^d}|f(x)|/(1+V(x))$. Since $\mathcal{P}_t^\epsilon : C_V \to C_V$ is bounded (see the last line in the proof of Lemma 8.2), a straightforward density argument shows that $\mathcal{P}_t^\epsilon$ defines a C_0 semigroup on C_V. Theorem 6.5 and the spectral radius theorem for bounded operators then together imply that $\sigma(\mathcal{P}_t^\epsilon) \subset \{1\} \cup \left\{z \in \mathbb{C} : |z| \le e^{-\epsilon\delta t}\right\}$. Standard semigroup theory (e.g. [Theorem 3.6, [51]]) yields a corresponding estimate for the spectrum of the generator L_ϵ, in particular $\sigma(L_\epsilon) \subset \{0\} \cup \{z \in \mathbb{C} : \mathbf{Re}\, z \le -\epsilon\delta\}$.

By duality, Theorem 6.5 also implies a corresponding statement on the convergence of the law of x_t^ϵ as a measure on $\mathbb{R}^d$ to μ_ϵ in a weighted total variation space. For a continuous function $V : \mathbb{R}^d \to [0,\infty)$ and $\mu, \nu \in \mathcal{M}(\mathbb{R}^d)$ we write

$$\|\mu - \nu\|_{TV,V} = \sup_{\|f\|_V \le 1} \int_{\mathbb{R}^d} f(x)(\mu(dx) - \nu(dx)). \tag{6.17}$$

Corollary 6.6. *Under the assumptions and notations of Theorem 6.5, there exists $K, \delta >$ 0 that do not depend on ϵ such that for all $\epsilon \in (0,1)$, $t > 0$, and measures $\mu \in \mathcal{M}(\mathbb{R}^d)$*

satisfying $\int V(x)\mu(dx) < \infty$ *there holds*

$$\|(\mathcal{P}_t^\epsilon)^*\mu - \mu_\epsilon\|_{TV,V} \leq Ke^{-\epsilon\delta t}\|\mu - \mu_\epsilon\|_{TV,V}. \tag{6.18}$$

Many techniques exist for studying exponential convergence to equilibrium of a Markov process. Perhaps the most well-known and flexible methods are Harris type theorems, which combine drift towards a "small set" and a type of local irreducibility there to yield an explicitly computable rate of convergence in weighted total variation or Wasserstein distances; see e.g. [70, 73, 104]. Related criterion for subgeometric rates of convergence have also been studied [30, 44, 47]. For examples of works using a Harris theorem framework in the setting of (6.3) we refer to [126] and [72] mentioned above. In finite-dimensional situations, an entirely different class of techniques exist that use the Kolmogorov equation (i.e., PDE approaches) and functional inequalities involving the equilibrium density; see e.g. [8, 11, 121] and the references therein. Most directly, for elliptic generators with the form $L = \Delta + X \cdot \nabla$, a Poincaré inequality in L^2_μ (μ being the stationary measure) implies exponential convergence to equilibrium in the same space (for related results and discussion see e.g. [11]). This is a consequence of the a priori estimate

$$\frac{d}{dt}\|\mathcal{P}_t f - \mu(f)\|^2_{L^2_\mu} = -2\|\nabla(\mathcal{P}_t f - \mu(f))\|^2_{L^2_\mu}, \tag{6.19}$$

where $\mathcal{P}_t$ denotes the Markov semigroup generated by L. Poincaré inequalities also play a crucial role in degenerate settings; see for example [Theorem 35, [121]], which shows that exponential convergence to equilibrium for the kinetic Fokker-Planck equation with a C^2 confining potential $V : \mathbb{R}^d \to \mathbb{R}$ (satisfying a natural upper bound) is implied by an L^2 Poincaré inequality for $e^{-V(x)}dx$. Methods for proving convergence to equilibrium based on weaker functional inequalities are also known. For example, weak Poincaré inequalities, which trace back to [91] and were extended to a more general form in [116] (see also [66, 77] for applications in degenerate settings). The key feature of weak Poincaré inequalities is that they allow for a small loss of a norm stronger than L^2_μ on the right-hand side, the most common example being

$$\|f - \mu(f)\|_{L^2_\mu} \leq \beta(s)\|\nabla f\|_{L^2_\mu} + s\|f\|_\infty, \tag{6.20}$$

where $\|f\|_\infty := \sup_{x\in\mathbb{R}^d} |f(x)|$ and the inequality is required to hold for every $s > 0$ and some nonincreasing function $\beta : (0,\infty) \to [1,\infty)$ that possibly blows up as $s \to 0$ (see [Theorem 1.4, [11]] for a related but more general inequality). As such, they are much more forgiving to prove than standard Poincaré inequalities, but when applied in (6.19) only result in a subgeometric rate of convergence and from a stronger norm to a weaker norm.

The only uniform-in-ϵ information on f_ϵ that we are able to prove are the pointwise bounds stated in Theorem 6.2, which are far from enough to imply a Poincaré inequality (see e.g. [10, 11, 121] for common conditions on a measure that yield a Poincaré inequality). Moreover, as discussed in earlier, uniform-in-ϵ irreducibility statements are not forthcoming from standard methods. As such, it is not clear what the starting point for a proof of (6.16) should be. Our idea is to extend to the hypoelliptic setting the interesting fact that any measure $\mu(dx) = Ce^{-V(x)}dx$ for $V : \mathbb{R}^d \to \mathbb{R}$ that is merely locally bounded satisfies an (elliptic) *weak* Poincaré inequality. This is done in Lemmas 11.1 and 8.4, where we prove a hypoelliptic version of (6.20) that implies the decay estimate

$$\|\mathcal{P}_t^\epsilon f - \mu_\epsilon(f)\|_{L^2_{\mu_\epsilon}} \leq \psi(\epsilon t)\|f - \mu_\epsilon(f)\|_{L^\infty} \tag{6.21}$$

for some function $\psi : [0, \infty) \to [0, \infty)$ with $\lim_{t\to\infty} \psi(t) = 0$ and every bounded, Borel measurable function $f : \mathbb{R}^d \to \mathbb{R}$. Combining with the hypoelliptic regularization of Lemma 6.3 the local equivalence of $\|\cdot\|_{L^2_{\mu_\epsilon}}$ and $\|\cdot\|_{L^2}$ given by Theorem 6.2, we are able to upgrade (6.21) to exponential decay in $\|\cdot\|_V$ using a classical Harris framework. This argument is described in detail in Chapter 8. To our knowledge, this particular scheme for obtaining exponential convergence by combining a weak Poincaré inequality with a local regularization estimate and drift condition has not appeared in the literature. We believe that this approach is of general interest and could be useful in other related problems.

The fact that Theorem 6.5 is optimal with respect to the scaling of $\gamma_\epsilon, C_\epsilon$ is described in the following.

Theorem 6.7 (Optimal $\epsilon \to 0$ scaling of Theorem 6.5)**.** *Let γ be small enough so that $V = e^{\gamma x^2}$ is a uniform Lyapunov function and suppose that Assumptions 1-3 are satisfied. If there exists $s \geq 0$ and $K, \delta > 0$ such that for every $\epsilon \in (0, 1)$, measurable function $f : \mathbb{R}^d \to \mathbb{R}$ satisfying $\|f\|_V < \infty$, and $t > 0$ there holds*

$$\|\mathcal{P}_t^\epsilon f - \mu_\epsilon(f)\| \leq Ke^{-\epsilon^s \delta t}\|f - \mu_\epsilon(f)\|,$$

then $s \geq 1$. *Similarly, if* $\delta > 0$ *and* $\{K_\epsilon\}_{\epsilon\in(0,1)}$ *are such that*

$$\|\mathcal{P}_t^\epsilon f - \mu_\epsilon(f)\|_V \leq K_\epsilon e^{-\epsilon\delta t}\|f - \mu_\epsilon(f)\|_V$$

for every $t > 0$ *and measurable* $f : \mathbb{R}^d \to \mathbb{R}$ *with* $\|f\|_V < \infty$, *then* $\liminf_{\epsilon\to 0} K_\epsilon > 0$.

Remark 20. The fact that Theorem 6.5 is optimal with respect to scaling in ϵ is deeply tied to the fact that we are working in finite dimensions. In general, fixing dimension and sending $\epsilon \to 0$ can yield very different results from sending $\epsilon \to 0$ in infinite dimensional problems due to possibility of anomalous dissipation. In fact, it is conceivable that for some infinite dimensional problems γ_ϵ could be larger than possible in finite dimensions.

6.2 Examples

There are a variety of systems that fit within the general form (6.3) and satisfy Assumptions 1–3. A relatively simple example is the *Lorenz-96* model for N real-valued oscillators $u_1, ..., u_N$ in a periodic ensemble (i.e., $u_{j\pm N} = u_j$). After rescaling to match (6.3), it reads

$$du_m = (u_{m+1} - u_{m-2})u_{m-1}dt - \epsilon u_m dt + \sqrt{2\epsilon}q_m dW_t^{(m)}, \tag{6.22}$$

where $\left\{W_t^{(m)}\right\}$ are independent Brownian motions and $\{q_m\}$ are fixed parameters. This system of equations was first put forward by Lorenz in [96] as a toy model for the advection of a scalar atmospheric quantity such as temperature, and has since been studied as a prototypical chaotic, high dimensional system (see e.g. [82, 97, 101]). The structural

Assumption 1 is easy to check. For the uniform spanning condition, Assumption 2, let

$$X_j = \partial_{u_j} \quad 1 \le j \le N,$$
$$X_{0,\epsilon} = \sum_{m=1}^{N} \left[(u_{m+1} - u_{m-2})u_{m-1} - \epsilon u_m\right] \partial_{u_m},$$

and observe that some straightforward computations yield

$$[X_j, X_{0,\epsilon}] = -\epsilon X_j + (u_{j+2} - u_{j-1})X_{j+1} + u_{j-2}X_{j-1} - u_{j+1}X_{j+2}, \tag{6.23}$$

$$[X_{j+1}, [X_j, X_{0,\epsilon}]] = -X_{j+2} \tag{6.24}$$

for every $1 \le j \le N$. By iterating the bracket computations above and noting that the right-hand side of (6.24) does not depend on ϵ, it follows immediately that Assumption 2 is satisfied provided that $q_1, q_2 \neq 0$. Assumption 3 follows from the bracket computation above and Theorem 2.5 in [76]. The key structure at play here is that while the nonlinear part of $X_{0,\epsilon}$ is even (which is generally a hindrance to proving controllability), it does not contain any diagonal terms; see e.g. discussions in [61, 76].

Next, Galerkin truncations (of arbitrary dimension) of the 2D Navier-Stokes equations in vorticity form set on a periodic box can be written in the form (6.3) with Assumption 1. The bracket structure is far more complicated than the simple example of Lorenz-96 above. Conditions on the forcing for Assumption 2 to hold were obtained in [126]. Assumption 3 again follows from [76] and the bracket computations needed to prove Assumption 2; see also [Example 3.8, [76]].

As a final example we have the SABRA shell model for turbulence, first introduced in [99]. Truncated to finite dimensional $(u_1, \dots, u_J) \in \mathbb{C}^J$ with boundary conditions $u_{-1} = u_0 = u_{J+1} = u_{J+2} = 0$ the model reads

$$\begin{aligned} \partial_t u_m = i2^m \left(\overline{u_{m+1}} u_{m+2} - \frac{\epsilon'}{2} \overline{u_{m-1}} u_{m+1} - \frac{\epsilon' - 1}{4} u_{m-2} u_{m-1} \right) \\ - \epsilon 2^{2m} u_m + \sqrt{\epsilon} q_m dW_t^{(m;R)} + i\sqrt{\epsilon} p_m dW_t^{(m;I)}, \end{aligned} \tag{6.25}$$

where q_m, p_m are real parameters and $\epsilon' \in (0,2) \setminus \{1\}$. For $\epsilon' \in (0,1)$ the system has only one positive invariant (the energy) and is meant to model 3D turbulence, while for $\epsilon' \in (1,2)$ the system has two positive invariants and is designed to capture properties of 2D turbulence. More discussion on (6.25) and shell models in general can be found in [41]. Assumption 1 is straightforward to verify. To determine simple conditions ensuring the bracket condition is satisfied, we rewrite the system in real variables $u_m = a_m + ib_m$. In this new formulation we have

$$\begin{aligned} Z_0 = &\sum_{m=1}^{J} 2^m \big[(a_{m+2} b_{m+1} - a_{m+1} b_{m+2}) \partial_{a_m} + (a_{m+1} a_{m+2} + b_{m+1} b_{m+2}) \partial_{b_m} \big] \\ &+ \sum_{m=1}^{J} \epsilon' 2^{m-1} \big[(a_{m-1} b_{m+1} - a_{m+1} b_{m-1}) \partial_{a_m} - (a_{m-1} a_{m+1} + b_{m-1} b_{m+1}) \partial_{b_m} \big] \\ &+ \sum_{m=1}^{J} (\epsilon' - 1) 2^{m-2} \big[(a_{m-2} b_{m-1} + a_{m-1} b_{m-2}) \partial_{a_m} - (a_{m-2} a_{m-1} - b_{m-2} b_{m-1}) \partial_{b_m} \big] \end{aligned}$$

Despite the appearance, Assumption 2 is actually essentially as easy to verify as in the case of Lorenz-96. Some computations reveal that

$$[\partial_{a_{j+1}}, [\partial_{a_j}, Z_0]] = 2^{j-1}\mathbf{1}_{2\le j\le J-1}\partial_{b_{j-1}} - (\epsilon' - 1)2^j\mathbf{1}_{1\le j\le J-2}\partial_{b_{j+2}},$$

$$[\partial_{b_{j+1}}, [\partial_{a_j}, Z_0]] = -2^{j-1}\mathbf{1}_{2\le j\le J-1}\partial_{a_{j-1}} + (\epsilon' - 1)2^j\mathbf{1}_{1\le j\le J-2}\partial_{a_{j+2}},$$

$$[\partial_{a_{j+1}}, [\partial_{b_j}, Z_0]] = 2^{j-1}\mathbf{1}_{2\le j\le J-1}\partial_{a_{j-1}} + (\epsilon' - 1)2^j\mathbf{1}_{1\le j\le J-2}\partial_{a_{j+2}},$$

$$[\partial_{b_{j+1}}, [\partial_{b_j}, Z_0]] = 2^{j-1}\mathbf{1}_{2\le j\le J-1}\partial_{b_{j-1}} + (\epsilon' - 1)2^j\mathbf{1}_{1\le j\le J-2}\partial_{b_{j+2}}.$$

Using these formulas it is easy to deduce, in the notation from Definition 6.1, the following facts:

- if $\partial_{a_1}, \partial_{a_2}, \partial_{b_1}, \partial_{b_2} \in \cup_{n=0}^{\infty} V_n$, then $\partial_{a_3}, \partial_{b_3} \in \cup_{n=0}^{\infty} V_n$;
- if $\partial_{a_{j-1}}, \partial_{a_j}, \partial_{a_{j+1}}, \partial_{b_{j-1}}\partial_{b_j}, \partial_{b_{j+1}} \in \cup_{n=0}^{\infty} V_n$ for some $2 \le j \le J-2$, then $\partial_{a_{j+2}}, \partial_{b_{j+2}} \in \cup_{n=0}^{\infty} V_n$.

From the two bullets above it is clear that Assumption 2 is satisfied provided that q_1, q_2, p_1, p_2 are all non-zero. As in the case of Lorenz-96, Assumption 3 follows from [76].

Chapter 7: Preliminaries

In this preliminary Chapter, we give precise statements and proof sketches of some standard results concerning (6.3), namely global well-posedness, the regularizing properties of the semigroup $\mathcal{P}_t^\epsilon$, and the existence and uniqueness of μ_ϵ. The results of this chapter are purely qualitative in nature.

7.1 Preliminary facts

We begin with the global well-posedness of (6.3) and some basic moment bounds. In what follows, $\left\{W_t^{(j)}\right\}_{j=1}^r$ are independent one-dimensional Wiener processes on a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and $\mathcal{F}_t$ denotes the σ-algebra generated by $\{W_s^{(j)} : 1 \leq j \leq r, 0 \leq s \leq t\}$ and the $\mathbf{P}$-null sets of $\mathcal{F}$. Moreover, we write $\mathbf{E}$ to denote expectation with respect to the measure $\mathbf{P}$.

Lemma 7.1. *Suppose that Assumption 1 holds (Assumption 2 is not needed). Let $X_0 \in L^2(\Omega; \mathbf{P})$ be a random variable independent of the σ-algebra generated by $\cup_{t\geq 0}\mathcal{F}_t$, and let $\mathcal{F}_t^{X_0}$ denote the σ-algebra generated by $\mathcal{F}_t$ and X_0. For $\epsilon \in (0,1)$, consider the SDE*

$$\begin{cases} dX_t^\epsilon = -\epsilon A X_t^\epsilon dt - \epsilon^\alpha B X_t^\epsilon dt - N(X_t^\epsilon)dt + \sqrt{2\epsilon}\sum_{j=1}^r Z_j dW_t^{(j)}, \\ X_0^\epsilon = X_0. \end{cases} \tag{7.1}$$

There exists a unique (up to indistinguishability), globally defined $\mathcal{F}_t^{X_0}$-adapted process $(X_t^\epsilon)_{t\geq 0}$ with continuous sample paths solving the integral form of (7.1) $\mathbf{P}$-a.s. and such that $\int_0^T \mathbf{E}|X_t^\epsilon|^2 dt < \infty$ for every $T \geq 0$. Let $X_{t,x}^\epsilon$ denote the unique solution with initial condition $X_0 = x \in \mathbb{R}^d$. If V is any uniform Lyapunov function (see Definition 6.4) with $\kappa, b > 0$ as in (6.15), then uniformly in $\epsilon \in (0,1)$ and $x \in \mathbb{R}^d$ there holds

$$\mathbf{E}V(X_{t,x}^\epsilon) \leq e^{-\epsilon\kappa t}V(x) + \frac{b}{\kappa}. \tag{7.2}$$

Proof. Since the noise is additive and the drift is smooth, uniqueness follows from the usual ODE argument using Grönwall's lemma. Due to the energy conservation property $N(x)\cdot x = 0$, global existence can be proven with an approximation scheme that relies on standard energy estimates and a routine stopping time argument. The details needed to carry out the procedure can all be found in [110] and [Section 3, [52]].

To prove the moment bound (7.2) we begin by applying Itô's formula to obtain

$$\begin{aligned} e^{\kappa\epsilon t}V(X_{t,x}^\epsilon) &= V(x) + \int_0^t e^{\kappa\epsilon s}(L_\epsilon V(X_{s,x}^\epsilon) + \epsilon\kappa V(X_{s,x}^\epsilon))ds \\ &\quad + \sqrt{2\epsilon}\sum_{j=1}^r \sum_{k=1}^d \int_0^t \frac{\partial V}{\partial x_k}(X_{s,x}^\epsilon) Z_j^{(k)} dW_s^{(j)}. \end{aligned} \tag{7.3}$$

Let $\tau_n(\omega) = \inf\{s \in [0,t] : |X_{s,x}^\epsilon| = n\}$. Applying (6.15) to estimate the first line of (7.3) and then localizing with τ_n (so that the stochastic integral becomes a martingale) we obtain, uniformly in $n \in \mathbb{N}$,

$$\mathbf{E}V(X_{t\wedge\tau_n,x}^\epsilon) \leq V(x)\mathbf{E}e^{-\kappa\epsilon t\wedge\tau_n} + \frac{\beta}{\kappa}. \tag{7.4}$$

Sending $n \to \infty$ the desired result follows from Fatou's lemma and the fact that $\tau_n \uparrow t$ $\mathbf{P}$-a.s. □

Next, we have a statement on the continuity of $X^\epsilon_{t,x}$ with respect to the initial condition and noise trajectory. In what follows, we write $W(t) \in \mathbb{R}^r$ for $(W^{(1)}_t, \dots, W^{(r)}_t)$.

Lemma 7.2. *Let $X^\epsilon_{t,x}$ be as in Lemma 7.1. If $x_n \to x$ in $\mathbb{R}^d$ then X^ϵ_{t,x_n} converges to $X^\epsilon_{t,x}$ $\mathbf{P}$-a.s. uniformly on compact time intervals. Moreover, the solution is continuous with respect to the Wiener trajectory in the sense that there exists a set $\Omega' \subseteq \Omega$ with full measure so that for every fixed $0 < T < \infty$ and $\omega_1, \omega_2 \in \Omega'$ one has that*

$$\sup_{0\le t\le T} |X^\epsilon_{t,x}(\omega_1) - X^\epsilon_{t,x}(\omega_2)| \to 0 \quad \textit{as} \quad \sup_{0\le t\le T} |W_t(\omega_1) - W_t(\omega_2)| \to 0. \tag{7.5}$$

Proof. We will just prove continuity with respect to the Wiener trajectory, as the statement concerning continuity with respect to the initial condition follows from a similar argument. For notational convenience we define $Z_0(x) = -\epsilon Ax - \epsilon^\alpha Bx - N(x)$. Fix $T > 0$ and $x \in \mathbb{R}^d$. For $j = 1, 2$ let $F_j : [0, T] \to \mathbb{R}^d$ be continuous and suppose that $x_j : [0, T] \to \mathbb{R}^d$ is a continuous solution to the integral equation

$$x_j(t) = x + \int_0^t Z_0(x_j(s))ds + F_j(t), \quad j = 1, 2.$$

Since we consider additive noise, to prove (7.5) it is enough to show that

$$\lim_{\delta\to 0} \sup_{F_2:\|F_1-F_2\|_{C([0,T];\mathbb{R}^d)}\le\delta} \sup_{0\le t\le T} |x_1(t) - x_2(t)| = 0. \tag{7.6}$$

Since x_1 is continuous, there exists $C > |x| + 1$ so that $\sup_{0 \leq t \leq T} |x_1(t)| \leq C$. For $\epsilon' > 0$ fixed and F_2 to be chosen close to F_1, let T_* be the maximal time so that $|x_1(t) - x_2(t)| \leq \epsilon'$ for all $t \in [0, T_*]$. By continuity we have $T_* > 0$, and moreover by a simple Grönwall argument there holds

$$\sup_{0 \leq t \leq T_*} |x_1(t) - x_2(t)| \lesssim_{C,T} \|F_1 - F_2\|_{C([0,T];\mathbb{R}^d)}. \tag{7.7}$$

Hence, as long as $\|F_1 - F_2\|_{C([0,T];\mathbb{R}^d)}$ is small in terms of T, C, and ϵ', it follows from a bootstrap argument that $T_* = T$. This yields (7.6). □

7.2 Qualitative regularity and Existence & Uniqueness of μ_ϵ

Recall that we write $\mathcal{M}(\mathbb{R}^d)$ for the space of Borel probability measures on $\mathbb{R}^d$ and $\mathcal{B}(\mathbb{R}^d)$ for the set of Borel sets. Also, we will denote the space of bounded, Borel measurable function $f : \mathbb{R}^d \to \mathbb{R}$ by $B_b(\mathbb{R}^d)$. In the setting of Lemma 7.1 the unique, global solution $X_{t,x}$ is a Markov process with respect to the filtration $\mathcal{F}_t$; i.e., for every $t, s > 0$, $x \in \mathbb{R}^d$, and $f \in B_b(\mathbb{R}^d)$ there holds

$$\mathbf{E}\left(f(X_{t+s,x}) | \mathcal{F}_s\right)(\omega) = \mathbf{E}\left(f\left(X_{t,X_s(\omega)}\right)\right). \tag{7.8}$$

Moreover, $(x, \omega) \mapsto X_{t,x}(\omega)$ is measurable for fixed $t \geq 0$. Together with (7.8), this allows one to prove that the transition probabilities $\mathcal{P}_t^\epsilon(x, A) = \mathbf{P}(X_{t,x}^\epsilon \in A)$ define a *Markov transition kernel*. That is, the following properties are satisfied:

- $\mathcal{P}_t^\epsilon(x, \cdot)$ is a probability measure on $\mathbb{R}^d$ for each $x \in \mathbb{R}^d$;

- $x \mapsto \mathcal{P}_t^\epsilon(x, A)$ is Borel measurable for each $A \in \mathcal{B}(\mathbb{R}^d)$;

- for every $t, s \geq 0$ and $A \in \mathcal{B}(\mathbb{R}^d)$ the semigroup property

$$\mathcal{P}_{t+s}^\epsilon(x, A) = \int_{\mathbb{R}^d} \mathcal{P}_t^\epsilon(y, A)\mathcal{P}_s^\epsilon(x, dy)$$

  holds.

The associated Markov semigroup $\mathcal{P}_t^\epsilon : B_b(\mathbb{R}^d) \to B_b(\mathbb{R}^d)$ is then defined by $\mathcal{P}_t^\epsilon f(x) = \mathbf{E}f(X_{t,x}^\epsilon)$.

The next lemma is about the regularizing properties of $\mathcal{P}_t^\epsilon$ and the uniqueness of its invariant measure.

Lemma 7.3. *Suppose that Assumptions 1 and 2 both hold. Then, the Markov semigroup $\mathcal{P}_t^\epsilon : B_b(\mathbb{R}^d) \to B_b(\mathbb{R}^d)$ is smoothing in the sense that if $f \in B_b(\mathbb{R}^d)$, then $\mathcal{P}_t^\epsilon f$ is smooth in space for each $t > 0$. Similarly, $(\mathcal{P}_t^\epsilon)^*\mu$ has a smooth density with respect to Lebesgue measure for any $\mu \in \mathcal{M}(\mathbb{R}^d)$ and $t > 0$. Moreover, $\mathcal{P}_t^\epsilon$ admits a unique invariant measure μ_ϵ and it has a smooth density f_ϵ satisfying*

$$\sup_{\epsilon \in (0,1)} \int_{\mathbb{R}^d} V(x) f_\epsilon dx \lesssim 1 \tag{7.9}$$

for any uniform Lyapunov function V.

Proof. For $\mu \in \mathcal{M}(\mathbb{R}^d)$, the measure $\mu_t = (\mathcal{P}_t^\epsilon)^*\mu$ is a distributional solution to Kolmogorov forward equation

$$\partial_t \mu_t = L_\epsilon^* \mu_t. \tag{7.10}$$

Since $\{\epsilon Ax + \epsilon^\alpha Bx + N, Z_1, \dots, Z_r\}$ satisfies the parabolic Hörmander condition, it is easy to see that $\partial_t - L_\epsilon^*$ satisfies Hörmander's condition on $\mathbb{R}^{d+1}$. Hence, the fact that μ_t has a smooth density with respect to Lebesgue measure for $t > 0$ is a direct consequence of Hörmander's theorem [79]. Similarly, it is classical consequence of Itô's formula that if $f \in C(\mathbb{R}^d)$, then $\mathcal{P}_t^\epsilon f$ is a distributional solution to the backward equation

$$\partial_t \mathcal{P}_t^\epsilon f = L_\epsilon \mathcal{P}_t^\epsilon f. \tag{7.11}$$

Using that $X_{t,x}^\epsilon$ has a smooth density with respect to Lebesgue measure for fixed x, one can show with a standard approximation argument that (7.11) holds when f is just bounded and measurable. The regularity of $\mathcal{P}_t^\epsilon f$ for $f \in B_b(\mathbb{R}^d)$ then follows again by Hörmander's theorem.

Existence of an invariant measure follows from the moment bound (7.2) and the Krylov-Boguliubov theorem (see e.g. [Theorem 3.1.1, [37]]). Since $\mathcal{P}_t^\epsilon$ is strong Feller (meaning that $\mathcal{P}_t^\epsilon f$ is continuous for $t > 0$ as soon as f is just bounded and measurable), to prove uniqueness it suffices to show that any invariant measure contains the origin in its support. This is a standard consequence of the dissipative structure of (7.1), the continuity with respect to the Wiener trajectory proven in Lemma 7.1, and the fact that $\mathbf{P}(\sup_{0\le t\le T} |W_t| \le \epsilon') > 0$ for any $\epsilon', T > 0$. Lastly, the moment bound (7.9) is proven by approximating V with $\min(V, n)$ for $n \in \mathbb{N}$, iteratively applying (7.2), and then sending $n \to \infty$. □

Remark 21. To more easily justify certain computations, it will be convenient at times to work with a regularized process $X_t^{\epsilon,\delta}$ generated by $L_\epsilon + \epsilon\delta\Delta$ for $\delta \in (0, 1]$. The results of Lemmas 7.1-7.3 hold equally well for the regularized process and moreover are uniform-in-δ. In particular, the associated Markov semigroup $\mathcal{P}_t^{\epsilon,\delta}$ admits a unique invariant measure $\mu_{\epsilon,\delta}$ with smooth density $f_{\epsilon,\delta}$ solving the problem

$$\begin{cases} L_\epsilon^* f_{\epsilon,\delta} + \epsilon\delta\Delta f_{\epsilon,\delta} = 0 \\ f_{\epsilon,\delta} \geq 0 \\ \int f_{\epsilon,\delta} = 1 \end{cases} \tag{7.12}$$

and satisfying

$$\sup_{\epsilon\in(0,1),\delta\in[0,1]} \int_{\mathbb{R}^d} V(x) f_{\epsilon,\delta}(x)dx < \infty \tag{7.13}$$

for any uniform Lyapunov function V. The process $X_{t,x}^{\epsilon,\delta}$ approximates $X_{t,x}^{\epsilon}$ in that if $\{\delta_n\}_{n=1}^{\infty}$ is a sequence with $\delta_n \to 0$, then

$$\lim_{n\to\infty} \sup_{0\leq t\leq T} |X_{t,x}^{\epsilon,\delta_n}(\omega) - X_{t,x}^{\epsilon}(\omega)| = 0 \quad \mathbf{P}\text{-a.s.}, \tag{7.14}$$

which can be proven with essentially the same technique used in Lemma 7.2. The argument that $f_{\epsilon,\delta}$ approximates f_ϵ in a suitable sense is a bit more subtle and carried out in Lemma A.2 of the Appendix.

Remark 22. As a consequence of Remark 21 and the uniqueness described in Lemma 7.3, for any $\epsilon \in (0,1)$ and $\delta \in [0,1]$ the only probability measure μ solving

$$(L_\epsilon^* + \epsilon\delta\Delta)\mu = 0$$

in the sense of distributions is $\mu_{\epsilon,\delta}$.

Remark 23. A straightforward computation using the fact that $N(x) \cdot x = Bx \cdot x = 0$ shows that $V_\delta(x) = e^{\gamma x^2}$ is a uniform Lyapunov function for any $\gamma \ll 1$. As a special case of (7.13) we thus have

$$\sup_{\epsilon \in (0,1), \delta \in [0,1]} \int_{\mathbb{R}^d} e^{\gamma x^2} f_{\delta,\epsilon}(x) dx < \infty. \tag{7.15}$$

Chapter 8: Proof outline of Theorem 6.5

In this chapter we prove our main quantitative geometric ergodicity result, Theorem 6.5, assuming the uniform bounds on the stationary measure provided by Theorem 6.2 and stating various intermediate results as lemmas to be proven in Chapter 11. As discussed earlier in Section 6.1.3, the main idea is to use the pointwise bounds on f_ϵ to prove a type of hypoelliptic weak Poincaré, which is combined with $L^2 \to L^\infty$ regularization for $\mathcal{P}_t^\epsilon$ and a Harris theorem to deduce exponential convergence to equilibrium. Our scheme in essence allows us to replace the irreducibility condition required by standard Harris theorems, which is typically phrased as a condition on transition probabilities (see e.g. [Assumption 2, [73]]), with the pointwise bounds of Theorem 6.2. Thus, our general approach to proving Theorem 6.5 may be useful for estimating the convergence rate to equilibrium in other problems where quantitative lower bounds on transition probabilities are not readily available, but good pointwise bounds on the stationary density can be obtained.

8.1 Harris framework

In this section, we set up the Harris framework that we use to quantify the spectral gap for $\mathcal{P}_t^\epsilon$. For an arbitrary Polish space $\mathcal{X}$ (i.e., a complete, separable metric space) we use the same notations $B_b(\mathcal{X})$, $\mathcal{M}(\mathcal{X})$, and $\mathcal{B}(\mathcal{X})$ defined earlier in the case that $\mathcal{X} = \mathbb{R}^d$. Recall also the notations $\|\cdot\|_V$ and $\|\cdot\|_{TV,V}$ defined in (6.10) and (6.17), respectively. These extend in the obvious way to general Polish spaces $\mathcal{X}$. Lastly, for any bounded, Borel measurable function $f : \mathcal{X} \to \mathbb{R}$ we write

$$\|f\|_\infty := \sup_{x \in \mathcal{X}} |f(x)|. \tag{8.1}$$

The Harris theorem we will use is straightforward to prove after a careful reading of [73] and is stated as follows.

Theorem 8.1 (Harris)**.** *Let $\mathcal{P}$ be a Markov semigroup on a Polish space $\mathcal{X}$. Suppose that there exists a continuous function $V : \mathcal{X} \to [0, \infty)$ such that:*

- *There exists $\bar{b} \geq 1$ such that for all $x \in \mathcal{X}$ there holds*

$$\mathcal{P}V(x) \leq \frac{1}{2}V(x) + \bar{b}. \tag{8.2}$$

- *There exists $\eta \in (0, 2)$ such that if $x, y \in \mathcal{X}$ satisfy $V(x) + V(y) \leq 10\bar{b}$ and $f : \mathcal{X} \to \mathbb{R}$ is a bounded, Borel measurable function with $\|f\|_\infty \leq 1$, then*

$$|\mathcal{P}f(x) - \mathcal{P}f(y)| \leq 2 - \eta. \tag{8.3}$$

Then, there exists $c \in (0, 1)$ satisfying $c \gtrsim \eta$ and $K \geq 1$ depending on $\bar{b}$ and η such that for any two measures $\mu, \nu \in \mathcal{M}(\mathcal{X})$ with $\int_{\mathcal{X}} V(x)\mu(dx) + \int_{\mathcal{X}} V(x)\nu(dx) < \infty$ and any

$n \in \mathbb{N}$ *there holds*

$$\|(\mathcal{P}^*)^n(\mu - \nu)\|_{TV,V} \le K(1-c)^n \|\mu - \nu\|_{TV,V}. \tag{8.4}$$

As a consequence, $\mathcal{P}$ *can admit only one stationary measure, and if* μ_∞ *is the unique stationary measure there exists* $\bar{K} \ge 1$ *depending on* $\bar{b}$ *and* η *so that for any* $n \in \mathbb{N}$ *and measurable function* $f : \mathcal{X} \to \mathbb{R}$ *with* $\|f\|_V < \infty$ *there holds*

$$\|\mathcal{P}^n f - \mu_\infty(f)\|_V \le \bar{K}(1-c)^n \|f - \mu_\infty(f)\|_V. \tag{8.5}$$

Proof. As in [73], for we consider the family of norms

$$\|\varphi\|_{\beta V} = \sup_{x \in \mathcal{X}} \frac{|\varphi(x)|}{1 + \beta V(x)}, \quad \beta > 0 \tag{8.6}$$

defined for measurable functions $\varphi : \mathcal{X} \to \mathbb{R}$. A careful reading of the statement and proof of Theorem 3.1 in [73] shows the following. Suppose that $\alpha, \gamma \in (0,1)$ and $K > 0$ are positive constants such that:

- For every $x \in \mathcal{X}$, $\mathcal{P}V(x) \le \gamma V(x) + K$.
- For every $\beta > 0$, $x, y \in \mathcal{X}$ with $V(x) + V(y) \le 4K/(1-\gamma)$, and measurable $\varphi : \mathcal{X} \to \mathbb{R}$ with $\|\varphi\|_\beta \le 1$, there holds

$$|\mathcal{P}\varphi(x) - \mathcal{P}\varphi(y)| \le 2(1-\alpha) + \gamma\beta V(x) + \gamma\beta V(y) + 2\beta K. \tag{8.7}$$

Then, for any two measures $\mu, \nu \in \mathcal{M}(\mathcal{X})$ with

$$\int_{\mathcal{X}} V(x)\mu(dx), \quad \int_{\mathcal{X}} V(x)\nu(dx) < \infty$$

there holds

$$\|\mathcal{P}(\mu - \nu)\|_{TV,(\alpha/2K)V} \leq \bar{\alpha}\|\mu - \nu\|_{TV,(\alpha/2K)V}, \tag{8.8}$$

where

$$\bar{\alpha} = \max\left(1 - \frac{\alpha}{2}, 1 - \frac{\alpha(1-\gamma)}{2\alpha + 2(1-\gamma)}\right). \tag{8.9}$$

We will now use our assumptions to verify the two bullets above for particular choices of the constants α, γ, and K. Fix $\gamma = 1/2$ and $K = \bar{b}$. Then, by our assumption in (8.2), the first bullet above holds. To find $\alpha \in (0, 1)$ such that (8.7) holds, we follow [67]. Fix $\beta > 0$ and $\varphi : \mathcal{X} \to \mathbb{R}$ with $\|\varphi\|_{\beta V} \leq 1$. Since $|\varphi(x)| \leq 1 + \beta V(x)$, we can write $\varphi = \varphi_1 + \varphi_2$ for measurable functions φ_1 and φ_2 with $|\varphi_1(x)| \leq 1$ and $|\varphi_2(x)| \leq \beta V(x)$ for every $x \in \mathcal{X}$. Let $x, y \in \mathcal{X}$ be such that $V(x) + V(y) \leq 4K/(1-\gamma)$. Since $4K/(1-\gamma) = 8\bar{b} \leq 10\bar{b}$, we may use the assumption (8.3) to obtain

$$\begin{aligned}
|\mathcal{P}\varphi(x) - \mathcal{P}\varphi(y)| &\leq |\mathcal{P}\varphi_1(x) - \mathcal{P}\varphi_1(y)| + |\mathcal{P}\varphi_2(x) - \mathcal{P}\varphi_2(y)| \\
&\leq 2 - \eta + |\mathcal{P}\varphi_2(x)| + |\mathcal{P}\varphi_2(y)| \\
&\leq 2 - \eta + \beta\gamma V(x) + \beta\gamma V(y) + 2\beta K \\
&= 2\left(1 - \frac{\eta}{2}\right) + \beta\gamma V(x) + \beta\gamma V(y) + 2\beta K.
\end{aligned}$$

Thus, (8.7) holds with $\alpha = \eta/2$. We conclude that for any two measures $\mu, \nu \in \mathcal{M}(\mathcal{X})$

with $\int_{\mathcal{X}} V(x)\mu(dx), \int_{\mathcal{X}} V(x)\nu(dx) < \infty$ there holds

$$\|\mathcal{P}^n(\mu - \nu)\|_{TV,(\eta/4\bar{b})V} \le \bar{\alpha}^n \|\mu - \nu\|_{TV,(\eta/4\bar{b})V}, \tag{8.10}$$

where

$$\bar{\alpha} = \max\left(1 - \frac{\eta}{4}, 1 - \frac{\eta}{4(\eta + 1)}\right) \le 1 - \frac{\eta}{12}. \tag{8.11}$$

Since $\eta/4\bar{b} \le 1$ we have

$$\frac{\eta}{4\bar{b}}\|\mathcal{P}^n(\mu - \nu)\|_{TV,V} \le \|\mathcal{P}^n(\mu - \nu)\|_{TV,(\eta/4\bar{b})V}$$

and $\|\mu - \nu\|_{TV,(\eta/4\bar{b})V} \le \|\mu - \nu\|_{TV,V}$ and hence (8.11) implies

$$\|\mathcal{P}^n(\mu - \nu)\|_{TV,V} \le \frac{4\bar{b}}{\eta}\left(1 - \frac{\eta}{12}\right)^n \|\mu - \nu\|_{TV}.$$

This proves (8.4). The final claim (8.5) then follows from (8.4) and an argument using the duality formula

$$\|\mu - \nu\|_{TV,V} = \sup_{\|f\|_V \le 1} \int_{\mathcal{X}} f(x)(\mu - \nu)(dx),$$

which holds for any μ and ν with $\int_{\mathcal{X}} V(x)\mu(dx), \int_{\mathcal{X}} V(x)\nu(dx) < \infty$. Here, we are using that $\int_{\mathcal{X}} V(x)\mu_\infty(dx)$ is finite for any invariant measure μ_∞ for $\mathcal{P}$. $\square$

8.2 Proof of quantitative geometric ergodicity

Using Theorem 8.1, we can reduce the proof of Theorem 6.5 to showing that for any uniform Lyapunov function V and fixed $\bar{b}$, $\mathcal{P} = \mathcal{P}_t^\epsilon$ satisfies (8.3) for $t \gg \epsilon^{-1}$ and $\eta > 0$ that does not depend on ϵ. This is the content of the following lemma.

Lemma 8.2. *Suppose that*

$$\lim_{t\to\infty} \sup_{\epsilon\in(0,1),\|f\|_\infty\le 1} \|\mathcal{P}^\epsilon_{t\epsilon^{-1}} f - \mu_\epsilon(f)\|_{L^\infty(B_R)} = 0. \tag{8.12}$$

for every fixed $R > 0$. Then, the result of Theorem 6.5 follows.

Proof. Let V be a uniform Lyapunov function for $\mathcal{P}_t^\epsilon$. By (7.2) we have

$$\mathcal{P}_t^\epsilon V(x) = e^{-\kappa\epsilon t} V(x) + \frac{b}{\kappa} := e^{-\kappa\epsilon t} V(x) + \bar{b} \tag{8.13}$$

for some $\bar{b}$ that does not depend on ϵ. Hence, if $C_* > 0$ is sufficiently large and $t_* = \epsilon^{-1} C_*$ then

$$\mathcal{P}^\epsilon_{t_*} V(x) \le \frac{1}{2} V(x) + \bar{b}. \tag{8.14}$$

Since, $\lim_{|x|\to\infty} V(x) = \infty$, there exists $R > 0$ such that $V(x) + V(y) \le 10\bar{b}$ implies that $x, y \in B_R$. By (8.12) and the triangle inequality, choosing C_* perhaps even larger we can ensure that

$$\sup_{\epsilon\in(0,1)} |\mathcal{P}^\epsilon_{t_*} f(x) - \mathcal{P}^\epsilon_{t_*} f(y)| \le 1 \tag{8.15}$$

for every $x, y \in \mathbb{R}^d$ satisfying $V(x) + V(y) \le 10\bar{b}$ and every measurable $f : \mathbb{R}^d \to \mathbb{R}$

with $\|f\|_\infty \leq 1$. By (8.14), (8.15), and Theorem 8.1 applied with $\mathcal{P} = \mathcal{P}^\epsilon_{t_*}$, there exists $K, \delta > 0$ such that for every $n \in \mathbb{N}$, $\epsilon \in (0,1)$, and measurable $f : \mathbb{R}^d \to \mathbb{R}$ with $\|f\|_V < \infty$ there holds

$$\|\mathcal{P}^\epsilon_{nt_*} f - \mu_\epsilon(f)\|_V \leq K e^{-\delta \epsilon n t_*} \|f - \mu_\epsilon(f)\|_V. \tag{8.16}$$

It only remains to upgrade (8.16) to continuous time. Let $t \geq 0$ and choose $n \in \mathbb{N}$ so that $t \in [nt_*, (n+1)t_*)$. Using the semigroup property, there exists $s \in [0, t_*)$ so that

$$\|\mathcal{P}^\epsilon_t f - \mu_\epsilon(f)\|_V = \|\mathcal{P}^\epsilon_s(\mathcal{P}^\epsilon_{nt_*} f - \mu_\epsilon(f))\|_V \leq K e^{-\delta \epsilon n t_*} \|\mathcal{P}^\epsilon_s\|_{V \to V} \|f - \mu_\epsilon(f)\|_V.$$

Whenever $n \geq 1$ one has $\delta n \epsilon t_* \geq \epsilon(\delta/2)t$, and so (6.16) follows provided

$$\sup_{0 \leq s \leq C_* \epsilon^{-1}} \|\mathcal{P}^\epsilon_s\|_{V \to V} \lesssim_{C_*} 1, \tag{8.17}$$

which is an immediate consequence of (8.13). Indeed, from the bound $\mathcal{P}^\epsilon_t V(x) \leq V(x) + \beta/\kappa$ we have

$$\|\mathcal{P}^\epsilon_t f\|_V = \sup_{x \in \mathbb{R}^d} \frac{|\mathcal{P}^\epsilon_t f|(x)}{1 + V(x)} \leq \|f\|_V \sup_{x \in \mathbb{R}^d} \frac{1 + \mathcal{P}^\epsilon_t V(x)}{1 + V(x)} \lesssim_{\kappa, b} \|f\|_V.$$

□

Remark 24. The limit (8.12) plays the role of a quantitative minorization condition, and is the main challenge in the proof of Theorem 6.5. As discussed earlier in Chapter 6, the standard proof that $\mathcal{P} = \mathcal{P}^\epsilon_t$ satisfies (8.3) for $t \gg \epsilon^{-1}$ yields an estimate on η, and consequently γ_ϵ, that is exponentially small in ϵ. To see roughly where the ϵ-dependence comes from, recall from the earlier discussion that the usual approach to obtaining a minorization condition uses the dissipative structure of (6.3) and the regularizing properties of the semigroup to deduce that distinct transition probabilities overlap near the origin for long times. This hinges on the fact that there is a nonzero probability that the driving Wiener process remains small over a fixed time window. It turns out (see e.g. [Lemma 3.1, [126]]) that the estimate on η that one can prove with this approach is limited by a lower bound for

$$\mathbf{P}\left(\sup_{0\leq s\leq \epsilon^{-1}}\left|\sum_{j=1}^{r} Z_j W_s^{(j)}\right| \ll \sqrt{\epsilon}\right),$$

which is exponentially small in ϵ.

Our proof of (8.12) is based on a two step procedure that makes crucial use of Theorem 6.2. First, we prove that $\mathcal{P}_t^\epsilon$ satisfies the quantitative $L^2_{\mu_\epsilon} \to L^\infty$ "parabolic" regularization estimate stated in Lemma 6.3. As mentioned earlier, this requires us to derive a slightly subtle space-time Hörmander inequality better adapted to a parabolic framework (Lemma 9.5). With such an inequality in hand, we will apply a suitably adapted Moser iteration, the details of which can be found in Section 11.2.

Given Lemma 6.3, the proof of (8.12) reduces to showing that

$$\lim_{t\to\infty} \sup_{\epsilon\in(0,1),\|f\|_\infty\leq 1} \|\mathcal{P}^\epsilon_{t\epsilon^{-1}} f - \mu_\epsilon(f)\|_{L^2_{\mu_\epsilon}} = 0. \tag{8.18}$$

In other words, we need to prove that $\mathcal{P}_t^\epsilon$ satisfies a $\|\cdot\|_\infty \to \|\cdot\|_{L^2_{\mu_\epsilon}}$ decay estimate with timescale ϵ^{-1} on functions which are mean zero with respect to μ_ϵ. This task is much more tractable than proving, say, an $L^2_{\mu_\epsilon} \to L^2_{\mu_\epsilon}$ exponential decay estimate for $\mathcal{P}^\epsilon_t$ (which would require a Poincaré inequality for μ_ϵ; see e.g. [11]), and is provided by the following lemma.

Lemma 8.3. *Under Assumption 3, there exists a function $\psi : [0, \infty) \to (0, \infty)$ with $\lim_{t\to\infty} \psi(t) = 0$ such that for every bounded, Borel measurable function $f : \mathbb{R}^d \to \mathbb{R}$ and $\epsilon \in (0, 1)$ there holds*

$$\|\mathcal{P}_t f - \mu_\epsilon(f)\|^2_{L^2_{\mu_\epsilon}} \leq \psi(\epsilon t)\|f - \mu_\epsilon(f)\|^2_{L^\infty}. \tag{8.19}$$

Using standard arguments (see the ODE computation in [Theorem 2.1, [116]]), we can prove Lemma 8.3 by showing that $\mathcal{P}^\epsilon_t$ satisfies a type of uniform-in-ϵ, hypoelliptic weak Poincaré inequality, which is stated as follows.

Lemma 8.4. *Suppose that Assumption 3 holds. Then, there exists a nonincreasing function $\beta : (0, \infty) \to [1, \infty)$ so that for every $s > 0$ and bounded, measurable $f : \mathbb{R}^d \to \mathbb{R}$ there holds, uniformly in $t_0 \geq 1$ and $\epsilon \in (0, 1)$,*

$$\begin{aligned}\int_{t_0+1/4}^{t_0+3/4} \int |\mathcal{P}^\epsilon_t f - \mu_\epsilon(f)|^2 d\mu_\epsilon dt \leq \beta(s) \sum_{j=1}^{r} \int_{t_0}^{t_0+1} \int |Z_j \mathcal{P}^\epsilon_t f|^2 d\mu_\epsilon dt \\ + s\|f - \mu_\epsilon(f)\|^2_{L^\infty}.\end{aligned} \tag{8.20}$$

The proof of Lemma 8.4 is the place where the uniform-in-ϵ estimates on f_ϵ are employed in an essential way. The main idea is to use the upper and lower bounds on f_ϵ to reduce Lemma 8.4 to proving a suitable Poincaré type inequality for Lebesgue measure on balls $B(0,R) \subset \mathbb{R}^d$. This inequality is the content of Lemma 11.1. Its proof uses the Hörmander inequality, Lemma 9.3, and a compactness-rigidity argument that relies on ideas from the theory of local controllability (see Lemma 11.2).

We now prove Theorem 6.5 assuming Lemmas 6.3 and 8.3.

Proof of Theorem 6.5. By Lemma 8.2, we just need to prove (8.12). Fix $R \geq 1$. By Lemma 6.3 and the monotonicity of $\mathcal{P}_t^\epsilon$ with respect to $L^2_{\mu_\epsilon}$ we have, for any $t \geq 2$,

$$\sup_{\epsilon\in(0,1),\|f\|_\infty\leq 1} \|\mathcal{P}^\epsilon_{t\epsilon^{-1}} f - \mu_\epsilon(f)\|_{L^\infty(B_R)} \lesssim_R \sup_{\epsilon\in(0,1),\|f\|_\infty\leq 1} \|\mathcal{P}^\epsilon_{\frac{t}{2}\epsilon^{-1}} f - \mu_\epsilon(f)\|_{L^2_{\mu_\epsilon}}.$$

Hence, applying Lemma 8.3 we obtain

$$\limsup_{t\to\infty} \sup_{\epsilon\in(0,1),\|f\|_\infty\leq 1} \|\mathcal{P}^\epsilon_{t\epsilon^{-1}} f - \mu_\epsilon(f)\|_{L^\infty(B_R)}$$
$$\lesssim \limsup_{t\to\infty} \sup_{\epsilon\in(0,1),\|f\|_\infty\leq 1} \sqrt{\psi(t/2)}\|f - \mu_\epsilon(f)\|_{L^\infty} \lesssim \limsup_{t\to\infty} \sqrt{\psi(t/2)} = 0,$$

as desired. □

Chapter 9: Uniform Hörmander Inequalities

The proof of Theorem 6.2, carried out in the next chapter, is based on adapting aspects of the classical De Giorgi-Nash-Moser estimates for elliptic PDEs (see e.g. [59, 119]) to our hypoelliptic setting. This requires a quantitative understanding of functional inequalities tuned to the a priori estimates available for solutions to $L_\epsilon^* f = 0$. These inequalities, which we refer to generally as *Hörmander inequalities*, play the role that the classical Sobolev embedding theorem does in elliptic PDEs. The present chapter is dedicated to statements and proof sketches of the Hörmander inequalities needed to prove Theorem 6.2 and Lemma 6.3.

We begin here by investigating the uniform-in-ϵ a priori estimates available for solutions to the problem

$$\begin{cases} L_\epsilon^* f_\epsilon = 0, \\ f_\epsilon \geq 0, \\ \int f_\epsilon = 1, \end{cases} \tag{9.1}$$

This serves as both motivation for the functional inequalities to be discussed in the present chapter and the natural starting point for a proof of Theorem 6.2. From a probabilistic point-of-view, the most immediate estimate is the exponential moment bound (7.15), which in the context of (9.1) follows formally by integrating $0 = V L_\epsilon^* f_\epsilon$ and then using

$\int f_\epsilon = 1$ along with the fact that $V(x) = e^{\gamma x^2}$ is a uniform Lyapunov function (see Definition 6.4) for $\gamma > 0$ sufficiently small. On the other hand, from the perspective of elliptic PDEs, the natural a priori bound is the energy estimate that holds for sufficiently regular, nonnegative subsolutions obtained by pairing $L_\epsilon^* f \geq 0$ with f and integrating by parts. Let $\mathscr{X}$ denote the natural energy norm defined by

$$\|f\|_{\mathscr{X}} := \|f\|_{L^2} + \sum_{j=1}^{r} \|Z_j f\|_{L^2}.$$

The contributions from N and B to the energy estimate both vanish due to $\nabla \cdot Bx = \nabla \cdot N = 0$, and so we obtain

$$L_\epsilon^* f \geq 0 \implies \|f\|_{\mathscr{X}} \lesssim \|f\|_{L^2}. \tag{9.2}$$

If L_ϵ^* is elliptic; i.e., there exists $c > 0$ such that

$$\sum_{j=1}^{r} |Z_j \cdot \xi|^2 \geq c|\xi|^2 \quad \forall \xi \in \mathbb{R}^d,$$

then a suitable localization of (9.2) combined with the Sobolev embedding $H^1 \hookrightarrow L^{2q}$ for some $q > 1$ yields a local gain of integrability for subsolutions. This observation is crucial to the classical De Giorgi-Nash-Moser estimates of elliptic theory. In the present hypoelliptic setting, it is not possible for the $\mathscr{X}$ norm alone to control the H^s norm for any $s > 0$, and consequently one cannot directly use the standard Sobolev embedding theorem to gain integrability. Thus, (9.2) must be supplemented with some type of uniform estimate on the drift vector field

$$Z_{0,\epsilon} := \epsilon A x + \epsilon^{\alpha} B x + N. \tag{9.3}$$

This bound is far more subtle than (9.2) and comes in the form of an estimate in the norm dual to $\mathscr{X}$. In particular, define the norm

$$\|f\|_{\mathscr{X}^*} := \sup_{\varphi \in C_0^\infty, \|\varphi\|_{\mathscr{X}} \leq 1} \int \varphi f.$$

Then, using $L_\epsilon^* f_\epsilon = 0$ and (9.2), we see that

$$\|Z_{0,\epsilon} f_\epsilon\|_{\mathscr{X}^*} = \sup_{\varphi \in C_0^\infty, \|\varphi\|_{\mathscr{X}} \leq 1} \epsilon \int \varphi \left(\sum_{j=1}^{r} Z_j^2 + \mathrm{Tr}(A) \right) f_\epsilon \lesssim \epsilon \|f_\epsilon\|_{\mathscr{X}} \lesssim \epsilon \|f_\epsilon\|_{L^2}. \tag{9.4}$$

We summarize the latter two a priori bounds discussed above in the following lemma.

Lemma 9.1 (Uniform-in-ϵ a priori estimates)**.** *Let $f \geq 0$ be a sufficiently smooth and well localized solution to $L_\epsilon^* f \geq 0$. Then, f satisfies the energy estimate*

$$\|f\|_{\mathscr{X}} \lesssim \|f\|_{L^2}. \tag{9.5}$$

If in addition $L_\epsilon^* f = 0$, then

$$\|Z_{0,\epsilon} f\|_{\mathscr{X}^*} \lesssim \|f\|_{L^2}. \tag{9.6}$$

All of the implicit constants above do not depend on ϵ.

Remark 25. Notice that Lemma 9.1 does *not* contain uniform-in-ϵ control on the L^2 norm of f_ϵ. Such a bound is more difficult to deduce and will be proven later in Section 10.1.

In light of Lemma 9.1, it is natural to seek a Sobolev inequality that uses both $\|f\|_{\mathscr{X}}$ and $\|Z_{0,\epsilon}f\|$. That is, defining

$$\|f\|_{H^1_{\text{hyp}}} := \|f\|_{\mathscr{X}} + \|Z_{0,\epsilon}f\|_{\mathscr{X}^*}, \tag{9.7}$$

one wishes to prove the following, which will play a fundamental role in the developments of Chapter 10.

Lemma 9.2 (Hörmander inequality for H^1_{hyp})**.** *There exists $s > 0$ such that for all $R \geq 1$ and $f \in C_0^\infty(B_R)$ there holds, uniformly in $\epsilon \in (0,1)$,*

$$\|f\|_{H^s} \lesssim R^{1-s}\|f\|_{H^1_{hyp}}. \tag{9.8}$$

Remark 26. The dependence on R in (9.8) is natural to expect from scaling considerations, though the precise power of R will not play an important role in our analysis. For our purposes, what matter is the uniformity in ϵ and the fact that the scaling is polynomial in R instead of, say, e^{R^2}. For the role of this latter fact, see the proof of (10.17) in the next chapter.

The idea that (9.7) is the natural norm for extending elliptic regularization to hypoelliptic operators with the general form of L_ϵ^* dates back to Hörmander's seminal paper on hypoellipticity [79] (see also discussions in [7]). In fact, except for determining the dependence on R, Lemma 9.2 follows directly from a careful reading of [79]. We should also mention the recent work [7], which is the first to use the notation H^1_{hyp}. In this paper, the authors develop a well-posedness theory in the complete space associated with a norm analogous to $\|\cdot\|_{H^1_{\text{hyp}}}$ for the kinetic Fokker-Planck equation that mimics the classical H^1 variational theory for elliptic PDEs. In our present study we need only to be concerned with a priori estimates, but the terminology neverthless remains quite natural.

The remainder of this chapter is dedicated to statements and proof sketches of various uniform-in-ϵ Hörmander inequalities. In Section 9.1 we set up the needed notation and recall some basic facts. In Section 9.2 we sketch the proof of a generalized version of Lemma 9.2 (Lemma 9.4) that is uniform in $\delta \in [0,1]$ and adapted to the a priori estimates provided by the regularized operator $\epsilon\delta\Delta + L_\epsilon^*$. We will also give a Hörmander inequality (Lemma 9.3) that allows one to include the L^∞ norm in the definition of $\|\cdot\|_{\mathscr{X}}$ and will be important in the proof of (6.13). Finally, in Section 9.3 we sketch the proof of a somewhat subtle time-dependent Hörmander inequality that is crucial to the proof of Lemma 6.3.

9.1 Notation and basic facts

In this brief section we define notation and collect some basic facts that will be needed in the proof sketches to follow. Throughout the remainder of the entire chapter, $\Omega \subset \mathbb{R}^d$ denotes an open, bounded set and $K \subset \Omega$ is compact.

We use the notation in [79] for the L^2-based Hölder regularity of a function u along a vector field $X \in T(\Omega)$. For any $0 < t_0 \ll 1$ and $s \in (0,1]$ we write

$$|u|_{X,s}^{t_0} = \sup_{|t|\le t_0} |t|^{-s}\|e^{tX}u - u\|_{L^2}, \quad u \in C_0^\infty(K). \tag{9.9}$$

This is well defined since e^{tX} maps $C_0^\infty(K)$ into $C_0^\infty(\Omega)$ provided that $|t|$ is sufficiently small depending only on K, Ω, and the derivatives of X. We also define an isotropic s-norm by

$$|u|_s^{t_0} = \sup_{|h|\le t_0} |h|^{-s}\|u(\cdot+h)-u(\cdot)\|_{L^2}, \quad u \in C_0^\infty(K). \tag{9.10}$$

In Section 9.3 we will need to consider differential operators of the form $\epsilon\partial_t + X$ for $\epsilon > 0$ and functions that depend on time. In this situation, we write

$$|u|_{\epsilon\partial_t+X,s}^{t_0} = \sup_{|\tau|\le t_0} |\tau|^{-s}\|e^{\tau X}u(\cdot+\epsilon\tau,\cdot)-u(\cdot,\cdot)\|_{L^2}, \quad u \in C_0^\infty((a,b)\times K). \tag{9.11}$$

The seminorm $|\cdot|_s^{t_0}$ is related to the usual homogeneous Sobolev spaces by the equivalence

$$\|u\|_{\dot{B}^s_{p,r}} \approx_{s,p,r} \left\| \frac{\|u(\cdot-y)-u(\cdot)\|_{L^p}}{|y|^s} \right\|_{L^r(\mathbb{R}^d;|y|^{-d}dy)}, \tag{9.12}$$

which holds for any $s \in (0,1)$ and $(p,r) \in [1,\infty]^2$; see e.g. [Theorem 2.36, [9]]. Here, $\dot{B}^s_{p,r}$ denotes the usual homogeneous Besov space. We refer to [9] for definitions and basic results. A straightforward consequence of (9.12) and $\|u\|_{\dot{B}^s_{2,\infty}} \le \|u\|_{\dot{B}^s_{2,2}} \approx \|u\|_{\dot{H}^s}$ is that for any $s \in (0,1)$ and $s' > s$ there holds

$$|u|_s^{t_0} \lesssim_s \|u\|_{H^s} \lesssim_{s'-s} C(t_0,s)\left(\|u\|_{L^2} + |u|_{s'}^{t_0}\right), \quad u \in C_0^\infty(K), \tag{9.13}$$

where C is nonincreasing in $|t_0|$.

As usual, define $\mathrm{ad}X(Y) := [X,Y]$ for $X,Y \in T(\Omega)$. Then, for $\{(X_j,s_j)\}_{j=0}^r$

$\subseteq T(\Omega) \times (0,1]$ and a multi-index $I = (i_1, \dots, i_k)$, $0 \le i_j \le r$ we write

$$X_I = \mathrm{ad}X_{i_k}\mathrm{ad}X_{i_{k-1}}\dots\mathrm{ad}X_{i_2}X_{i_1},$$
$$\frac{1}{s(I)} = \sum_{j=1}^{k}\frac{1}{s_{i_j}}, \quad m(I) = \frac{1}{s(I)}. \tag{9.14}$$

9.2 Time-independent Hörmander inequalities

In this section we discuss the Hörmander inequalities that we will require in the proof of Theorem 6.2. We begin by defining the natural Hörmander norm pairs for working with the regularized operator $\epsilon\delta\Delta + L_\epsilon^*$. For $\{X_j\}_{j=1}^r \subseteq T(\Omega)$, an open set $\Omega' \subseteq \Omega$, and $\delta \in [0,1]$ we define

$$\|g\|_{\mathscr{X}_\delta(\Omega')} := \|g\|_{L^2(\Omega')} + \sum_{j=1}^{r}\|X_j g\|_{L^2(\Omega')} + \sqrt{\delta}\|\nabla g\|_{L^2(\Omega')},$$
$$\|g\|_{\mathscr{X}_\delta^*(\Omega')} := \sup_{\varphi\in C_0^\infty(\Omega'), \|\varphi\|_{\mathscr{X}_\delta(\Omega')}\le 1}\int_{\Omega'}\varphi g; \tag{9.15}$$

$$\|g\|_{\tilde{\mathscr{X}}_\delta(\Omega')} := \|g\|_{\mathscr{X}_\delta(\Omega')} + \|g\|_{L^\infty(\Omega')},$$
$$\|g\|_{\tilde{\mathscr{X}}_\delta^*(\Omega')} := \sup_{\varphi\in C_0^\infty(\Omega'), \|\varphi\|_{\tilde{\mathscr{X}}_\delta(\Omega')}\le 1}\int_{\Omega'}\varphi g. \tag{9.16}$$

Typically, the functions g we consider have compact support in Ω and $\Omega' = \Omega$. In this case, we do not indicate any domain in the notation. Also, by an abuse of notation, we will write ($\mathscr{X}_\delta$, $\mathscr{X}_\delta^*$), ($\mathscr{X}$, $\mathscr{X}^*$), etc., regardless of whether the vector fields involved are a general collection $\{X_j\}_{j=1}^r \subseteq T(\Omega)$ or the specific vector fields $\{Z_j\}_{j=1}^r \subseteq T(\mathbb{R}^d)$ from (6.1) since the meaning will always be clear from context.

The following lemma is a generalized and quantitative version of [(3.4), [79]] that holds uniformly in the regularization parameter δ and is indifferent to whether or not L^∞ is included in the Hörmander norm. The proof we give is a straightforward adaptation of the techniques from [79]. Recall the terminology from Definition 6.1.

Lemma 9.3 (Quantitative Hörmander inequality)**.** *Let $\Omega \subset \mathbb{R}^d$ be an open, bounded set, $K \subset \Omega$ be compact, and $\bar{\mathscr{X}}$ be either $\mathscr{X}_\delta$ or $\tilde{\mathscr{X}}_\delta$. Suppose that $\{X_j\}_{j=0}^r \subseteq T(\Omega)$ satisfies the uniform Hörmander condition on Ω with constants $(N_0, C_0) \in \mathbb{N} \times (0, \infty)$. There exists $s(N_0) > 0$ and a constant C such that for all $u \in C_0^\infty(K)$ and $\delta \in [0,1]$ there holds*

$$\|u\|_{H^s} \leq C\left(\|u\|_{\bar{\mathscr{X}}} + \|X_0 u\|_{\bar{\mathscr{X}}^*}\right).$$

The constant C depends on $\{X_j\}_{j=0}^r$ only through r, N_0, C_0, and an upper bound on $\sum_{j=0}^r \|X_j\|_{C^k(\Omega)}$ for some $k(N_0) > 0$ sufficiently large.

Proof sketch. Recall the definition of the norm $\mathscr{X}$ from the introduction to this chapter. For $\delta_1, \delta_2 \in [0,1]$, and functions $g \in C_0^\infty(\Omega)$ we define the Hörmander norm pair

$$\begin{aligned} \|g\|_{\mathscr{X}_{\delta_1,\delta_2}} &:= \|g\|_{\mathscr{X}} + \delta_1 \|\nabla g\|_{L^2} + \delta_2 \|g\|_{L^\infty}, \\ \|g\|_{\mathscr{X}^*_{\delta_1,\delta_2}} &:= \sup_{\varphi \in C_0^\infty(\Omega), \|\varphi\|_{\mathscr{X}_{\delta_1,\delta_2}} \leq 1} \int \varphi g. \end{aligned} \tag{9.17}$$

Our goal is to show that uniformly in $\delta_1, \delta_2 \in [0,1]$ there holds

$$\|u\|_{H^s} \leq C\left(\|u\|_{\mathscr{X}_{\delta_1,\delta_2}} + \|X_0 u\|_{\mathscr{X}^*_{\delta_1,\delta_2}}\right), \quad u \in C_0^\infty(K), \tag{9.18}$$

where s and C are as in the statement of the lemma. In the remainder of this proof, $C > 0$

denotes any such constant and u denotes an arbitrary function in $C_0^\infty(K)$.

Let $s_j = 1$ for $j = 1, \dots, r$, and $s_0 = 1/2$. Let $\sigma, s' > 0$ satisfy

$$N_0^{-1} \lesssim \sigma < s' < \min_{X_I \in V_{N_0}} s(I), \tag{9.19}$$

where X_I and V_{N_0} are as in (9.14) and Definition 6.1, respectively. Then, let $\mathcal{J}$ be the set of multi-indices with $\sigma m(I) \le 1$ that contain both zero and nonzero indices, and for $t > 0$ to be taken sufficiently small define

$$\bar{M}(u) = \|u\|_{\mathscr{X}} + \sum_{I \in \mathcal{J}} |u|_{X_I, s(I)} + |u|_\sigma^t. \tag{9.20}$$

Observe that $\bar{M}(u)$ is nothing more than the quantity $M(u)$ defined in the equation preceding [(5.6), [79]] but with the dual norm removed. It is clear from a reading of [79] that [Lemma 5.2, [79]] and [(5.16), [79]] both hold with $M(u)$ replaced $\bar{M}(u)$.

Let S_t denote the regularizer defined in [79] directly after the statement of [Theorem 5.1, [79]], and set $v_{t,\tau} = (e^{\tau X_0} S_t)^*(e^{\tau X_0} S_t u - S_t u)$ for $0 \le \tau \le t^2$. To prove (9.18), we follow the proof of [(3.4), [79]] exactly, except we replace the estimate of the second term in [(5.15), [79]] with

$$\begin{aligned} \left| \int_\Omega (X_0 u) v \right| &\le \|X_0 u\|_{\mathscr{X}^*_{\delta_1,\delta_2}} \|v_{t,\tau}\|_{\mathscr{X}_{\delta_1,\delta_2}} \\ &\lesssim \|X_0 u\|^2_{\mathscr{X}^*_{\delta_1,\delta_2}} + \|v_{t,\tau}\|^2_{\mathscr{X}} + \delta_1^2 \|\nabla v_{t,\tau}\|^2_{L^2} + \delta_2^2 \|v_{t,\tau}\|^2_{L^\infty}. \end{aligned}$$

Bounding $\|v_{t,\tau}\|_{\mathscr{X}}$ using [(5.16), [79]] with M replaced by $\bar{M}$, and then proceeding as in the computations after [(5.6), [79]] results in the following modified version of [(3.4), [79]]:

$$|u|^t_{s'} \lesssim \|u\|_{\mathscr{X}} + \|X_0 u\|_{\mathscr{X}^*_{\delta_1,\delta_2}} + \delta_1 \sup_{0<|\tau|\le t^2} \|\nabla v_{t,\tau}\|_{L^2} + \delta_2 \sup_{0<|\tau|\le t^2} \|v_{t,\tau}\|_{L^\infty}. \tag{9.21}$$

The estimate is uniform in δ_1, δ_2 and holds for t sufficiently small. Moreover, a careful reading of [79] shows that in addition to depending of course on K and Ω, both the implicit constant and the smallness requirement on t in (9.21) depend only on r, C_0, N_0, and an upper bound on $\sum_{j=0}^r \|X_j\|_{C^k(\Omega)}$ for some $k(N_0) > 0$. Applying (9.13), we thus obtain that for $\sigma < s < s'$ there holds

$$\|u\|_{H^s} \le C(\|u\|_{\mathscr{X}} + \|X_0 u\|_{\mathscr{X}^*_{\delta_1,\delta_2}} + \delta_1 \sup_{0<|\tau|\le t^2} \|\nabla v_{t,\tau}\|_{L^2} + \delta_2 \sup_{0<|\tau|\le t^2} \|v_{t,\tau}\|_{L^\infty}). \tag{9.22}$$

It remains to bound the latter two terms of (9.22) in terms of $\|u\|_{\mathscr{X}_{\delta_1,\delta_2}}$. Let $T_t = S_t$, e^{tX_0}, S_t^*, or $(e^{tX_0})^*$. From the definition of S_t (it is a finite product of operators that smooth along the vector fields X_I, $I \in \mathcal{J}$) it is straightforward to check that if V_1 and V_2 are open sets with $V_1 \subset\subset V_2 \subset\subset \Omega$, then for t sufficiently small depending only on r, N_0, V_1, V_2, and $\sum_{j=0}^r \|X_j\|_{C^k(\Omega)}$, for any $g \in C_0^\infty(V_1)$ there holds

$$T_t g \in C_0^\infty(V_2), \tag{9.23}$$

$$\|\nabla T_t g\|_{L^2} \le C\|g\|_{H^1}, \tag{9.24}$$

$$\|T_t g\|_{L^\infty} \le C\|g\|_{L^\infty}. \tag{9.25}$$

Combining this with (9.22) and recalling the definition of $v_{t,\tau}$ completes the proof. $\square$

For $\mathscr{X}_\delta$, $\mathscr{X}_\delta^*$ as in (9.15) with X_j replaced by Z_j and $g \in C_0^\infty(\mathbb{R}^d)$, let

$$\|g\|_{H^1_{\text{hyp},\delta}} := \|g\|_{\mathscr{X}_\delta} + \|Z_{0,\epsilon} g\|_{\mathscr{X}_\delta^*}, \tag{9.26}$$

which is nothing more than the natural δ-regularization of the H^1_{hyp} norm defined in (9.7). We now apply Lemma 9.3 to obtain a Hörmander inequality for $H^1_{\text{hyp},\delta}$ that is uniform in both $\delta \in [0,1]$ and $\epsilon \in (0,1)$. It is one of the key ingredients in the proofs of Lemmas 10.1 and 10.2 carried out in Chapter 10.

Lemma 9.4 (Hörmander inequality for $H^1_{\text{hyp},\delta}$)**.** *Let $R \geq 1$. There exists $s > 0$ such that for any $g \in C_0^\infty(B_R)$ there holds, uniformly in $\epsilon \in (0,1)$ and $\delta \in [0,1]$,*

$$\|g\|_{H^s} \lesssim R^{1-s} \|g\|_{H^1_{hyp,\delta}}.$$

Proof. Let $\bar{g}(x) = g(Rx)$ so that $\bar{g} \in C_0^\infty(B_1)$. Define

$$Z_0 = N + \epsilon R^{-p+1} A x + \epsilon^\alpha R^{-p+1} B x,$$

where p is the homogeneity degree of N. By Assumption 2, $\{Z_0, Z_1, \ldots, Z_r\}$ satisfies Hörmander's condition on B_2 with constants $(N_0, C_0) \in \mathbb{N} \times (0,\infty)$ that do not depend on ϵ, and so by Lemma 9.3 there exists $s(N_0) > 0$ such that

$$\|\bar{g}\|_{H^s} \lesssim \|\bar{g}\|_{\mathscr{X}_\delta} + \|Z_0 \bar{g}\|_{\mathscr{X}_\delta{}^*}. \tag{9.27}$$

The implicit constant in (9.27) depends on r, N_0, and C_0, but not on ϵ or δ. Now, if $\varphi \in C_0^\infty$ with $\|\varphi\|_{\mathscr{X}_\delta} \leq 1$, then by rescaling we have

$$\int_{B_1} \varphi Z_0 \bar{g} \leq \left\| \varphi\left(\frac{\cdot}{R}\right) \right\|_{\mathscr{X}_\delta} R^{-d-p+1} \|Z_{0,\epsilon} g\|_{\mathscr{X}_\delta{}^*} \leq R^{-d/2-p+1} \|Z_{0,\epsilon} g\|_{\mathscr{X}_\delta{}^*}.$$

Combining with $\|\bar{g}\|_{\mathscr{X}_\delta} \leq R^{-d/2+1} \|g\|_{\mathscr{X}_\delta}$ we obtain

$$\|\bar{g}\|_{\mathscr{X}_\delta} + \|Z_0 \bar{g}\|_{\mathscr{X}_\delta{}^*} \leq R^{-d/2+1} \|g\|_{H^1_{\text{hyp},\delta}}.$$

Since $\|\bar{g}\|_{H^s} = R^{-d/2+s} \|g\|_{H^s}$ it follows then from (9.27) that

$$R^{-d/2+s} \|g\|_{H^s} \lesssim R^{-d/2+1} \|g\|_{H^1_{\text{hyp},\delta}},$$

as desired. □

Remark 27. Since Lemma 9.3 does not use the parabolic Hörmander condition, Lemma 9.4 holds just as well when the uniform spanning condition in Theorem 6.2 is replaced with the analogous statement requiring only Hörmander's condition.

9.3 Hörmander inequality for spaces involving time

In this section, we discuss a parabolic Hörmander inequality that is natural for proving uniform-in-ϵ $L^2 \to L^\infty$ regularization estimates for the semigroup generated by $\epsilon^{-1} L_\epsilon$.

We begin with some notation. For an open, bounded set $\Omega \subset \mathbb{R}^d$, an open set

$\Omega' \subseteq \Omega$, $\{X_j\}_{j=1}^r \subseteq T(\Omega)$, $t_0 \in \mathbb{R}$, and $t > 0$ we define the Hörmander norm pair

$$\|g\|_{L^2((t_0,t_0+t);\mathscr{X}(\Omega'))} := \left(\int_{t_0}^{t_0+t} \|g(\tau)\|^2_{\mathscr{X}(\Omega')} d\tau \right)^{1/2}, \tag{9.28}$$

$$\|g\|_{L^2((t_0,t_0+t);\mathscr{X}^*(\Omega'))} := \sup_{\varphi \in C_0^\infty((t_0,t_0+t)\times\Omega'), \|\varphi\|_{L^2((t_0,t_0+t);\mathscr{X}(\Omega'))} \leq 1} \int_{\mathbb{R}\times\Omega'} \varphi g. \tag{9.29}$$

We use all of the same notations when $(\mathscr{X}, \mathscr{X}^*)$ is replaced with a different Hörmander norm pair.

In the above setting and notations, the parabolic Hörmander inequality is as follows.

Lemma 9.5 (Uniform parabolic Hörmander inequality)**.** *Let $K \subset \Omega$ be compact and suppose that $\{X_j\}_{j=0}^r \subseteq T(\Omega)$ satisfies the uniform parabolic Hörmander condition on Ω with constants $(N_0, C_0) \in \mathbb{N} \times (0, \infty)$. Fix $t_0 \in \mathbb{R}$, $t \in [1, 10]$, and $0 < \eta \leq 1/4$. There exists $s(N_0) > 0$ and a constant C such that for all $u \in C_0^\infty((t_0 + \eta, t_0 + t - \eta) \times K)$ there holds, uniformly in $\epsilon > 0$ and $\delta \in [0, 1]$,*

$$\int_{t_0}^{t_0+t} \|u(\tau, \cdot)\|^2_{H^s(\mathbb{R}^d)} d\tau \leq C \left(\|u\|^2_{L^2((t_0,t_0+t);\mathscr{X}_\delta)} + \|(\epsilon\partial_t + X_0)u\|^2_{L^2((t_0,t_0+t);\mathscr{X}_\delta^*)} \right).$$

The constant C is uniformly bounded with respect to η varying over compact time intervals away from the origin, and depends on $\{X_j\}_{j=0}^r$ only through r, N_0, C_0, and an upper bound on $\sum_{j=0}^r \|X_j\|_{C^k(\Omega)}$ for some $k = k(N_0)$ sufficiently large.

The remainder of this section is devoted to the proof of Lemma 9.5. We will assume throughout that $t = 1$, $t_0 = 0$, and $\eta = 1/4$ since the general case is no different. Moreover, by the same arguments from the proof of Lemma 9.3 it suffices to consider the case when $\delta = 0$. We will also suppress the superscript notation in (9.9)–(9.11), with the understanding that the increment is always taken sufficiently small depending only on K and finitely many derivatives of $\{X_j\}_{j=0}^r$. Lastly, unless otherwise stated, all implicit constants in this section satisfy the same properties as C from the lemma statement.

The proof of Lemma 9.5 is again based very closely on [79]. However, the generalization is a little more subtle than in Lemma 9.3 because to make use of the parabolic Hörmander condition we need to give the time direction a distinguished role, something not done in [79]. The first step is to generalize [Theorem 4.3, [79]], which says that a function $u : \mathbb{R}^d \to \mathbb{R}$ with some regularity along the vector fields $\{X_j\}_{j=0}^r$ must have some regularity in all directions. Recall from Definition 6.1 and Section 9.1 the definitions of V_j and the seminorms $|\cdot|_{X,s}$, $|\cdot|_{\epsilon\partial_t + X,s}$ for $X \in T(\Omega)$ and $s \in (0, 1]$.

Lemma 9.6. *Let $\{(X_j, s_j)\}_{j=0}^r \subseteq T(\Omega) \times [1/2, 1]$ and $\gamma \gtrsim N_0^{-1}$. For $|\tau|$ sufficiently small and any $X_I \in V_j$ with $j \lesssim_{N_0} 1$ there holds, uniformly in $\epsilon \in (0, 1)$,*

$$\int_0^1 \|e^{\tau^{m(I)} X_I} u(t,\cdot) - u(t,\cdot)\|_{L^2}^2 dt \lesssim \tau^2 |u|_{\epsilon\partial_t + X_0, s_0}^2 + \tau^2 \int_0^1 \left(\sum_{j=1}^r |u(t)|_{X_j, s_j}^2 + |u(t)|_\gamma^2 \right) dt \tag{9.30}$$

for every $u \in C_0^\infty((1/4, 3/4) \times K)$. As a consequence, when $\{X_j\}_{j=0}^r \subseteq T(\Omega)$ satisfies the parabolic Hörmander condition on Ω with constants (N_0, C_0) there exists $s(N_0) > 0$ so that for any $u \in C_0^\infty((1/4, 3/4) \times K)$ there holds, uniformly in $\epsilon \in (0, 1)$,

$$\int_0^1 |u(t)|_s^2 dt \lesssim \int_0^1 \|u(t)\|_{L^2}^2 dt + \sum_{j=1}^r \int_0^1 |u(t)|_{X_j, s_j}^2 dt + |u|_{\epsilon\partial_t + X_0, s_0}^2. \tag{9.31}$$

Proof sketch. Let I be a multi-index and $N \in \mathbb{N}$ be such that $N > |I|$, where as usual we write $|I|$ to denote the length of I. There exists a finite product decomposition (see [(4.13), [79]] and the discussion leading up to it)

$$e^{\tau^{m(I)}X_I} = \left(\Pi e^{\pm\tau^{m_j}X_j}\right) e^{\tau^{m(I_1)}X_{I_1}} \dots e^{\tau^{m(I_\ell)}X_{I_\ell}} H_N^\tau, \tag{9.32}$$

where each multi-index I_j, $1 \le j \le \ell$ satisfies $|I| < |I_j| \le N$, and

$$H_N^\tau v(x) = v(g(x,\tau)), \quad v \in C_0^\infty(K)$$

for a smooth mapping $g : K \times (-t_0, t_0) \to \Omega$ satisfying

$$\sup_{x\in K} |g(x,\tau) - x| = \mathcal{O}(|\tau|^N), \quad |\tau| \le t_0$$

for t_0 sufficiently small. The decomposition (9.32) is obtained by iteratively using that from the Cambell-Baker-Hausdorff formula one has

$$e^{-\tau X} e^{-\tau Y} e^{\tau X} e^{\tau Y} = e^{\tau^2[X,Y]+\dots}, \quad X, Y \in T(\Omega) \tag{9.33}$$

in the sense of formal power series, where $+\dots$ denotes a series of iterated commutators of length at least three formed with τX and τY; see e.g. [pg. 162, [79]]. Since $[\partial_t, X] = [\partial_t, Y] = 0$ for $X, Y \in T(\Omega)$ viewed as constant in time vector fields on $\mathbb{R}^{d+1}$, it is clear that (9.33) remains true with Y in the left-hand side replaced by $\epsilon\partial_t + Y$ for any $\epsilon > 0$. It follows that, when lifted to an operator on functions of spacetime, (9.32) holds with every occurrence of X_0 on the right-hand side replaced by $\epsilon\partial_t + X_0$. In particular, the error H_N^τ in the Taylor expansion acts only on the spatial variables:

$$H_N^\tau u(t,x) = u(t, g(x,\tau)), \quad u \in C_0^\infty((1/4,3/4)\times K).$$

By choosing $N \lesssim N_0$ such that $N\gamma \geq 1$, it follows then from [(4.11), [79]] and [Lemma 4.2, [79]] that for any $X_I \in V_j$ with $j \lesssim_{N_0} 1$ and $u \in C_0^\infty((1/4,3/4)\times K)$ there holds

$$\begin{aligned}\int_0^1 \|e^{\tau^{m(I)}X_I}u(t,\cdot) - u(t,\cdot)\|_{L^2}^2 dt \lesssim |\tau|^2 |u|_{\epsilon\partial_t + X_0}^2 + |\tau|^2 \sum_{j=1}^r \int_0^1 |u(t)|_{X_j,s_j}^2 dt \\ + \sum_{j=1}^{\ell} \int_0^1 \|e^{\tau^{m(I_j)}X_{I_j}}u(t,\cdot) - u(t,\cdot)\|_{L^2}^2 dt + |\tau|^2 \int_0^1 |u(t)|_\gamma^2 dt,\end{aligned} \tag{9.34}$$

where each I_j, $1 \leq j \leq \ell$ satisfies $|I| < |I_j| \leq N$. Using (9.34), the proof of (9.30) follows from the induction argument in [Lemma 4.6, [79]].

Now we turn to (9.31). Applying (9.30) and the arguments that lead to [(4.14), [79]] yields that for σ and s' as in (9.19) there holds

$$\begin{aligned}\sup_{0<|h|\ll 1} |h|^{-2s'} \int_0^1 \|u(t,\cdot+h) - u(t,\cdot)\|_{L^2}^2 dt \\ \lesssim |u|_{\epsilon\partial_t + X_0, s_0}^2 + \sum_{j=1}^r \int_0^1 |u(t)|_{X_j,s_j}^2 dt + \int_0^1 |u(t)|_\sigma^2 dt.\end{aligned} \tag{9.35}$$

Let now $\sigma < s < s'$. By (9.12) and Fubini's theorem we have the bound

$$\begin{aligned}\int_0^1 \|u(t)\|_{\dot H^s}^2 dt \approx \int_0^1 \|u(t)\|_{\dot B^s_{2,2}}^2 dt \\ \lesssim \sup_{0<|h|\ll 1} |h|^{-2s'} \int_0^1 \|u(t,\cdot+h) - u(t,\cdot)\|_{L^2}^2 dt + \int_0^1 \|u(t)\|_{L^2}^2 dt.\end{aligned}$$

Since $s > \sigma$, for every $\delta > 0$ there exists C_δ such that

$$|u(t)|_\sigma \lesssim \|u(t)\|_{\dot{B}^\sigma_{2,\infty}} \lesssim \|u(t)\|_{\dot{B}^\sigma_{2,2}} \leq \delta\|u(t)\|_{\dot{H}^s} + C_\delta\|u(t)\|_{L^2}.$$

The previous two estimates together with (9.35) yield (9.31) with $|u(t)|_s$ replaced by $\|u(t)\|_{H^s}$. The proof is then complete since $|u(t)|_s \lesssim \|u(t)\|_{H^s}$. □

With Lemma 9.6 at our disposal, the proof of Lemma 9.5 is a straightforward generalization of [Section 5, [79]]. Throughout the entire proof we write $L^2\mathscr{X}$ and $L^2\mathscr{X}^*$ to mean the norms taken on the time interval $(0,1)$. Also, for convenience we define $X_{0,\epsilon} = \epsilon\partial_t + X_0$.

Proof sketch of Lemma 9.5. For $\sigma > 0$, $\{s_j\}_{j=0}^r$, and $\mathcal{J}$ all as in the proof of Lemma 9.3, let

$$\begin{aligned}\tilde{M}(u) = \|u\|^2_{L^2\mathscr{X}} + \sum_{I\in\mathcal{J}} \sup_{0<|\tau|\ll 1} |\tau|^{-2}\int_0^1 \|e^{\tau^{m(I)}X_I}u(t) - u(t)\|^2_{L^2}dt \\ + \int_0^1 |u(t)|^2_\sigma dt.\end{aligned} \tag{9.36}$$

Note that $\tilde{M}(u)$ is not equivalent to $\int_0^1 |\bar{M}(u(t))|^2 dt$ because in the second term the supremum over the increment is outside of the time integral. Using Lemma 9.6, it follows from the arguments between [(5.6), [79]] and [(5.11), [79]] that to complete the proof of Lemma 9.5 it suffices to show that for $\tau > 0$ sufficiently small there holds

$$\int_0^1 \|e^{\tau^2 X_{0,\epsilon}}S_\tau u(t) - S_\tau u(t)\|^2_{L^2}dt \lesssim \tau^2\tilde{M}(u) + \tau^2\|X_{0,\epsilon}u\|^2_{L^2\mathscr{X}^*}, \tag{9.37}$$

where S_τ denotes the same regularizer introduced in the proof of Lemma 9.3.

To prove (9.37) we proceed as in [79] and define

$$f(s) = \left(\int_0^1 \| e^{sX_{0,\epsilon}} S_\tau u(t) - S_\tau u(t) \|_{L^2}^2 dt \right)^{1/2} \quad 0 < |s| \le \tau^2$$

with the goal of showing that $f(\tau) \lesssim |\tau| \sqrt{\tilde{M}(u)} + |\tau| \|X_{0,\epsilon}\|_{L^2 \mathscr{X}^*}$. Since S_τ does not regularize in the time variable we clearly have $[S_\tau, X_{0,\epsilon}] = [S_\tau, X_0]$, and so differentiating f^2 with respect to s gives

$$\begin{aligned} \frac{1}{2} \frac{d}{ds} f^2(s) = &\left\langle e^{sX_{0,\epsilon}} [X_0, S_\tau] u, e^{sX_{0,\epsilon}} S_\tau u - S_\tau u \right\rangle \\ &+ \left\langle X_{0,\epsilon} u, (e^{sX_{0,\epsilon}} S_\tau)^* (e^{sX_{0,\epsilon}} S_\tau u - S_\tau u) \right\rangle, \end{aligned} \tag{9.38}$$

where $\langle \cdot, \cdot \rangle$ denotes the L^2-inner product over $(0,1) \times \Omega$. To estimate the right-hand side of the expression above we need the following lemma, which is a variation of [Lemma 5.2, [79]].

Lemma 9.7. *Let V be an open set with $V \subset\subset \Omega$. For $\tau > 0$ sufficiently small and every $v \in C_0^\infty((0,1) \times V)$ there holds*

$$\int_0^1 \| \tau^{1/\sigma} \nabla S_\tau v(t) \|_{L^2}^2 dt \lesssim \tau^2 \tilde{M}(v) \tag{9.39}$$

$$\sum_{I \in \mathcal{J}} \int_0^1 \| \tau^{m(I)} X_I S_\tau v(t) \|_{L^2}^2 dt \lesssim \tau^2 \tilde{M}(v) \tag{9.40}$$

$$\sum_{j=0}^{r} \int_0^1 [\tau^{m_j} X_j, S_\tau] v(t) \|_{L^2}^2 dt \lesssim \tau^2 \tilde{M}(v). \tag{9.41}$$

A remark on the proof is in order because the second term in $\tilde{M}$ is weaker than

$$\sum_{I\in\mathcal{J}} \int_0^1 |u(t,\cdot)|^2_{X_I,s(I)} dt,$$

and so the lemma does not follow simply by integrating the estimates in [Lemma 5.2, [79]].

Proof sketch of Lemma 9.7. The left-hand side of (9.39) is bounded using the latter term in the definition of $\tilde{M}$. The estimate follows in the same manner as [(5.12), [79]] because in this term the supremum over the increment is inside the time integral.

Now we turn to (9.40). Let φ_X denote the regularizer defined in [Section 5, [79]]. By Minkowski's inequality, Jensen's inequality, and the equation preceding [(5.1), [79]] there holds

$$\int_0^1 \|\tau^{m(I)} X_I \varphi_{\tau^{m(I)}X_I} v(t)\|^2_{L^2} dt \leq \int_0^1 \int_{-1}^1 \|e^{s\tau^{m(I)}X_I} v(t) - v(t)\|^2_{L^2} |\varphi'(s)|^2 ds dt.$$

Employing also Fubini's theorem we then get

$$\begin{aligned}
&\int_0^1 \|\tau^{m(I)} X_I \varphi_{\tau^{m(I)}X_I} v(t)\|^2_{L^2} dt \\
&\leq \tau^2 \int_{-1}^1 |s|^{2/m(I)} |\varphi'(s)|^2 \left((|s|^{1/m(I)}|\tau|)^{-2} \int_0^1 \|e^{(|s|^{1/m(I)}\tau)^{m(I)}X_I} v(t) - v(t)\|^2_{L^2} dt \right) ds \\
&\leq \tau^2 \tilde{M}(v) \int_{-1}^1 |s|^{2/m(I)} |\varphi'(s)|^2 ds \lesssim \tau^2 \tilde{M}(v).
\end{aligned}$$

Hence, (9.40) holds with the summation replaced by a fixed $I \in \mathcal{J}$ and S_τ replaced by the individual regularizer $\varphi_{\tau^{m(I)}X_I}$. The induction trick in the proof of [Lemma 5.2, [79]] then works to upgrade to (9.40). Adapting the methods from [Lemma 5.2, [79]] to obtain (9.41) is done similarly. □

Using (9.41) in (9.38) we have

$$f(s)f'(s) \lesssim \tau^{-1}\sqrt{\tilde{M}(u)}f(s) + \|X_{0,\epsilon}u\|^2_{L^2\mathscr{X}^*} + \|(e^{sX_{0,\epsilon}}S_\tau)^*(e^{sX_{0,\epsilon}}S_\tau u - S_\tau u)\|^2_{L^2\mathscr{X}}.$$

From the elementary ODE computation proceeding [(5.15), [79]], it follows that to complete the proof of (9.37) it suffices to show that the latter term above can be controlled by $\tilde{M}(u)$. To this end, for $v \in C_0^\infty((0,1)\times\Omega)$ we define

$$\tilde{N}_\tau(v) = \|v(t)\|^2_{L^2\mathscr{X}} + \sum_{I\in\mathcal{J}}\int_0^1 \|\tau^{m(I)-1}X_I v(t)\|^2_{L^2}dt + \int_0^1 \|\tau^{1/\sigma-1}\nabla v(t)\|^2_{L^2}dt.$$

Applying Lemma 9.7 and the arguments that lead to [(5.17), [79]] and [(5.18), [79]] gives that for any open set $V \subset\subset \Omega$ and $v \in C_0^\infty((0,1)\times V)$, as long as τ is sufficiently small there holds

$$\tilde{N}_\tau(S_\tau v) \lesssim \tilde{M}(v), \tag{9.42}$$

$$\tilde{N}_\tau(e^{sX_I}v) \lesssim \tilde{N}_\tau(v), \qquad 0 \le |s| \le \tau^{m(I)}, \quad I \in \mathcal{J}. \tag{9.43}$$

Because $[\partial_t, X_0] = 0$ and $e^{\pm s\epsilon\partial_t}$ is bounded with respect to $L^2\mathscr{X}$ we have

$$\|(e^{sX_{0,\epsilon}}S_\tau)^*(e^{sX_{0,\epsilon}}S_\tau u(t) - S_\tau u(t))\|^2_{L^2\mathscr{X}} \lesssim \tilde{N}_\tau((e^{sX_0}S_\tau)^*e^{sX_0}S_\tau u) + \tilde{N}_\tau((e^{sX_0}S_\tau)^*S_\tau u).$$

Now, (9.42) and (9.43) along with the form of S_τ imply that $(e^{sX_0}S_\tau)^*$ is bounded with respect to $\tilde{N}_\tau$. Hence, we have

$$\|(e^{sX_{0,\epsilon}}S_\tau)^*(e^{sX_{0,\epsilon}}S_\tau u(t) - S_\tau u(t))\|^2_{L^2\mathscr{X}} \lesssim \tilde{N}_\tau(e^{sX_0}S_\tau u) + \tilde{N}_\tau(S_\tau u) \lesssim \tilde{M}(u),$$

which completes the proof. □

Chapter 10: Estimates on the Stationary Measure

In this Chapter we prove Theorem 6.2. The upper bound is obtained in Sections 10.1-10.4. In Section 10.1, we prove a uniform-in-ϵ L^2 upper found for f_ϵ using an interpolation argument together with the moment bound (7.9) and the uniform Hörmander inequality given in Lemma 9.4. In Section 10.2, we use a hypoelliptic Moser iteration to upgrade the L^2 upper bound to the uniform L^∞_{loc} estimate

$$\sup_{\epsilon\in(0,1)} \|f_\epsilon\|_{L^\infty(B_R)} \lesssim_R 1. \tag{10.1}$$

Section 10.3 is dedicated proving the lower bound (6.13). This involves adapting ideas from [64] and the classical De Giorgi proof of Hölder regularity (see e.g. [119]). We conclude the Chapter in Section 10.4, wherein we upgrade (10.1) to the Gaussian upper bound stated in Theorem 6.2.

Before proceeding to the main portion of this Chapter, recall from Remark 21 that in order to more easily justify formal calculations, we introduce the following regularization: $\forall \delta, \epsilon > 0$, define $f_{\epsilon,\delta} \geq 0$ with $\int f_{\epsilon,\delta} = 1$ to be the unique solution to the problem

$$\delta\Delta f_{\epsilon,\delta} + \frac{1}{\epsilon} L^*_\epsilon f_{\epsilon,\delta} = 0. \tag{10.2}$$

As discussed in Remarks 21-23 at the end of Chapter 7, this problem is well posed and $f_{\epsilon,\delta}$ satisfies the moment bound (7.13). Note also that from classical elliptic theory [59] there holds $\forall R > 0$,

$$||f_{\epsilon,\delta}||_{H^k(B_R)} \lesssim_{R,k,\delta,\epsilon} 1. \tag{10.3}$$

Moreover, we will prove in the Appendix that $f_{\epsilon,\delta} \in L^2$ for every $\delta > 0$, and that $\lim_{\delta\to 0} f_{\epsilon,\delta} = f_\epsilon$ in H^k_{loc} for each fixed $\epsilon > 0$ and $k \in \mathbb{N}$. Our general strategy in proving estimates for f_ϵ will be to obtain estimates for $f_{\epsilon,\delta}$ that are uniform in both ϵ and δ, and then to pass to the limit $\delta \to 0$.

10.1 Uniform L^2 estimate for f_ϵ

The purpose of this section is to prove the following lemma.

Lemma 10.1. *We have the uniform L^2 bound*

$$\sup_{\epsilon,\delta\in(0,1)} \|f_{\epsilon,\delta}\|_{L^2} \lesssim 1. \tag{10.4}$$

As a consequence,

$$\sup_{\epsilon\in(0,1)} \|f_\epsilon\|_{H^1_{hyp}} \lesssim \sup_{\epsilon\in(0,1)} \|f_\epsilon\|_{L^2} \lesssim 1. \tag{10.5}$$

Proof. We begin with the proof of (10.4), which is the bulk of the work. Let $\bar{\chi}(x) \in C_0^\infty(B_1)$ be radially symmetric and such that $\bar{\chi} = 1$ for $|x| < 1/2$. Define then $\chi(x) = \bar{\chi}(x/2) - \bar{\chi}(x)$, which is now a $C_0^\infty(B_2 \setminus B_{1/2})$ function. Further, define $\chi_R = \chi(x/R)$

and note that

$$1 = \bar{\chi} + \sum_{R=2^j : j\geq 0} \chi_R. \tag{10.6}$$

Step 1: estimates on $\bar{\chi} f_{\epsilon,\delta}$: Multiplying (10.2) by $\bar{\chi}$ and using the energy property $N(x) \cdot x = Bx \cdot x = 0$ together with the radial symmetry of $\bar{\chi}$, we obtain

$$(\epsilon\delta\Delta + L_\epsilon^*)\, \bar{\chi} f_{\epsilon,\delta} = [\epsilon\delta\Delta, \bar{\chi}] f_{\epsilon,\delta} + \epsilon \sum_{j=1}^{r} [Z_j^2, \bar{\chi}] f_{\epsilon,\delta} + \epsilon [Ax \cdot \nabla, \bar{\chi}] f_{\epsilon,\delta}. \tag{10.7}$$

Not to be confused with the commutator of vector fields, we write here for example $[\Delta, \bar{\chi}]$ to simply mean the operator that acts on smooth functions g via $[\Delta, \bar{\chi}]g = \Delta(\bar{\chi} g) - \bar{\chi}\Delta g$. Pairing with $\bar{\chi} f_{\epsilon,\delta}$ gives the a priori estimate

$$\delta \left|\left| \nabla(\bar{\chi} f_{\epsilon,\delta}) \right|\right|_{L^2}^2 + \sum_{j=1}^{r} \left|\left| Z_j(\bar{\chi} f_{\epsilon,\delta}) \right|\right|_{L^2}^2 \lesssim \left|\left| f_{\epsilon,\delta} \right|\right|_{L^2}^2 . \tag{10.8}$$

Similarly, we pair with a test function $v \in C_0^\infty(\mathbb{R}^d)$ satisfying $\|v\|_{\mathscr{X}_\delta} \leq 1$ and obtain, using (10.8),

$$\begin{aligned}
\left| \int v Z_{0,\epsilon} \bar{\chi} f_{\epsilon,\delta} dx \right| &\lesssim \epsilon\delta \left|\left| \nabla(\bar{\chi} f_{\epsilon,\delta}) \right|\right|_{L^2} \left|\left| \nabla v \right|\right|_{L^2} + \epsilon \sum_{j=1}^{r} \left|\left| Z_j(\bar{\chi} f_{\epsilon,\delta}) \right|\right|_{L^2} \left|\left| Z_j v \right|\right|_{L^2} \\
&\quad + \epsilon\delta \left|\left| f_{\epsilon,\delta} \right|\right|_{L^2} \left|\left| \nabla v \right|\right|_{L^2} + \epsilon \sum_{j=1}^{r} \left|\left| Z_j v \right|\right|_{L^2} \left|\left| f_{\epsilon,\delta} \right|\right|_{L^2} \\
&\quad + \epsilon \left|\left| f_{\epsilon,\delta} \right|\right|_{L^2} \left|\left| v \right|\right|_{L^2} \\
&\lesssim \epsilon \| f_{\epsilon,\delta} \|_{L^2}.
\end{aligned}$$

Combining with (10.8) we then have, uniformly in δ, ϵ,

$$||\bar{\chi} f_{\epsilon,\delta}||_{H^1_{\text{hyp},\delta}} \lesssim ||f_{\epsilon,\delta}||_{L^2} . \tag{10.9}$$

Thus, by Lemma 9.4 and Sobolev embedding, $\exists \theta \in (0,1)$ (depending on dimension but not ϵ) such that

$$||\bar{\chi} f_{\epsilon,\delta}||_{L^2} \lesssim ||\bar{\chi} f_{\epsilon,\delta}||_{L^1}^{1-\theta} ||\bar{\chi} f_{\epsilon,\delta}||_{H^s}^{\theta} \lesssim ||\bar{\chi} f_{\epsilon,\delta}||_{L^1}^{1-\theta} ||\bar{\chi} f_{\epsilon,\delta}||_{H^1_{\text{hyp},\delta}}^{\theta} \lesssim ||f_{\epsilon,\delta}||_{L^2}^{\theta} . \tag{10.10}$$

Step 2: estimates on $\chi_R f_{\epsilon,\delta}$: For any $R \geq 1$, by applying the same arguments as in the case of $\bar{\chi}$ and using $\|\nabla^j \chi_R\|_{L^\infty} \lesssim R^{-j}$ to control the commutator error terms, we similarly obtain

$$||\chi_R f_{\epsilon,\delta}||_{H^1_{\text{hyp},\delta}} \lesssim ||f_{\epsilon,\delta}||_{L^2} . \tag{10.11}$$

Therefore, again by Lemma 9.4 and Sobolev embedding, $\exists \theta \in (0,1)$ such that

$$||\chi_R f_{\epsilon,\delta}||_{L^2} \lesssim R\, ||\chi_R f_{\epsilon,\delta}||_{L^1}^{1-\theta} ||f_{\epsilon,\delta}||_{L^2}^{\theta} . \tag{10.12}$$

Step 3: L^2 **estimates:** By (10.6), Young's inequality, (10.10), and (10.12), we have

$$||f_{\epsilon,\delta}||_{L^2} \leq ||\bar{\chi} f_{\epsilon,\delta}||_{L^2} + \sum_{2^j : j \geq 0} ||\chi_{2^j} f_{\epsilon,\delta}||_{L^2} \tag{10.13}$$

$$\lesssim ||f_{\epsilon,\delta}||^{\theta}_{L^2} + ||f_{\epsilon,\delta}||^{\theta}_{L^2} \sum_{j \geq 0} 2^j \, ||\chi_{2^j} f_{\epsilon,\delta}||^{1-\theta}_{L^1} \tag{10.14}$$

$$\lesssim ||f_{\epsilon,\delta}||^{\theta} \left(1 + \sum_{j \geq 0} 2^{-\frac{j}{\theta}} + \sum_{j \geq 0} 2^{\frac{2j}{1-\theta}} \|\chi_{2^j} f_{\epsilon,\delta}\|_{L^1} \right) \tag{10.15}$$

$$\lesssim ||f_{\epsilon,\delta}||^{\theta}_{L^2} \left(1 + \sum_{2^j : j \geq 0} \left|\left| \langle \cdot \rangle^{\frac{2}{1-\theta}} \chi_{2^j} f_{\epsilon,\delta} \right|\right|_{L^1} \right). \tag{10.16}$$

Applying (7.15) to control the L^1 norm gives

$$||f_{\epsilon,\delta}||_{L^2} \lesssim ||f_{\epsilon,\delta}||^{\theta}_{L^2}. \tag{10.17}$$

The bound (10.4) then follows from $\theta < 1$ and $\|f_{\epsilon,\delta}\|_{L^2} < \infty$. □

With (10.4) at hand, the second inequality in (10.5) follows by sending $\delta \to 0$ and appealing to Lemma A.2 (specifically, we are using here (A.9)). It remains to prove the first inequality in (10.5). Let $\bar{\chi}_R(x) := \bar{\chi}(x/R)$. The same computations that led to (10.9) and (10.11) yield

$$\|\bar{\chi}_R f_{\epsilon}\|_{H^1_{\text{hyp}}} \lesssim \|f_{\epsilon}\|_{L^2}, \tag{10.18}$$

where importantly the implicit constant does not depend on R. Sending $R \to \infty$ yields the desired inequality, completing the proof.

Remark 28. The bound (10.18) holds equality well with $\delta > 0$. Thus, a consequence of the proof above (which we will require later) is that

$$\sup_{\epsilon,\delta\in(0,1)} \|f_{\epsilon,\delta}\|_{H^1_{\text{hyp},\delta}} \lesssim 1. \tag{10.19}$$

10.2 Hypoelliptic Moser iteration

In this section we carry out a Moser iteration scheme to obtain a local gain of integrability. Combining with the results of Section 10.1 will complete the proof of (10.1). The gain of integrability is stated precisely as follows.

Lemma 10.2. *Let $\delta \in (0,1)$ and suppose that $f \in C^\infty(\mathbb{R}^d)$ satisfies $f \geq 0$ and*

$$(\epsilon\delta\Delta + L_\epsilon^*)f \geq 0. \tag{10.20}$$

Then, for any $R \geq 1$, uniformly in $\epsilon, \delta \in (0,1)$ there holds

$$\|f\|_{L^\infty(B_R)} \lesssim_R \|f\|_{L^2(B_{2R})}.$$

Proof. Let $f \in C^\infty(\mathbb{R}^d)$ satisfy $f \geq 0$ and $(\epsilon\delta\Delta + L_\epsilon^*)f \geq 0$. By replacing f with $f + \epsilon'$ and then sending $\epsilon' \to 0$ we may assume without loss of generality that $f > 0$. Fix $R \geq 1$ and for each $k \geq 0$ define $R_k = R(1 + 2^{-k})$. With s as given in Lemma 9.4, let $\alpha > 1$ be such that $H^s \hookrightarrow L^{2\alpha}$ and define $w_k = f^{\alpha^k}$. We prove that $\exists C > 0$ (depending only on R and dimension) such that for $k \geq 0$,

$$||w_k||_{L^{2\alpha}(B_{R_{k+1}})} \leq C^k \, ||w_k||_{L^2(B_{R_k})} \, . \tag{10.21}$$

By the convexity of $z \mapsto z^\beta$,

$$\delta \Delta w_k + \sum_{j=1}^{r} Z_j^2 w_k + \frac{1}{\epsilon} Z_{0,\epsilon} w_k + \alpha^k \mathrm{Tr} A w_k \geq 0. \tag{10.22}$$

Let $\chi_k \in C_0^\infty(B_{R_k})$ be a radially-symmetric, smooth cutoff function satisfying $\chi_k(x) = 1$ for $|x| \leq R_{k+1}$ and $|D^\beta \chi_k| \lesssim R^{-1} 2^{|\beta| k}$ for every multi-index β with $|\beta| \leq 2$. Denoting $v_k = \chi_k w_k$ and using (10.22) we obtain

$$\delta \Delta v_k + \sum_{j=1}^{r} Z_j^2 v_k + \frac{1}{\epsilon} Z_{0,\epsilon} v_k + \alpha^k \mathrm{Tr} A v_k - \mathcal{C} \geq 0, \tag{10.23}$$

where

$$\mathcal{C} = [\delta \Delta, \chi_k] w_k + \sum_{j=1}^{r} [Z_j^2, \chi_k] w_k + [Ax \cdot \nabla, \chi_k] w_k. \tag{10.24}$$

Pairing with $v_k = \chi_k w_k$ we obtain the a priori estimate

$$\delta \|\nabla v_k\|_{L^2}^2 + \sum_{j=1}^{r} \|Z_j v_k\|_{L^2}^2 \lesssim \alpha^k \, ||v_k||_{L^2}^2 + 2^{2k} \, ||w_k||_{L^2(B_{R_k})}^2 . \tag{10.25}$$

Let g be the unique solution to the Dirichlet problem

$$\begin{cases} \delta \Delta g + \sum_{j=1}^{r} Z_j^2 g + \frac{1}{\epsilon} Z_{0,\epsilon} g + \alpha^k \mathrm{Tr} A v_k - \mathcal{C} = 0 \\ g|_{\partial B_{2R+1}} = 0. \end{cases} \tag{10.26}$$

By the weak elliptic maximum principle we have $v_k \leq g$ and, in particular, for all L^p, we have $||v_k||_{L^p} \leq ||g||_{L^p}$. Moreover, we have the a priori estimate

$$\delta \, ||\nabla g||_{L^2}^2 + \sum_{j=1}^{r} \|Z_j g\|_{L^2}^2 + ||g||_{L^2}^2 \lesssim_R \alpha^{2k} \, ||v_k||_{L^2}^2 + 2^{4k} \, ||w_k||_{L^2(B_{R_k})}^2 . \tag{10.27}$$

Multiplying by a radially-symmetric, smooth cutoff $\chi \in C_0^\infty(B_{2R+1/2})$ with $\chi(x) = 1$ for $|x| \leq 2R$ and applying the arguments we used in the proof of Lemma 10.1 we obtain

$$\|\chi g\|_{H^1_{\text{hyp},\delta}} \lesssim_R \alpha^{2k} \|v_k\|_{L^2} + 2^{2k} \|w_k\|_{L^2(B_{R_k})}, \tag{10.28}$$

and so by Lemma 9.4 we have

$$\|w_k\|_{L^{2\alpha}(B_{R_{k+1}})} \leq ||v_k||_{L^{2\alpha}} \leq ||\chi g||_{L^{2\alpha}} \lesssim_R \alpha^{2k} \, ||v_k||_{L^2} + 2^{2k} \, ||w_k||_{L^2(B_{R_k})} . \tag{10.29}$$

This completes the proof of the iteration (10.21).

The bound (10.21) implies that for some $C > 0$ (depending only on R and dimension) there holds

$$||f||_{L^{2\alpha^{k+1}}(B_{R_{k+1}})} \leq C^{k\alpha^{-k}} \, ||f||_{L^{2\alpha^k}(B_{R_k})} , \tag{10.30}$$

which by iteration gives

$$||f||_{L^{2\alpha^{k+1}}(B_{R_{k+1}})} \leq C^{\sum_{j=0}^{k} j\alpha^{-j}} \, ||f||_{L^2(B_{2R})} \tag{10.31}$$

for every $k \geq 0$. Using that $\alpha > 1$, we pass to the limit $k \to \infty$ and obtain the desired result. ∎

Remark 29. The comparison principle argument used in the proof above is a modification of an idea from [64].

We are now ready to complete the proof of (10.1).

Proof of (10.1). Combining Lemmas 10.1 and 10.2 we have, for every $R \geq 1$,

$$\sup_{\epsilon,\delta\in(0,1)} \|f_{\epsilon,\delta}\|_{L^\infty(B_R)} \lesssim_R \sup_{\epsilon,\delta\in(0,1)} \|f_{\epsilon,\delta}\|_{L^2(B_{2R})} \lesssim 1. \tag{10.32}$$

The bound (10.1) then follows from sending $\delta \to 0$ and using Lemma A.2. ∎

10.3 Proof of the lower bound for f_ϵ

In this section we prove the uniform-in-ϵ lower bound for f_ϵ stated in (6.13). The calculations will be somewhat more technical than those in Sections 10.1 and 10.2. As such, we will begin by outlining the key ideas. The main lemmas that go into the proof of (6.13) will be stated in this section and proven in Section 10.3.1 and Section 10.3.2.

The observation that motivates our general approach is that (10.1) and (7.15) together imply that for every $R \geq 1$ sufficienty large there are constants $c_1, c_2 > 0$ such that uniformly in $\epsilon \in (0,1)$ there holds

$$|\{x \in B_R : f_\epsilon \geq c_1\}| \geq c_2. \tag{10.33}$$

In other words, f_ϵ stays uniformly bounded away from zero on a set of positive measure.

A classical idea in the Hölder regularity theory for second-order elliptic equations with rough coefficients is that weak solutions "cannot oscillate too much," which in the context of nonnegative solutions can be made precise as follows: if $0 \leq u \leq 1$ solves $Lu = 0$ on B_2 for a suitable elliptic operator L, then $|\{x \in B_1 : u(x) \geq 1/2\}| > 0$ implies that u must remain uniformly bounded below away from 0 *everywhere* on the smaller set $B_{1/2}$; see for example the review [119] (in particular, Proposition 9) and the references therein. The now-standard techniques used to prove such results allow one to show that when L_ϵ^* is elliptic the bound (10.33) implies

$$\inf_{x \in B_R} f_\epsilon(x) \gtrsim_{R,\epsilon,c_1,c_2} 1. \tag{10.34}$$

We have indicated that the implicit constant a priori depends on ϵ, although we will see below that this is not the case.

Our strategy is to extend the argument that yields (10.34) from (10.33) to the hypoelliptic setting (and with a constant independent of ϵ). The proof is based on a clever trick from De Giorgi's approach to Hölder regularity for elliptic PDEs with rough coefficients (see e.g. [119]). We now briefly recall the main idea. For $\theta \in (0,1)$ consider the rescaled functions

$$w_k = \left(1 - \theta^{-k}(f_\epsilon/c_1)\right)_+, \tag{10.35}$$

which are structured so that

$$|\{x \in B_R : w_k = 0\}| \geq c_2, \tag{10.36}$$

$$|\{x \in B_R : w_k \geq 1-\theta\}| \geq \int_{B_R} |w_{k+1}|^2. \tag{10.37}$$

Observe now that a uniform lower bound on f_ϵ follows provided that

$$\liminf_{k\to\infty} \|w_k\|_{L^2(B_R)} = 0$$

uniformly in $\epsilon \in (0,1)$. Indeed, $L_\epsilon^* w_k \geq 0$ by the convexity of $z \mapsto z_+$, and hence Lemma 10.2 (suppose for the sake of discussion that it holds for $f \in H^1$ and $\delta = 0$) gives, for all $k \in \mathbb{N}$,

$$\inf_{x\in B_{R/2}} f_\epsilon(x) \geq c_1\theta^k \left(1 - \|w_k\|_{L^\infty(B_{R/2})}\right) \geq c_1\theta^k \left(1 - C(R)\|w_k\|_{L^2(B_R)}\right).$$

The proof that $\|w_k\|_{L^2}$ eventually gets small uses an iteration argument that hinges on the ability to control how quickly a nonnegative subsolution $L_\epsilon^* f \geq 0$ can oscillate. More precisely, one needs to show that

$$|\{x \in B_R : w_k \geq 1-\theta\}| \geq \kappa > 0, \quad |\{x \in B_R : w_k = 0\}| \geq c_2, \tag{10.38}$$

and $L_\epsilon^* w_k \geq 0$ together imply

$$|\{x \in B_R : 0 < w_k < 1-\theta\}| \gtrsim_{\kappa,c_1,c_2,R} 1. \tag{10.39}$$

If this bound is true, then for $\kappa > 0$ chosen as small as we wish, for some sufficiently large k, we must have

$$\|w_{k+1}\|^2_{L^2(B_R)} < \kappa(R).$$

Indeed, if $\|w_{k+1}\|^2_{L^2(B_R)} \geq \kappa(R) > 0$, then (10.38) is satisfied due to (10.37), and so the claim follows from (10.39) and the fact that the sets $\{\{0 < w_k < 1-\theta\}\}_{k=1}^{\infty}$ are pairwise disjoint; see the proof of Lemma 10.7 for more explanation. It is important to note that proving (6.13) requires the constant in (10.39) to be independent of ϵ.

In the elliptic setting, the "intermediate value estimate" (10.39) is provided by the natural H^1 energy estimate and the classical De Giorgi isoperimetric inequality (see e.g. [Lemma 10, [119]]), which explicitly quantifies a lower bound for $|\{x \in B_R : 0 < w_k < 1-\theta\}|$ in terms of $\|w_k\|_{H^1(B_R)}$ and the quantities in (10.38). This approach does not apply in the hypoelliptic setting because $L^*_\epsilon w_k \geq 0$ is not sufficient to provide a uniform-in-ϵ bound on $\|w_k\|_{H^1(B_R)}$. Nevertheless, for each fixed $R \geq 1$ we are able to prove an intermediate value lemma that holds uniformly in $0 < \epsilon \ll 1$. It is stated as follows (recall the notation from (10.2)).

Lemma 10.3 (An intermediate value lemma)**.** *Fix $R \geq 1$ and $\alpha_1, \alpha_2 > 0$. There exists $\epsilon_0 > 0$, $\mu > 0$, and $\theta \in (0, 1/2)$ such that if $\epsilon \leq \epsilon_0$, $\delta \in (0,1)$, and $w \in C^\infty(B_{2R})$ with $0 \leq w \leq 1$ satisfies*

$$0 \leq \delta \Delta w + \frac{1}{\epsilon} L^*_\epsilon w \leq \frac{1}{\sqrt{\epsilon}} \left(1 + \delta |\nabla f_{\epsilon,\delta}|^2 + \sum_{j=1}^{r} |Z_j f_{\epsilon,\delta}|^2 \right) \tag{10.40}$$

on B_{2R}, then the inequalities

$$|\{w = 0\} \cap B_R| \geq \alpha_1$$

and

$$|\{w \geq 1 - \theta\} \cap B_R| \geq \alpha_2$$

together imply

$$|\{0 < w < 1 - \theta\} \cap B_R| \geq \mu.$$

The proof of Lemma 10.3 is carried out in Section 10.3.1 and follows a compactness -rigidity argument motivated by [Lemma 14, [64]]. The desired compactness is deduced with a uniform Hörmander inequality. In particular, we will apply Lemma 9.3 in the case where L^∞ is included in the Hörmander norm to obtain $\|w\|_{H^s} \lesssim 1$. We use this lemma instead of Lemma 9.4 since (10.40) is too weak to provide a uniform estimate on $\|Z_{0,\epsilon} w\|_{\mathscr{X}^*}$. The rigidity step consists of passing to the limit and deriving a contradiction with the supposed counter-example obtained at $\theta = \epsilon = \mu = 0$ satisfying the ϵ-independent estimates provided by (10.40). It turns out that this only requires one to know that there cannot exist a non-constant characteristic function ξ satisfying (in the sense of distributions) $N \cdot \nabla \xi = 0$ and $Z_j \cdot \nabla \xi = 0$ for $j = 1, \ldots, r$, which in fact follows directly from Hörmander's theorem. Later on when proving Theorem 6.5 we will need a stronger and much more interesting rigidity statement (see Lemma 11.2).

We have already sketched the main ideas in using a uniform-in-ϵ intermediate value lemma to obtain a local lower bound, though due to the complexity of Lemma 10.3 there are some additional details to fill in. This is done in Section 10.3.2, wherein we prove the following.

Lemma 10.4. *Suppose that Lemma 10.3 holds. Then, for all $R \geq 1$ there exists $\epsilon_*(R) > 0$ such that*

$$\inf_{\epsilon \in (0,\epsilon_*)} \inf_{|x| \leq R} f_\epsilon(x) \gtrsim_R 1. \tag{10.41}$$

If Assumption 3 is satisfied, then (6.13) also holds.

10.3.1 Proof of the intermediate value lemma

The purpose of this section is to prove Lemma 10.3.

Proof of Lemma 10.3. If the lemma fails, then there exists a sequence $\{(\delta_n, \epsilon_n)\}_{n=1}^\infty \subseteq (0,1) \times (0,1)$ with $\lim_n \epsilon_n = 0$ and $\{w_n\}_{n=1}^\infty \subseteq C^\infty(B_{2R})$ satisfying the following properties:

- $0 \leq w_n \leq 1$
- $|\{w_n = 0\} \cap B_R| \geq \alpha_1$
- $|\{w_n \geq 1 - \frac{1}{n}\} \cap B_R| \geq \alpha_2$
- $|\{0 < w_n < 1 - \frac{1}{n}\} \cap B_R| < \frac{1}{n}$;

and moreover

$$0 \leq \delta_n \epsilon_n \Delta w_n + L_{\epsilon_n}^* w_n \leq \sqrt{\epsilon_n} \left(1 + \delta_n |\nabla f_{\epsilon_n,\delta_n}|^2 + \sum_{j=1}^r |Z_j f_{\epsilon_n,\delta_n}|^2 \right). \tag{10.42}$$

By the uniform estimate $\|w_n\|_{L^\infty} \le 1$ and the Banach-Alaoglu theorem, $\exists w \in L^\infty$ such that

$$w_n \rightharpoonup_* w$$

in L^∞ up to extracting a subsequence (not relabled).

Now we obtain the needed compactness. Let $\chi \in C_0^\infty(B_{2R})$ be radially symmetric with $0 \le \chi \le 1$ and $\chi(x) = 1$ for $|x| \le R$. From the lower bound in (10.42) and the arguments that led to (10.25), we have

$$\delta_n \|\nabla(\chi w_n)\|_{L^2}^2 + \sum_{j=1}^r \|Z_j(\chi w_n)\|_{L^2}^2 \lesssim \int_{B_{2R}} |w_n|^2 dx \lesssim |B_{2R}|\,, \tag{10.43}$$

where the constant is independent of n using $0 \le w_n \le 1$. Moreover, pairing (10.42) with $\chi\varphi$ for $\varphi \in C_0^\infty$ yields

$$\begin{aligned}\left|\int Z_{0,\epsilon_n}(\chi w_n)\varphi\right| &\lesssim \sum_{j=1}^r \|Z_j\varphi\|_{L^2}\|Z_j(\chi w_n)\|_{L^2} + \delta_n\|\nabla\varphi\|_{L^2}\|\nabla(\chi w_n)\|_{L^2} \\ &\quad + \|w_n\|_{L^2(B_{2R})}\left(\|\varphi\|_{L^2} + \delta_n\|\nabla\varphi\|_{L^2} + \sum_{j=1}^r \|Z_j\varphi\|_{L^2}\right) \\ &\quad + \|\varphi\|_{L^\infty}\left(1 + \|f_{\epsilon_n,\delta_n}\|^2_{H^1_{\text{hyp},\delta_n}}\right).\end{aligned}$$

Combining with (10.43) and (10.19) it follows that

$$\left|\int Z_{0,\epsilon_n}(\chi w_n)\varphi\right| \lesssim \|\varphi\|_{L^\infty} + \|\varphi\|_{L^2} + \sum_{j=1}^r \|Z_j\varphi\|_{L^2} + \sqrt{\delta_n}\|\nabla\varphi\|_{L^2}. \tag{10.44}$$

In the notations $\tilde{\mathscr{X}}_\delta$ and $\tilde{\mathscr{X}}_\delta^*$ from (9.16), the bounds (10.43) and (10.44) together imply

$$\|\chi w_n\|_{\tilde{\mathscr{X}}_{\delta_n}} + \|\chi w_n\|_{\tilde{\mathscr{X}}_{\delta_n}^*} \lesssim 1 \tag{10.45}$$

uniformly in n. Applying Lemma 9.3 we conclude that for some $s > 0$,

$$\sup_{n\geq 1} \|\chi w_n\|_{H^s} \lesssim 1.$$

Therefore, by compact embedding (up to extracting another subsequence) $w_n \to w$ strongly in $L^p(B_R)$ for some $p > 2$. In particular, $w_n \to w$ in measure. Moreover, using (10.19), passing $n \to \infty$ in the sense of distributions in (10.42) we obtain that $w \in L^2(B_R)$ is a distributional solution to

$$N \cdot \nabla w = 0$$

on B_R. Convergence in measure and lower semicontinuity moreover provide

- $0 \leq w \leq 1$
- $|\{w = 0\} \cap B_R| \geq \alpha_1$
- $|\{w = 1\} \cap B_R| \geq \alpha_2$
- $|\{0 < w < 1\} \cap B_R| = 0.$

Now, by the Banach-Alaoglu theorem, lower semicontinuity, and (10.43) we have

$$Z_j w \in L^2(B_R), \quad 1 \leq j \leq r. \tag{10.46}$$

Since w is a characteristic function on B_R, (10.46) implies that $Z_j w = 0$ for $j = 1, \ldots, r$.[1] By Assumption 2, the collection $\{N, Z_1, \ldots, Z_r\}$ satisfies Hörmander's condition on $\mathbb{R}^d$. Since $\sum_{j=1}^r Z_j^2 w + Nw = 0$ in the sense of distributions, it follows from Hörmander's theorem that w is a smooth function. Thus, w must be constant, which contradicts the second and third bullets. This completes the proof of Lemma 10.3. □

10.3.2 Concluding the proof using the intermediate value lemma

In this section we prove Lemma 10.4.

We will need a regularized version of the function $z \to z_+$ that smooths out the kink at the origin in such a way so that the signed term that appears when passing solutions through the resulting convex function does not blow up too fast.

Lemma 10.5. *For all $\epsilon > 0$ $\exists \phi_\epsilon : \mathbb{R} \to \mathbb{R}$ that satisfies the following properties:*

- *ϕ_ϵ is smooth with $\|\phi_\epsilon''\|_{L^\infty(\mathbb{R})} \lesssim \epsilon^{-1/4}$*
- *$\phi_\epsilon'' \geq 0$*
- *$\phi_\epsilon(x) = x$ when $x \geq \epsilon^{1/4}$*
- *$\phi_\epsilon(x) = 0$ when $x \leq -\epsilon^{1/4}$*
- *$\phi_\epsilon(x)$ is nondecreasing with $\|\phi_\epsilon'\|_{L^\infty} \lesssim 1$ and $\phi_\epsilon(x) > 0$ for $x > -\epsilon^{1/4}$*

[1] Since Z_j is constant, by changing coordinates we may assume that $Z_j = (1, 0, 0, \ldots, 0)$. Denote $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and let $I(x_2, \ldots, x_d) = \int_{|x_1| < \sqrt{R^2 - x_2^2 - \ldots - x_d^2}} |\partial_{x_1} w(x_1, x_2, \ldots, x_d)|^2 dx_1$. Since w takes only the values 0 and 1, $I(x_2, \ldots, x_d) \in \{0, \infty\}$. Now, we have $\|Z_j w\|_{L^2(B_R)} = \int_{\sqrt{x_2^2 + \ldots + x_d^2} < R} I(x_2, \ldots, x_d) dx_2 \ldots dx_d < \infty$. Thus, $I(x_2, \ldots, x_d) < \infty$ for almost every $(x_2, \ldots, x_d)$ with respect the $d-1$ dimensional Lebesgue measure on $\{x_2^2 + \ldots + x_d^2 < R^2\}$ and we conclude that $\|Z_j w\|_{L^2(B_R)} = 0$.

Proof. Let $\varphi \in C_0^\infty([-\epsilon^{1/4}, \epsilon^{1/4}])$ be symmetric and satisfy $\varphi(x) > 0$ for $|x| < \epsilon^{1/4}$, $\int \varphi = 1$, and $\|\varphi\|_{L^\infty} \lesssim \epsilon^{-1/4}$. Define ϕ_ϵ to be the mollification

$$\phi_\epsilon(x) = \int_{-\epsilon^{1/4}}^{\epsilon^{1/4}} \varphi(y)(x+y)_+ dy. \tag{10.47}$$

From a straightforward calculation we see that

$$\phi_\epsilon(x) = \begin{cases} 0 & x \leq -\epsilon^{1/4} \\ x\int_{-x}^{\epsilon^{1/4}} \varphi(y)dy + \int_{-x}^{\epsilon^{1/4}} y\varphi(y)dy & -\epsilon^{1/4} < x < \epsilon^{1/4} \\ x & x \geq \epsilon^{1/4}. \end{cases} \tag{10.48}$$

The properties asserted above follow directly. □

Proof of Lemma 10.4. A byproduct of proving Lemmas 10.1 and 10.2 is that

$$\sup_{\epsilon,\delta \in (0,1)} \|f_{\epsilon,\delta}\|_{L^\infty(B_R)} \lesssim_R 1, \quad R > 0. \tag{10.49}$$

Combining this with $\int f_{\epsilon,\delta} = 1$ and (7.15), we see that there exist positive constants c_1, c_2, and R_0 independent of ϵ and δ such that

$$|\{f_{\epsilon,\delta} \geq c_1\} \cap B_{R_0}| \geq c_2. \tag{10.50}$$

Let $\tilde{f}_{\epsilon,\delta} = f_{\epsilon,\delta}/c_1$. For any $\theta \in (0,1)$ we define the sequence of functions

$$\tilde{w}_{k,\theta}^{\epsilon,\delta} = 1 - \left(\frac{4}{\theta}\right)^k \tilde{f}_{\epsilon,\delta}, \quad w_{k,\theta}^{\epsilon,\delta} = \phi_\epsilon(\tilde{w}_{k,\theta}^{\epsilon,\delta}), \tag{10.51}$$

where ϕ_ϵ is the function guaranteed by Lemma 10.5. When ϵ and δ are clear from context, we suppress them from the notation and simply write $w_{k,\theta}$. A direct consequence of the construction and (10.50) is that for any $\theta \in (0, 1/2)$, $\delta \in (0, 1)$, $\epsilon \in (0, 1/16)$, and $k, \ell \in \mathbb{N}$ there holds

$$|\{w_{k,\theta} = 0\} \cap B_{R_0}| \geq c_2, \tag{10.52}$$

$$\{w_{k+1,\theta} > 0\} \subseteq \{w_{k,\theta} \geq 1 - \theta\}, \tag{10.53}$$

$$\{0 < w_{k,\theta} < 1 - \theta\} \cap \{0 < w_{\ell,\theta} < 1 - \theta\} = \emptyset, \quad k \neq \ell. \tag{10.54}$$

Moreover, we have the following lemma, which says that for ϵ and θ fixed, the sequence $\{w_{k,\theta}\}$ satisfies the inequalities in Lemma 10.3 as long as k is not too large.

Lemma 10.6. *Let $\theta_* \in (0, 1)$, $k_* \in \mathbb{N}$, and $R > 0$. There exists $\epsilon_*(k_*, \theta_*, R)$ so that whenever $\epsilon \in (0, \epsilon_*)$, $\delta \in (0, 1)$, and $k \in \{1, \dots, k_*\}$ the following is satisfied pointwise for $|x| < 2R$:*

$$0 \leq \delta \Delta w_{k,\theta_*} + \frac{1}{\epsilon} L_\epsilon^* w_{k,\theta_*} \leq \frac{1}{\sqrt{\epsilon}} \left(1 + \delta |\nabla f_{\epsilon,\delta}|^2 + \sum_{j=1}^r |Z_j f_{\epsilon,\delta}|^2 \right). \tag{10.55}$$

Proof. A direct computation reveals that

$$\begin{aligned} \delta \Delta w_{k,\theta_*} + \frac{1}{\epsilon} L_\epsilon^* w_{k,\theta_*} = \phi_\epsilon''(\tilde{w}_{k,\theta_*}) \left(\frac{4}{\theta_*} \right)^{2k} \left(\delta |\nabla \tilde{f}_{\epsilon,\delta}|^2 + \sum_{j=1}^r |Z_j \tilde{f}_{\epsilon,\delta}|^2 \right) \\ + \text{Tr}(A) \phi_\epsilon'(\tilde{w}_{k,\theta_*}) \left(\frac{4}{\theta_*} \right)^k \tilde{f}_{\epsilon,\delta} + \text{Tr}(A) w_{k,\theta_*}. \end{aligned}$$

The lower bound in (10.55) is then immediate for any $k \in \mathbb{N}$ due to $\phi_\epsilon'', \phi_\epsilon' \geq 0$. As for the upper bound, by (10.49), $0 \leq w_{k,\theta_*} \leq 1$, $\|\phi''_\epsilon\|_{L^\infty} \lesssim \epsilon^{-1/4}$, and $\|\phi'_\epsilon\|_{L^\infty} \lesssim 1$, there exists a constant $C(R)$ such that for any $k \leq k_*$ we have

$$\delta \Delta w_{k,\theta_*} + \frac{1}{\epsilon} L^*_\epsilon w_{k,\theta_*} \leq C(R)\epsilon^{-1/4} \left(\frac{4}{\theta_*}\right)^{2k_*} \left(1 + \delta|\nabla f_{\epsilon,\delta}|^2 + \sum_{j=1}^{r} |Z_j f_{\epsilon,\delta}|^2\right).$$

Choosing

$$\epsilon_*^{1/4} \leq \frac{1}{C(R)} \left(\frac{\theta_*}{4}\right)^{2k_*}$$

yields (10.55). □

The main application of Lemma 10.3 is the following.

Lemma 10.7. *Let $R \geq R_0$, $\kappa > 0$, and $\delta \in (0,1)$. There exists $\theta_* \in (0,1/2)$, $\epsilon_* \in (0,1/16)$, and $K \in \mathbb{N}$, all depending only on κ and R, so that whenever $\epsilon \in (0,\epsilon_*)$ there exists $k_* \in \mathbb{N}$ with $k_* \leq K$ such that*

$$\int_{B_R} |w_{k_*,\theta_*}|^2 \leq \kappa. \tag{10.56}$$

Proof. Let $\epsilon_0 > 0$, $\theta_* \in (0,1/2)$, and $\mu > 0$ denote the parameters guaranteed by applying Lemma 10.3 at radius R with $\alpha_1 = c_2$ and $\alpha_2 = \kappa$. This fixes θ_* from the lemma statement. Note that since c_2 is universal, ϵ_0, θ_*, and μ depend only on κ and R.

Let K be the first natural number that exceeds $1 + 2|B_R|/\mu$ and observe that K depends only on R and κ. By Lemma 10.6 there exists $\bar{\epsilon}(K,\theta_*,R) < 1/16$ such that (10.55) holds whenever $\epsilon \in (0,\bar{\epsilon})$ and $k \leq K$. Let $\epsilon_* = \min(\epsilon_0,\bar{\epsilon})$. To complete the proof, it suffices to show that for every $\epsilon \in (0, \epsilon_*)$ there exists $k_* \leq K$ such that

$$\int_{B_R} |w_{k_*,\theta_*}|^2 \leq \kappa.$$

If this is not the case, then there exists $\epsilon' \in (0, \epsilon_*)$ such that

$$\int_{B_R} |w_{k,\theta_*}^{\epsilon',\delta}|^2 > \kappa \tag{10.57}$$

for all $k \leq K$. In the remainder of this proof we write $w_{k,\theta_*} = w_{k,\theta_*}^{\epsilon',\delta}$. Since $0 \leq w_{k,\theta_*} \leq 1$, it follows from (10.53) and (10.57) that for every $k \leq K-1$ we have

$$|\{w_{k,\theta_*} \geq 1-\theta_*\} \cap B_R| \geq \int_{B_R} |w_{k+1,\theta_*}|^2 \geq \kappa. \tag{10.58}$$

Combining with (10.52) and (10.55), we see that for every $k \leq K-1$, the function w_{k,θ_*} satisfies the hypotheses of Lemma 10.3 at radius R with $\alpha_1 = c_2$ and $\alpha_2 = \kappa$. Since $\epsilon' < \epsilon_0$, we obtain that for every $k \leq K-1$ there holds

$$|\{0 < w_{k,\theta_*} < 1-\theta_*\} \cap B_R| \geq \mu, \tag{10.59}$$

which along with (10.54) implies that $|B_R| \geq (K-1)\mu \geq 2|B_R|$, a contradiction. □

We are now ready to complete the proof of Lemma 10.4. For $R \geq R_0$ and $\kappa(R)$ to be chosen sufficiently small we apply Lemma 10.7 to obtain ϵ_*, θ_*, and K, all depending only on R, so that whenever $\epsilon \in (0, \epsilon_*)$ there exists $k_* \leq K$ such that

$$\int_{B_R} \left| w_{k_*,\theta_*}^{\epsilon,\delta} \right|^2 \leq \kappa. \tag{10.60}$$

From the lower bound in Lemma 10.6 (which $w_{k,\theta_*}^{\epsilon,\delta}$ satisfies for any $\epsilon, \delta \in (0,1)$ and $k \in \mathbb{N}$) it follows by Lemma 10.2 that there exists a constant $C(R)$ such that

$$\|w_{k_*,\theta_*}\|_{L^\infty(B_{R/2})} \leq C(R)\|w_{k_*,\theta_*}\|_{L^2(B_R)} \leq C(R)\sqrt{\kappa}. \tag{10.61}$$

Let κ be small enough so that $C(R)\sqrt{\kappa} \leq 1/2$. Since ϕ_ϵ is monotone increasing with $\phi(1/2) = 1/2$ this implies that $\sup_{|x|\leq R/2} \tilde{w}_{\theta_*,k_*}(x) \leq 1/2$. Directly from (10.51) we get that

$$\inf_{\epsilon\in(0,\epsilon_*)} \inf_{\delta\in(0,1)} \inf_{|x|\leq R/2} f_{\epsilon,\delta}(x) \geq \frac{c_1}{2}\left(\frac{\theta_*}{4}\right)^K \gtrsim_R 1. \tag{10.62}$$

The bound (10.41) then follows from Lemma A.2. If Assumption 3 is satisfied, then Lemma A.3 and (10.41) together yield (6.13), which completes the proof. □

Remark 30. Under Assumption 3, the arguments of this section yield

$$\inf_{\epsilon,\delta\in(0,1),|x|\leq R} f_{\epsilon,\delta}(x) \gtrsim_R 1 \tag{10.63}$$

for every $R > 0$. Indeed, this follows immediately from (10.62) and Lemma A.3.

10.4 Global bounds from local ones

The purpose of this section is to upgrade (10.1) to the Gaussian upper bound (6.12). By Lemma A.2 it suffices to prove the following.

Lemma 10.8. *There exists $\lambda > 0$ so that*

$$\sup_{\epsilon,\delta\in(0,1)} f_{\epsilon,\delta}(x) \lesssim e^{-\lambda|x|^2/2}. \tag{10.64}$$

Proof. Let $G_\lambda(x) = \exp(-\lambda|x|^2/2)$ for $\lambda > 0$. Note that $Bx \cdot \nabla G_\lambda = N \cdot \nabla G_\lambda = 0$ because G_λ is radially symmetric. Hence, denoting $Z_j = (Z_j^{(1)}, \dots, Z_j^{(r)})$, we have

$$\begin{aligned}(L_\epsilon^* + \epsilon\delta\Delta)G_\lambda(x) = &\left(\lambda^2 \sum_{j=1}^r |Z_j \cdot x|^2 - \lambda \sum_{k,j} (Z_j^{(k)})^2 + \mathrm{Tr}(A) - \lambda Ax \cdot x\right) \epsilon G_\lambda(x) \\ &+ \left(\lambda^2 |x|^2 - d\lambda\right) \epsilon\delta G_\lambda(x),\end{aligned} \tag{10.65}$$

where recall that d denotes the dimension of the space. Since A is positive definite, there exists $\lambda_0 > 0$ sufficiently small and $R_0 \geq 1$ sufficiently large, both depending only on A, $\{Z_j\}_{j=1}^r$, and d, so that $(L_\epsilon^* + \epsilon\delta\Delta)G_{\lambda_0}(x) < 0$ whenever $|x| \geq R_0$. With R_0 fixed, we have from (10.32) that there exists a constant $C_0(A, \{Z_j\}_{j=1}^r, d)$ such that

$$\sup_{\epsilon,\delta\in(0,1)} \|f_{\epsilon,\delta}\|_{L^\infty(B_{2R_0})} \leq C_0. \tag{10.66}$$

We then define the upper barrier function

$$G^+(x) = 2C_0 e^{4\lambda_0 R_0^2/2} G_{\lambda_0}(x). \tag{10.67}$$

To prove the lemma, it suffices to show that for all $x \in \mathbb{R}^d$ there holds

$$\sup_{\epsilon,\delta\in(0,1)} f_{\epsilon,\delta}(x) \lesssim G^+(x). \tag{10.68}$$

Let $\chi \in C_0^\infty(B_{2R_0})$ be a smooth cutoff function with $\chi(x) = 1$ for $|x| \leq R_0$. Then, define $g = \chi f_{\epsilon,\delta}$. Note that by construction we have

$$f_{\epsilon,\delta}(x) \leq g(x) \quad \forall |x| \leq R_0, \tag{10.69}$$

$$g(x) < \min\left(G^+(x), \frac{3}{2} f_{\epsilon,\delta}(x)\right) \quad \forall x \in \mathbb{R}^d. \tag{10.70}$$

Due to (10.69) and (7.15) we can assume that R_0 is large enough so that

$$\int g \geq \frac{1}{2}. \tag{10.71}$$

As in Chapter 7, let $\mathcal{P}_t^{\epsilon,\delta}$ denote the Markov semigroup generated by $L_\epsilon + \epsilon\delta\Delta$. Let μ be the measure on $\mathbb{R}^d$ with density $(\int g)^{-1} g$ and let $\tilde{g}_t : \mathbb{R}^d \to \mathbb{R}$ denote the density of $(\mathcal{P}_t^{\epsilon,\delta})^* \mu$. Then, $g_t := (\int g)\tilde{g}_t$ is the unique global smooth solution to the problem

$$\begin{cases} \partial_t g_t = L_\epsilon^* g_t + \epsilon\delta\Delta g_t, \\ g_0 = g(x). \end{cases} \tag{10.72}$$

Define

$$t_* = \inf\{t \geq 0 : \text{there exists } x \text{ such that } g_t(x) = G^+(x)\}$$

with the convention that $t_* = \infty$ if $g_t(x) < G^+(x)$ for every $x \in \mathbb{R}^d$ and $t > 0$. By the convergence

$$\limsup_{t\to\infty} \|\tilde{g}_t - f_{\epsilon,\delta}\|_{L^1} \leq \limsup_{t\to\infty} \|(\mathcal{P}_t^{\epsilon,\delta})^*\mu - \mu_{\epsilon,\delta}\|_{TV} = 0$$

for any $\epsilon, \delta > 0$ and (10.71), there exists a sequence $\{t_k\}_{k=1}^{\infty}$ such that

$$\lim_{k\to\infty} g_{t_k}(x) = \left(\int g\right) f_{\epsilon,\delta}(x) \geq \frac{1}{2} f_{\epsilon,\delta}(x)$$

for almost every $x \in \mathbb{R}^d$. Thus, since $f_{\epsilon,\delta}$ is smooth, to prove (10.68) it is enough to show that $t_* = \infty$.

As in the proof of Lemma A.1, it can be shown that $\forall k > 0$, $\exists \gamma_k > 0$ that does not depend on ϵ or δ such that for all $T < \infty$ and $\delta > 0$,

$$\sup_{t\in[0,T]} \|e^{\gamma_k |x|^2} g_t\|_{H^k} \lesssim_{T,\delta,k} 1.$$

Directly from the equation, the same estimate holds for $\partial_t g_t$ provided that γ_k is replaced with $\gamma_k/2$, and so $e^{\frac{\gamma_k}{2}|x|^2} g_t$ takes values continuously in H^k. In particular, there is some $\gamma' > 0$ sufficiently small that does not depend on ϵ or δ so that $e^{\gamma'|x|^2} g_t \in C([0,T]; L^\infty)$ for any $T > 0$. Hence, if $\lambda_0 < \gamma'$ and $t_* < \infty$ there exists a "first crossing time" $t_* > 0$, i.e., (t_*, x_*) is such that $g_{t_*}(x_*) = G^+(x_*)$ and $g_t(x) \leq G^+(x)$ for all $t \leq t_*$ and $x \in \mathbb{R}^d$. Suppose for the sake of contradiction that $t_* < \infty$. We have two cases.

Case 1: $|x_*| < R_0$: By (10.70) and the fact $(\mathcal{P}_t^{\epsilon,\delta})^*$ preserves positivity, we have

$g_t(x) \leq (3/2)f_{\epsilon,\delta}$ for all $t \geq 0$. Combining with (10.66) we obtain

$$\sup_{t\geq 0, |x|<R_0} g_t(x) \leq \frac{3C_0}{2}.$$

Since $G^+(x) \geq 2C_0$ whenever $|x| \leq 2R_0$ we conclude that $|x_*| < R$ is impossible.

Case 2: $|x_*| \geq R_0$: Since $g \in C_t^1 C_x^2((0,\infty) \times \mathbb{R}^d)$, it follows from (10.70) and a classical barrier function argument that $(L_\epsilon^* + \epsilon\delta\Delta)G^+(x_*) \geq 0$. This is a contradiction because we chose λ_0 and R_0 so that $(L_\epsilon^* + \epsilon\delta\Delta)G^+(x) < 0$ whenever $|x| \geq R_0$. $\square$

Chapter 11: Geometric Ergodicity

The main goal of this chapter is to carry out the details of the proof outline of Theorem 6.5 given in Chapter 8. This consists of proving Lemmas 8.4, 8.3, and 6.3. We will also prove, in Section 11.3, the optimality result Theorem 6.7.

11.1 $L^\infty \to L^2_{\mu_\epsilon}$ decay for $\mathcal{P}_t^\epsilon$

In this section we will prove Lemmas 8.4 and 8.3.

11.1.1 Proof of Lemma 8.4

The first step in the proof of Lemma 8.4 is to prove the following general functional inequality, which can be interpreted as a type of hypoelliptic weak Poincaré inequality. Recall the notations defined in (9.28) and (9.29).

Lemma 11.1. *Let $R > 0$. For every $\delta > 0$ there exists a constant $C_{\delta,R}$ such that for all*

$\epsilon \in [0,1]$, $t_0 \geq 0$, and $f \in C^\infty((t_0, t_0+1) \times B_{R+1})$ there holds

$$\begin{aligned}\|f - \bar{f}\|_{L^2((t_0+1/4, t_0+3/4)\times B_R)} \leq \delta \|f\|_{L^\infty((t_0,t_0+1)\times B_{R+1})} \\ + C_{\delta,R} \sum_{j=1}^{r} \|Z_j f\|_{L^2((t_0,t_0+1)\times B_{R+1})} \\ + C_{\delta,R} \|(\partial_t + Z_{0,\epsilon}) f\|_{L^2((t_0,t_0+1),\mathscr{X}^*(B_{R+1}))},\end{aligned}$$

where

$$\bar{f} = 2|B_R|^{-1} \int_{t_0+1/4}^{t_0+3/4} \int_{B_R} f$$

is the average value of f on $(t_0+1/4, t_0+3/4) \times B_R$.

We will prove the lemma above with a compactness-rigidity argument that is in the same spirit as the compactness proof of the classical Poincaré inequality $\|f\|_{L^2} \leq CR\|\nabla f\|_{L^2}$. As one might expect, the compactness will be obtained with a Hörmander inequality, in particular Lemma 9.3. The rigidity step is more subtle. The result that we need is stated as follows.

Lemma 11.2 (Rigidity lemma)**.** *Let $\Omega \subset \mathbb{R}^d$ be an open, connected set. Suppose that $\{X_j\}_{j=0}^k \subseteq T(\Omega)$ satisfies Hörmander's condition on Ω. Let $f \in L^2_{loc}(\mathbb{R}^d)$ be a distributional solution to*

$$X_0 f = 0.$$

If $X_j f = 0$ for each $j = 1, \ldots, k$, then f is constant on Ω.

Proof. The assumptions of the lemma imply that $f \in L^2_{\text{loc}}(\Omega)$ is a distributional solution to $\sum_{j=1}^k X_j^2 f + X_0 f = 0$, and so by Hörmander's theorem [Theorem 1.1, [79]] it follows that $f \in C^\infty(\Omega)$. Thus, it suffices to prove the result for smooth functions satisfying $X_j f = 0$ for $j = 0, 1, \ldots, k$ in the classical sense on Ω.

Let $B(x, r)$ denote the ball of radius r centered at x with the usual Euclidean metric, and for $c \in \mathbb{R}$, let

$$S_c = \{x \in \Omega : \exists r > 0 \text{ such that } f \equiv c \text{ on } B(x, r)\}.$$

We will prove that there is some c so that S_c is open, nonempty, and relatively closed in Ω. The fact that S_c is open for each c follows by its definition. To prove that $\exists c$ such that S_c is both relatively closed and nonempty, it suffices to show that for every $x \in \Omega$ $\exists r_x > 0$ so that f is constant on $B(x, r_x)$. Fix $x_0 \in \Omega$ and let $\mathcal{U} \subset \Omega$ be an open ball containing x_0. Define the *$\mathcal{U}$-reachable set at x_0* to be the points $x_1 \in \mathcal{U}$ such that there exist bounded, measurable functions $\{c_j : [0, 1] \to \mathbb{R}\}_{j=0}^k$ and a curve $\gamma : [0, 1] \to \mathcal{U}$ such that $\gamma(0) = x_0$, $\gamma(1) = x_1$, and

$$\gamma'(t) = \sum_{j=0}^{k} c_j(t) X_j(\gamma(t)) \quad \text{a.e. } t \in [0, 1]. \tag{11.1}$$

A well-known fact in the theory of local controllability is that the $\mathcal{U}$-reachable set at x_0 is an open neighborhood of x_0 as soon as $\{X_j\}_{j=0}^k$ satisfies Hörmander's condition on $\mathcal{U}$; see e.g. [Theorem 2.2, [75]]. Since f is constant along any curve γ satisfying (11.1) we conclude that it must be constant on some open ball containing x_0, completing the proof. □

We are now ready to prove Lemma 11.1.

Proof of Lemma 11.1. It suffices to prove the inequality for $t_0 = 0$. Suppose for the sake of contradiction that the result is false. Then, there exists $\delta > 0$ and a sequence $\{(f_n, \epsilon_n)\}_{n=1}^{\infty} \subseteq C^\infty((0,1) \times B_{R+1}) \times [0,1]$ such that

$$\begin{aligned} \|f_n - \bar{f}_n\|_{L^2((1/4,3/4)\times B_R)} \geq \delta \|f_n\|_{L^\infty((0,1)\times B_{R+1})} + n \sum_{j=1}^{r} \|Z_j f_n\|_{L^2((0,1)\times B_{R+1})} \\ + n\|(\partial_t + Z_{0,\epsilon_n}) f_n\|_{L^2((0,1);\mathscr{X}^*(B_{R+1}))} \end{aligned} \tag{11.2}$$

for every $n \in \mathbb{N}$. Let

$$g_n = \frac{f_n - \bar{f}_n}{\|f_n - \bar{f}_n\|_{L^2((1/4,3/4)\times B_R)}}.$$

Dividing (11.2) by $\|f_n - \bar{f}_n\|_{L^2((1/4,3/4)\times B_R)}$ and using that

$$\|g_n\|_{L^\infty} \leq \frac{2}{\|f_n - \bar{f}_n\|_{L^2((1/4,3/4)\times B_R)}} \|f_n\|_{L^\infty((0,1)\times B_{R+1})}$$

we obtain

$$\begin{aligned} 1 \geq \frac{\delta}{2} \|g_n\|_{L^\infty((0,1)\times B_{R+1})} + n \sum_{j=1}^{r} \|Z_j g_n\|_{L^2((0,1)\times B_{R+1})} \\ + n\|(\partial_t + Z_{0,\epsilon_n}) g_n\|_{L^2((0,1);\mathscr{X}^*(B_{R+1}))}. \end{aligned} \tag{11.3}$$

Let $\chi \in C_0^\infty((0,1) \times B_{R+1})$ be a smooth cutoff function with $0 \leq \chi \leq 1$ and $\chi \equiv 1$ on $(1/8, 7/8) \times B_{R+1/2})$. From (11.3) it follows readily that

$$\begin{aligned} \|\chi g_n\|_{L^2((0,1)\times B_{R+1})} + \sum_{j=1}^{r} \|Z_j(\chi g_n)\|_{L^2((0,1)\times B_{R+1})} \\ + \|(\partial_t + Z_{0,\epsilon_n})(\chi g_n)\|_{L^2((0,1);\mathscr{X}^*(B_{R+1}))} \lesssim \delta^{-1}. \end{aligned} \tag{11.4}$$

By Assumption 2, $\{\partial_t + Z_{0,\epsilon_n}, Z_1, \ldots Z_r\}$ satisfies the uniform Hörmander condition on

$(0,1) \times B_{R+1}$ with constants that do not depend on ϵ_n (depending on R however). Thus, Lemma 9.3 implies that there exists $s > 0$ such that

$$\sup_{n \in \mathbb{N}} \|\chi g_n\|_{H^s(\mathbb{R} \times \mathbb{R}^d)} \lesssim \delta^{-1}.$$

By compact embedding there exists $g_\infty \in L^2((1/8, 7/8) \times B_{R+1/2})$ such that (up to a subsequence that we do not relabel) $g_n \to g_\infty$ strongly in $L^2((1/8, 7/8) \times B_{R+1/2})$. Moreover,

$$\int_{1/4}^{3/4} \int_{B_R} g_\infty = 0, \tag{11.5}$$

$$\sum_{j=1}^{r} \int_{1/8}^{7/8} \int_{B_{R+1/2}} |Z_j g_\infty|^2 = 0, \tag{11.6}$$

$$\int_{1/4}^{3/4} \int_{B_R} |g_\infty|^2 = 1. \tag{11.7}$$

By extracting a subsequence can ensure that $\epsilon_n \to \epsilon_\infty \in [0, 1]$. Since any function $\varphi \in C_0^\infty((1/8, 7/8) \times B_{R+1/2})$ can be extended by zero to a function $\varphi \in C_0^\infty((0, 1) \times B_{R+1})$ with $\|\varphi\|_{\mathscr{X}} < \infty$, we must have $(\partial_t + Z_{0,\epsilon_\infty}) g_\infty = 0$ in the sense of distributions on $(1/8, 7/8) \times B_{R+1/2}$ by (11.3). Thus, due to Lemma 11.2 we have that g_∞ is constant on $(1/8, 7/8) \times B_{R+1/2}$, which contradicts the combination of (11.5) and (11.7). □

With the proof of Lemma 11.1 complete we are now in a position to prove Lemma 8.4, which we restate here for convenience.

Lemma 11.3. *Suppose that Assumption 3 holds. Then, there exists a nonincreasing function* $\beta : (0, \infty) \to [1, \infty)$ *so that for every* $s > 0$ *and bounded, measurable* f :

$\mathbb{R}^d \to \mathbb{R}$ *there holds, uniformly in* $t_0 \geq 1$ *and* $\epsilon \in (0,1)$,

$$\int_{t_0+1/4}^{t_0+3/4} \int |\mathcal{P}_t^\epsilon f - \mu_\epsilon(f)|^2 d\mu_\epsilon dt \leq \beta(s) \sum_{j=1}^{r} \int_{t_0}^{t_0+1} \int |Z_j \mathcal{P}_t^\epsilon f|^2 d\mu_\epsilon dt + s\|f - \mu_\epsilon(f)\|_{L^\infty}^2.$$

Remark 31. Our proof gives no estimate on the rate at which $\beta(s)$ blows up as $s \to 0$. Instead, $\beta(s) \geq C_{\delta(s),R(s)}$ for $C_{\delta(s),R(s)}$ obtained by applying Lemma 11.1 with $\delta(s), R(s) > 0$ satisfying $\lim_{s\to 0} R(s) = \infty$ and $\lim_{s\to 0} \delta(s) = 0$.

Proof. For simplicity we omit the ϵ dependence in the notation.

If $s \geq 1/2$ then the claimed inequality is trivial. Fix $s < 1/2$ and let $g(t) = \mathcal{P}_t f - \mu(f)$. By the moment bound (7.13), there exists $R(s)$ sufficiently large so that $\mu(B_R^c) \leq s/2$ uniformly in ϵ. Using that $\mathcal{P}_t$ propagates L^∞ bounds we have then

$$\int_{t_0+1/4}^{t_0+3/4} \int |g(t)|^2 d\mu dt \leq \int_{t_0+1/4}^{t_0+3/4} \int_{|x|\leq R} |g(t)|^2 d\mu dt + \frac{s}{4}\|f - \mu(f)\|_{L^\infty}^2. \tag{11.8}$$

The goal now is to bound the first term on the right-hand side of (11.8). Let

$$\mu_R(C) = \frac{\mu(C \cap B_R)}{\mu(B_R)}, \quad C \in \mathcal{B}(\mathbb{R}^d)$$

and

$$g_R = 2\int_{t_0+1/4}^{t_0+3/4} \int g(t) d\mu_R dt.$$

By adding and subtracting g_R we have

$$\begin{aligned}\int_{t_0+1/4}^{t_0+3/4}\int_{|x|\le R}|g(t)|^2 d\mu dt &= \mu(B_R)\int_{t_0+1/4}^{t_0+3/4}\int_{|x|\le R}|g(t)|^2 d\mu_R dt\\ &\le 2\mu(B_R)\int_{t_0+1/4}^{t_0+3/4}\int |g(t)-g_R|^2\, d\mu_R dt + \mu(B_R)|g_R|^2.\end{aligned}$$

Now, for each $t \ge 0$ we have $\int g(t) d\mu = 0$, and so

$$\begin{aligned}|g_R|^2 &= \frac{4}{\mu(B_R)^2}\left(\int_{t_0+1/4}^{t_0+3/4}\int_{|x|>R} g(t) d\mu dt\right)^2\\ &\le \left(\frac{\mu(B_R^c)}{\mu(B_R)}\right)^2 \|f-\mu(f)\|_{L^\infty}^2\\ &\le \frac{s}{4}\|f-\mu(f)\|_{L^\infty}^2,\end{aligned}$$

where in the last inequality we used that

$$\left(\frac{\mu(B_R^c)}{\mu(B_R)}\right)^2 \le \left(\frac{s/2}{3/4}\right)^2 = \frac{4s^2}{9} \le \frac{s}{4}.$$

Combining our estimates thus far and using that $\mathrm{Var}_\nu h \le \mathbb{E}_\nu(h-c)^2$ for any measure $\nu \in \mathcal{M}(\mathbb{R}^d)$, $c \in \mathbb{R}$, and $h \in L^2_\nu$ we obtain

$$\begin{aligned}\int_{t_0+1/4}^{t_0+3/4}\int_{|x|\le R}|g(t)|^2 d\mu dt \le 2\mu(B_R)\int_{t_0+1/4}^{t_0+3/4}\int |g(t)-\bar{g}_R|^2\, d\mu_R dt\\ + \frac{s}{4}\|f-\mu(f)\|_{L^\infty}^2,\end{aligned} \tag{11.9}$$

where we have introduced

$$\bar{g}_R = \frac{2}{|B_R|}\int_{t_0+1/4}^{t_0+3/4}\int_{|x|\le R} g(t)dxdt.$$

Using (6.12), there exists a constant $c > 0$ that does not depend on ϵ such that

$$\begin{aligned} 2\mu(B_R)\int_{t_0+1/4}^{t_0+3/4}\int_{|x|\le R}|g(t)-\bar{g}_R|^2 d\mu_R dt \\ \le c\int_{t_0+1/4}^{t_0+3/4}\int_{|x|\le R}|g(t)-\bar{g}_R|^2 dxdt. \end{aligned} \tag{11.10}$$

By Lemma 11.1 applied with $\delta = \sqrt{s/(8c)}$ and the fact that

$$\partial_t g + Z_{0,\epsilon} g = \epsilon\sum_{j=1}^{r} Z_j^2(\mathcal{P}_t f),$$

there exist constants C_s and C_s' such that

$$\begin{aligned} &c\int_{t_0+1/4}^{t_0+3/4}\int_{|x|\le R}|g(t)-\bar{g}_R|^2 dxdt \\ &\le C_s\sum_{j=1}^{r}\int_{t_0}^{t_0+1}\int_{|x|\le R+1}|Z_j\mathcal{P}_t f|^2 dxdt + \frac{s}{2}\|f-\mu(f)\|_{L^\infty}^2 \\ &+\epsilon^2 C_s \sup_{\varphi\in C_0^\infty(I_{t_0}\times B_{R+1}),\|\varphi\|_{L^2(I_{t_0};\mathscr{X})}\le 1}\left|\sum_{j=1}^{r}\int_{t_0}^{t_0+1}\int_{|x|\le R+1}\varphi Z_j^2(\mathcal{P}_t f)dxdt\right|^2 \\ &\le C_s'\sum_{j=1}^{r}\int_{t_0}^{t_0+1}\int_{|x|\le R+1}|Z_j\mathcal{P}_t f|^2 dxdt + \frac{s}{2}\|f-\mu(f)\|_{L^\infty}^2, \end{aligned} \tag{11.11}$$

where we have written $I_{t_0} = (t_0, t_0+1)$. From (6.13) and the fact that R depends only s

we have

$$C'_s \sum_{j=1}^{r} \int_{t_0}^{t_0+1} \int_{|x|\le R+1} |Z_j \mathcal{P}_t f|^2 dx dt \le C''_s \sum_{j=1}^{r} \int_{t_0}^{t_0+1} \int_{|x|\le R+1} |Z_j \mathcal{P}_t f|^2 d\mu dt,$$

which along with (11.9), (11.10), and (11.11) yields

$$\begin{aligned} \int_{t_0+1/4}^{t_0+3/4} \int_{|x|\le R} |g(t)|^2 d\mu dt \le C''_s \sum_{j=1}^{r} \int_{t_0}^{t_0+1} \int_{|x|\le R+1} |Z_j \mathcal{P}_t f|^2 d\mu dt \\ + \frac{3s}{4} \|f - \mu(f)\|^2_{L^\infty}. \end{aligned} \tag{11.12}$$

Combining (11.12) and (11.8) completes the proof. □

We now conclude the section by using Lemma 8.4 to prove Lemma 8.3.

Proof of Lemma 8.3. Let $f : \mathbb{R}^d \to \mathbb{R}$ be bounded and Borel measurable. Using the identity

$$g L_\epsilon g = \frac{1}{2} L_\epsilon (g^2) - \epsilon \sum_{j=1}^{r} |Z_j g|^2$$

for any $g \in C^2(\mathbb{R}^d)$, $L^*_\epsilon f_\epsilon = 0$, and $\partial_t(\mathcal{P}^\epsilon_t f) = L_\epsilon(\mathcal{P}^\epsilon_t f - \mu_\epsilon(f))$ we obtain

$$\frac{1}{2}\frac{d}{d\tau} \int |\mathcal{P}^\epsilon_\tau f - \mu_\epsilon(f)|^2 d\mu_\epsilon = -\epsilon \sum_{j=1}^{r} \int |Z_j \mathcal{P}^\epsilon_\tau f|^2 d\mu_\epsilon. \tag{11.13}$$

Integrating over $\tau \in (t, t+1)$ for $t \ge 1$ and using Lemma 8.4 gives, for any $s > 0$,

$$\begin{aligned} \|\mathcal{P}^\epsilon_{t+1} f - \mu_\epsilon(f)\|^2_{L^2_{\mu_\epsilon}} - \|\mathcal{P}^\epsilon_t f - \mu_\epsilon(f)\|^2_{L^2_{\mu_\epsilon}} \le -\frac{2\epsilon}{\beta(s)} \int_{t+1/4}^{t+3/4} \|\mathcal{P}^\epsilon_\tau f - \mu_\epsilon(f)\|^2_{L^2_{\mu_\epsilon}} d\tau \\ + \frac{2\epsilon s}{\beta(s)} \|f - \mu_\epsilon(f)\|^2_{L^\infty}. \end{aligned} \tag{11.14}$$

Let $E(t) = \|\mathcal{P}_t^\epsilon f - \mu_\epsilon(f)\|^2_{L^2_{\mu_\epsilon}}$. By (11.13), the energy $E(t)$ is nonincreasing, and so the previous estimate implies

$$E(t+1) \leq \left(1 + \frac{\epsilon}{\beta(s)}\right)^{-1} E(t) + \left(1 + \frac{\epsilon}{\beta(s)}\right)^{-1} \frac{2\epsilon s}{\beta(s)} \|f - \mu_\epsilon(f)\|^2_{L^\infty} \quad \forall\, t \geq 1. \tag{11.15}$$

Iterating over $t = 1, 2, \ldots, n-1$ yields

$$E(n) \leq \left(1 + \frac{\epsilon}{\beta(s)}\right)^{-(n-1)} E(1) + \frac{2\epsilon s}{\beta(s)} \|f - \mu_\epsilon(f)\|^2_{L^\infty} \sum_{k=1}^{n-1} \left(1 + \frac{\epsilon}{\beta(s)}\right)^{-k} \tag{11.16}$$

$$\leq \left(1 + \frac{\epsilon}{\beta(s)}\right)^{-(n-1)} \|f - \mu_\epsilon(f)\|^2_{L^\infty} + 2s\|f - \mu_\epsilon(f)\|^2_{L^\infty}. \tag{11.17}$$

This implies that there exists a universal constant $\delta > 0$ such that for all $\epsilon \in (0, 1)$ there holds

$$E(n) \leq e^{-n\delta\epsilon/\beta(s)} \|f - \mu_\epsilon(f)\|^2_{L^\infty} + 2s\|f - \mu_\epsilon(f)\|^2_{L^\infty}, \quad \mathbb{N} \ni n \geq 2. \tag{11.18}$$

Let $\bar{\psi} : [0, \infty) \to (0, 1]$ be defined by

$$\bar{\psi}(t) = \inf\{s > 0 : e^{-t\delta/\beta(s)} \leq s\}. \tag{11.19}$$

It is clear that $\bar{\psi}$ is non-increasing with $\lim_{t\to\infty} \bar{\psi}(t) = 0$, and moreover by (11.18) we have proven that

$$E(n) \leq 3\bar{\psi}(n\epsilon)\|f - \mu_\epsilon(f)\|^2_{L^\infty}, \quad \mathbb{N} \ni n \geq 2.$$

Because $E(t)$ is nonincreasing it follows that

$$\|\mathcal{P}_t^\epsilon f - \mu_\epsilon(f)\|_{L^2_{\mu_\epsilon}}^2 \leq 3\bar{\psi}(\lfloor t \rfloor \epsilon)\|f - \mu_\epsilon(f)\|_{L^\infty}^2 \leq 3\bar{\psi}(\epsilon t - 1)\|f - \mu_\epsilon(f)\|_{L^\infty}^2, \quad t \geq 2.$$

Since $\|\mathcal{P}_t^\epsilon f - \mu_\epsilon(f)\|_{L^2_{\mu_\epsilon}} \leq \|f - \mu_\epsilon(f)\|_{L^\infty}$ for any $t \geq 0$, the proof of (8.19) is complete by setting

$$\psi(t) = \begin{cases} 3 & t \leq 2, \\ 3\bar{\psi}(t-1) & t > 2. \end{cases}$$

□

11.2 $L^2 \to L^\infty$ regularization for $\mathcal{P}_t$

In this section we prove Lemma 6.3, which proceeds by a parabolic version of the arguments in Section 10.2.

Proof of Lemma 6.3. Since $\mathcal{P}_t^\epsilon$ is strong Feller (see Lemma 7.3), it follows from the semigroup property and the monotonicity (11.13) that it suffices to prove the result for continuous f. As in earlier calculations, it is convenient to regularize the problem with $\delta\Delta$ and pass to the limit. Let $\tilde{\mathcal{P}}_t^{\epsilon,\delta}$ denote the Markov semigroup generated by $\delta\Delta + \epsilon^{-1}L_\epsilon$ and as before write $\mu_{\epsilon,\delta}$ for its unique invariant measure. For $k \geq 0$, define $R_k = R(1 + 2^{-k})$ and $t_k = \frac{1}{4} - 2^{-k-3}$. Let $\alpha \in (1, 4)$ be such that with s given as in Lemma 9.5 there holds

$$||g||_{L_t^{2\alpha} L_x^{2\alpha}} \lesssim ||g||_{L_t^\infty L_x^2} + ||g||_{L_t^2 H_x^s}, \tag{11.20}$$

and for $k \geq 0$ define $w_k = (\tilde{\mathcal{P}}_t^{\epsilon,\delta} f)^{\alpha^k}$. We will show that $\exists C > 0$ such that for every $k \geq 0$ there holds

$$\|w_k\|_{L_t^{2\alpha} L_x^{2\alpha}((t_{k+1}, 2-t_{k+1}) \times B_{R_{k+1}})} \leq C^k \|w_k\|_{L_t^2 L_x^2((t_k, 2-t_k) \times B_{R_k})}. \tag{11.21}$$

Let $\chi_k \in C_0^\infty((t_k, 2-t_k) \times B_{R_k})$ be a time-dependent, radially-symmetric in space, smooth cutoff function satisfying $\chi_k(t,x) = 1$ for $|x| \leq R_{k+1}$ and $2 - t_{k+1} \geq t \geq t_{k+1}$. Moreover, we may choose χ_k so that $|\partial_t \chi_k| \lesssim 2^k$ and $|D_x^\beta \chi| \lesssim R^{-1} 2^{|\beta| k}$ for every multi-index with $|\beta| \leq 2$. Let $v_k = \chi_k w_k$. By splitting $f = \max(f,0) - \max(-f,0)$ and regularizing with a small constant we may assume without loss of generality that $f > 0$. From the convexity and smoothness of $z \mapsto z^\beta$ away from the origin, for all $k \geq 0$ we then have

$$\partial_s v_k \leq \delta \Delta v_k + \sum_{j=1}^r Z_j^2 v_k - \frac{1}{\epsilon} Z_{0,\epsilon} v_k + S_k, \tag{11.22}$$

where

$$S_k = -[\chi_k, \partial_s] w_k + [\chi_k, \delta\Delta + \sum_{j=1}^r Z_j^2 - Ax \cdot \nabla] w_k. \tag{11.23}$$

Let g be a solution to the Dirichlet problem

$$\begin{cases} \partial_s g = \delta \Delta g + \sum_{j=1}^r Z_j^2 g - \frac{1}{\epsilon} Z_{0,\epsilon} g + S_k & (t,x) \in (0,2) \times B_{2R+1} \\ g|_{t=0} = 0 & \\ g|_{|x| = 2R+1} = 0. & \end{cases} \tag{11.24}$$

By the weak parabolic maximum principle, there holds $v_k \leq g$. Pairing (11.24) with g and using Grönwall's lemma we obtain

$$||g||_{L^\infty_t L^2_x} + ||g||_{L^2_t \mathscr{X}_\delta} \lesssim 2^{2k} \|w_k\|_{L^2_t L^2_x((t_k, 2-t_k)\times B_{R_k})}. \tag{11.25}$$

Introducing a radially-symmetric in space cutoff $\chi \in C_0^\infty((1/16, 31/16) \times B_{2R+1/2})$ with $\chi(t,x) = 1$ for $(t,x) \in (1/8, 15/8) \times 2R$ and using (11.24) again we then deduce

$$\|\chi g\|_{L^2 \mathscr{X}_\delta} + \|(\epsilon\partial_t + Z_{0,\epsilon})(\chi g)\|_{L^2 \mathscr{X}^*_\delta} \lesssim 2^{2k} \|w_k\|_{L^2_t L^2_x((t_k, 2-t_k)\times B_{R_k})}. \tag{11.26}$$

Therefore, by Assumption 2 and the parabolic Hörmander inequality, Lemma 9.5, we obtain the bound

$$||\chi g||_{L^\infty_t L^2_x} + ||\chi g||_{L^2_t H^s_x} \lesssim_R 2^{2k} \|w_k\|_{L^2_t L^2_x((t_k, 2-t_k)\times B_{R_k})}. \tag{11.27}$$

By $v_k \leq g$, (11.20), and the definition of χ_k, there is a constant $C > 0$ such that

$$\|w_k\|_{L^{2\alpha}_{t,x}((t_{k+1}, 2-t_{k+1})\times B_{R_{k+1}})} \leq ||v_k||_{L^{2\alpha}_{t,x}} \leq C^k \|w_k\|_{L^2_{t,x}((t_k, 2-t_k)\times B_{R_k})}, \tag{11.28}$$

which completes the proof of (11.21).

By definition, (11.21) implies

$$\left|\left|\tilde{\mathcal{P}}^{\epsilon,\delta}_s f\right|\right|_{L^{2\alpha^{k+1}}_{t,x}((t_{k+1}, 2-t_{k+1})\times B_{R_{k+1}})} \leq C^{\sum_{j=0}^k j\alpha^{-j}} \left|\left|\tilde{\mathcal{P}}^{\epsilon,\delta}_s f\right|\right|_{L^2_t\left(\frac{1}{8}, \frac{15}{8}; L^2_x(B_{2R})\right)}. \tag{11.29}$$

Passing to the limit and using $\sum_{j=0}^{\infty} j\alpha^{-j} < \infty$ along with the definitions of t_k and R_k yields (passing also to the limit $\delta \to 0$ by using (7.14))

$$\left|\left|\tilde{\mathcal{P}}^{\epsilon}_s f\right|\right|_{L^\infty_t\left(\frac{1}{4},\frac{7}{4};L^\infty_x(B_R)\right)} \lesssim_R \left|\left|\tilde{\mathcal{P}}^{\epsilon}_s f\right|\right|_{L^2_t\left(\frac{1}{8},\frac{15}{8};L^2_x(B_{2R})\right)}. \tag{11.30}$$

Finally, by the uniform lower bound (6.13) followed by the monotonicity (11.13) we have

$$\left|\left|\tilde{\mathcal{P}}^{\epsilon}_s f\right|\right|_{L^2_t\left(\frac{1}{8},\frac{15}{8};L^2_x(B_{2R})\right)} \leq \int_0^2 \left|\left|\tilde{\mathcal{P}}^{\epsilon}_s f\right|\right|^2_{L^2(B_{2R})} ds \lesssim_R \int_0^2 \left|\left|\tilde{\mathcal{P}}^{\epsilon}_s f\right|\right|^2_{L^2_{\mu_\epsilon}} ds \lesssim ||f||^2_{L^2_{\mu_\epsilon}}. \tag{11.31}$$

Combining (11.30) and (11.31) completes the proof of Lemma 6.3. □

11.3 Optimality of Theorem 6.5

In this section we prove Theorem 6.7. The idea is essentially that if one starts the process $(x^\epsilon_t)_{t\geq 0}$ at the origin, then the expected value of the energy $\mathbf{E}|x^\epsilon_t|^2$ must take at least time $t \gtrsim \epsilon^{-1}$ to reach equilibrium. Recall here that $(x^\epsilon_t)_{t\geq 0}$ denotes the solution to (6.3). For the basic properties of $(x^\epsilon_t)_{t\geq 0}$, see Lemma 7.1.

Proof of Theorem 6.7. We will only prove the statement that rules out $s < 1$, since the other is treated in the same way. Suppose for the sake of contradiction that there exists $s < 1$ and $K, \delta > 0$ so that for all measurable functions $f : \mathbb{R}^d \to \mathbb{R}$ with $\|f\|_V < \infty$ there holds

$$\|\mathcal{P}^\epsilon_t f - \mu_{\epsilon_n}(f)\|_V \leq K e^{-\delta \epsilon_n^s t} \|f - \mu_\epsilon(f)\|_V \tag{11.32}$$

for all $t \geq 0$ and $\epsilon \in (0, 1)$. We will derive a contradiction by considering $f : \mathbb{R}^d \to \mathbb{R}$

defined by $f(x) = x^2$, which clearly satisfies $\|f\|_V < \infty$.

By Itô's formula we have

$$\frac{1}{2}\mathbf{E}|x_t^\epsilon|^2 = \frac{1}{2}\mathbf{E}|x_0^\epsilon|^2 - \epsilon \int_0^t \mathbf{E}(Ax_s^\epsilon \cdot x_s^\epsilon)ds + \epsilon t \sum_{j=1}^r |Z_j|^2. \tag{11.33}$$

In statistical steady state this reduces to

$$\mu_\epsilon(f) \approx \int_{\mathbb{R}^d} (Ax \cdot x)\mu_\epsilon(dx) = \sum_{j=1}^r |Z_j|^2. \tag{11.34}$$

Next, applying (11.33) with $x_0^\epsilon \equiv 0$ gives

$$(\mathcal{P}_t^\epsilon f)(0) \lesssim \epsilon t. \tag{11.35}$$

Combining the previous two equations we see that there are constants $c, \eta > 0$ sufficiently small so that

$$|(\mathcal{P}_{c\epsilon^{-1}}^\epsilon f)(0) - \mu_\epsilon(f)| \geq \eta. \tag{11.36}$$

Hence, by (11.32) and the upper bound in (11.34) we have

$$\frac{\eta}{2} \leq \frac{|(\mathcal{P}_{c\epsilon^{-1}}^\epsilon f)(0) - \mu_\epsilon(f)|}{1 + V(0)} \leq \|\mathcal{P}_{c\epsilon^{-1}}^\epsilon f - \mu_\epsilon(f)\|_V \lesssim e^{-\delta c \epsilon^{s-1}}. \tag{11.37}$$

Since $s < 1$, sending $\epsilon \to 0$ yields the desired contradiction. $\square$

Appendix A: Additional Qualitative Regularity Properties

Recall from Chapter 7 that $\mu_{\epsilon,\delta}(dx) = f_{\epsilon,\delta}(x)dx$ denotes the unique invariant measure of the semigroup $\mathcal{P}_t^{\epsilon,\delta}$ generated by $\epsilon\delta\Delta + L_\epsilon$. In this Appendix we prove some qualitative properties of $f_{\epsilon,\delta}$ and rigorously justify the precise sense in which the limit $f_{\epsilon,\delta} \to f_\epsilon$ as $\delta \to 0$ holds.

We begin by showing that $f_{\epsilon,\delta}$ is indeed in L^2 provided that the regularization parameter δ is positive, which was used crucially in the proof of Lemma 10.1. Since this is a distinctively PDE type estimate, it requires an argument beyond the classical probabilistic ones appealed to in Chapter 7. The main idea is to take a smooth initial density ρ and use the Kolmogorov equation satisifed by $\rho_t := (\mathcal{P}_t^{\epsilon,\delta})^*\rho$ to show that the time averages $t^{-1}\int_0^t \rho_s ds$ remain uniformly bounded in H^1. Then, compactness and an argument similar to the proof of the Krylov-Boguliobov theorem yields the result. Rigorous justification of this sketch however entails some technicalities.

Lemma A.1. *The smooth density $f_{\epsilon,\delta}$ of the measure $\mu_{\epsilon,\delta}$ constructed in Chapter 7 is in L^2 whenever $\delta > 0$.*

Proof. The first step in the proof is to rigorously justify energy estimates for the PDE generated by $L_\epsilon^* + \epsilon\delta\Delta$. To this end, fix $\epsilon, \delta > 0$ and for $n \in \mathbb{N}$ let $\chi_n \in C_0^\infty(B_{2n})$ be a radially symmetric cutoff satisfying $\chi_n(x) = 1$ for $|x| \leq n$. With the precise setting being as in the beginning of Section 7.1, let $X_{t,x}^{(n)}$ denote the unique, global solution with initial condition $x \in \mathbb{R}^d$ to the SDE

$$\begin{aligned} dX_{t,x}^{(n)} = -\epsilon A X_{t,x}^{(n)} dt - \chi_n(X_{t,x}^{(n)}) \left(\epsilon^{\alpha} B X_{t,x}^{(n)} + N(X_{t,x}^{(n)})\right) dt \\ + \sqrt{2\epsilon} \sum_{j=1}^{r} Z_j dW_t^{(j)} + \sqrt{2\epsilon\delta} dB_t, \end{aligned} \tag{A.1}$$

where B_t denotes a standard d-dimensional Brownian motion independent from and defined on the same complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ as the Brownian motions $\{W_t^{(j)}\}_{j=1}^{r}$. Of course, the filtration $\mathcal{F}_t$ defined in Section 7.1 is now appropriately augmented to account for the presence of B_t. Let $\mathcal{P}_t^n$ denote the Markov semigroup generated by $(X_t^{(n)})_{t\geq 0}$.

Let $\rho \in C_0^{\infty}(B_1)$ be a probability density function and set $\mu(dx) = \rho(x)dx$, $(\mathcal{P}_t^{\epsilon,\delta})^*\mu = \rho_t(x)dx$, and $(\mathcal{P}_t^n)^*\mu = \rho_t^{(n)}(x)dx$. Define also $F_n = (\epsilon^{\alpha} Bx + N)\chi_n \in C_0^{\infty}(B_{2n})$, where for notational convenience we suppress the dependence on ϵ. The Kolmogorov equation for $\rho_t^{(n)}$ takes the form

$$\partial_t \rho_t^{(n)} = \mathcal{L}\rho_t^{(n)} + \nabla \cdot (F_n \rho_t^{(n)}) \tag{A.2}$$

where $\mathcal{L}$ is a Fokker-Planck operator that can be expressed as

$$\mathcal{L}\rho_t^{(n)} = \epsilon \nabla \cdot \left(D\nabla \rho_t^{(n)} + Ax\rho_t^{(n)} \right). \tag{A.3}$$

for some symmetric, positive definite matrix D that depends on δ. For simplicity, we suppress the dependence of $\mathcal{L}$ on ϵ and δ from the notation. The semigroup $e^{\mathcal{L}t}$ is given by integration against a smooth, explicitly computable kernel. Namely,

$$e^{\mathcal{L}t}f = \frac{C}{\sqrt{J(\Lambda_t)}} \int f(y) \exp\left(-\langle \Lambda_t^{-1}(x - e^{-At}y), x - e^{-At}y\rangle_{\mathbb{R}^d}\right) dy,$$

where

$$\Lambda_t = \int_0^t e^{-A(t-s)} ADA e^{-A(t-s)} ds$$

and $J(\Lambda_t)$ denotes the determinant of the transformation Λ_t. This formula allows one to show that there exists $\gamma > 0$ sufficiently small such for any $f \in C_0^\infty$ and $T > 0$ there holds (using standard multi-index notation, $\alpha \in \mathbb{N}^d$)

$$\sup_{0\leq t\leq T} \|e^{\mathcal{L}t}f\|_{N^k} := \sup_{0\leq t\leq T} \sum_{|\alpha|\leq k} \left\| e^{\gamma|x|^2} D^\alpha e^{\mathcal{L}t} f \right\|_{L^2} \leq C(k, T, \text{spt}(f), \|f\|_{C^k}) < \infty.$$

Since $\rho_t^{(n)}$ is a smooth solution to (A.2), it satisfies the Duhamel formula

$$\rho_t^{(n)} = e^{\mathcal{L}t}\rho + \int_0^t e^{\mathcal{L}(t-s)} \nabla \cdot (F_n \rho_t^{(n)})(s) ds,$$

from which we deduce that $\sup_{0\leq t\leq T} \|\rho_t^{(n)}\|_{N^k} < \infty$ on every finite time interval. One can then rigorously justify energy estimates using (A.2). In particular, from standard calculations we can show that

$$\sup_{0\leq t\leq T} \sum_{|\alpha|\leq k} \|e^{(1+|\alpha|)^{-1}\gamma|x|^2} D^\alpha \rho_t^{(n)}\|_{L^2} < \infty$$

with an n independent (though ϵ and δ dependent) upper bound. Sending $n \to \infty$ and passing to a subsequence (which we do not relabel), we extract a limit $\tilde{\rho}_t$ with $\rho_t^{(n)} \to \tilde{\rho}_t$ strongly in H^k for any $k < \infty$. Moreover, $\tilde{\rho}_t \geq 0$, $\int \tilde{\rho}_t = 1$ for all $t \geq 0$, and $\tilde{\rho}_t$ solves the Kolmogorov equation for ρ_t. By using an integrating factor to remove the zero order term in (A.2) and appealing to the uniqueness of sufficiently regular solutions to Kolmogorov backward equations (see e.g. [Theorem 8.1.1, [110]]) we conclude that $\tilde{\rho}_t = \rho_t$. In particular, for every $k > 0$ there exists some $\gamma_k > 0$ such that $e^{\gamma_k x^2}\rho_t$ remains uniformly bounded in H^k on finite time intervals.

The considerations above allow us to justify energy estimates on the equation for ρ_t, namely

$$\begin{cases} \partial_t \rho_t(x) = L_\epsilon^* \rho_t(x) + \epsilon\delta\Delta\rho_t(x) & (t,x) \in (0,\infty) \times \mathbb{R}^d \\ \rho_0(x) = \rho(x) & x \in \mathbb{R}^d. \end{cases} \tag{A.4}$$

Pairing the equation with ρ_t, we have

$$\frac{d}{dt}\|\rho_t\|_{L^2}^2 + \|\nabla\rho_t\|_{L^2}^2 \lesssim_\delta \|\rho_t\|_{L^2}^2. \tag{A.5}$$

Integrating this estimate and applying the Gagliardo-Nirenberg to bound $\|\rho_t\|_{L^2} \lesssim \|\nabla\rho_t\|_{L^2}^\theta \|\rho_t\|_{L^1}^{1-\theta} = \|\nabla\rho_t\|_{L^2}^\theta$ we obtain

$$\int_0^t \|\nabla\rho_s\|_{L^2}^2 ds \lesssim_\delta 1 + t. \tag{A.6}$$

Thus,

$$\int_0^t \|\rho_s\|_{H^1} ds \lesssim t + \int_0^t \|\nabla\rho_s\|_{L^2}^2 ds \lesssim_\delta 1 + t, \tag{A.7}$$

where in the first inequality we used again $\|\rho_t\|_{L^2} \lesssim 1 + \|\nabla \rho_t\|_{L^2}$. For $n \geq 1$ we define the probability density function

$$\rho_{n,\mathrm{KB}} = \frac{1}{n}\int_0^n \rho_s ds.$$

By (A.7), the sequence $\{\rho_{n,\mathrm{KB}}\}_{n=1}^{\infty}$ is uniformly bounded in H^1, and so passing to a subsequence (which we do not relabel) we obtain a limit $\rho_\infty \in H^1$ with $\lim_{n\to\infty} \rho_{n,\mathrm{KB}} = \rho_\infty$ weakly in H^1 and strongly in L^2 on compact subsets. We may also assume that $\rho_{n,\mathrm{KB}} \to \rho_\infty$ pointwise a.e., so $\rho_\infty \geq 0$ a.e. Moreover, due to (7.2) there holds $\sup_{t\geq 0} \int e^{\gamma|x|^2}\rho_t(x)dx < \infty$ for $\gamma > 0$ small enough, which when combined with the strong L^2_{loc} convergence $\rho_{n,\mathrm{KB}} \to \rho_\infty$ implies that $\int \rho_\infty = 1$. Similar to the proof of the Krylov-Bogoliubov theorem we can show that ρ_∞ solves $(L_\epsilon^* + \epsilon\delta\Delta)\rho_\infty = 0$ in the sense of distributions. By the uniqueness described in Lemma 7.3 we conclude that $f_{\epsilon,\delta} = \rho_\infty \in H^1$, which completes the proof. □

Next, we have a lemma regarding the elliptic regularization, which justifies our approximation arguments with $f_{\epsilon,\delta}$.

Lemma A.2. *For all $\epsilon > 0$, $k \geq 0$, and $R > 0$,*

$$\sup_{\delta\in[0,1]} ||f_{\epsilon,\delta}||_{H^k(B_R)} \lesssim_{k,\epsilon,R} 1. \tag{A.8}$$

For each fixed $\epsilon > 0$ there exists a subsequence $\{\delta_n\} \subset (0,1)$ with $\lim_{n\to\infty}\delta_n = 0$ such *that for all $k \geq 0$ and $R > 0$ there holds*

$$\lim_{n\to\infty} ||f_{\epsilon,\delta_n} - f_\epsilon||_{H^k(B_R)} = 0. \tag{A.9}$$

Proof. Let $s \in (0,1)$ be as given in Lemma 9.3. Let $k \leq sJ$ for $J \in \mathbb{N}$ fixed and define a decreasing sequence of radially-symmetric, smooth cutoff functions χ_j which satisfy $\chi_j(x) = 1$ for $|x| \leq R + J - j$ and $\chi_j(x) = 0$ for $|x| > R + J - j + 1$. Define $\langle \nabla \rangle^s$ as the Fourier multiplier

$$\widehat{\langle \nabla \rangle^s u}(\xi) = \left(1 + |\xi|^2\right)^{s/2} \widehat{u}(\xi).$$

Let $v_0 = \chi_0 f_{\epsilon,\delta}$ and $v_j = \langle \nabla \rangle^{sj} \chi_j f_{\epsilon,\delta}$. Then,

$$\epsilon\delta\Delta v_j + L_\epsilon^* v_j + \langle \nabla \rangle^{sj} [\chi_j, \epsilon\delta\Delta + \epsilon \sum_{k=1}^{r} Z_k^2] f_{\epsilon,\delta} + \langle \nabla \rangle^{sj} [\chi_j, Ax \cdot \nabla] f_{\epsilon,\delta} + \mathcal{C}_j = 0, \tag{A.10}$$

where we denote

$$\mathcal{C}_j = [\langle \nabla \rangle^{sj}, Z_{0,\epsilon} \cdot \nabla] \chi_j f_{\epsilon,\delta}. \tag{A.11}$$

Note that

$$\begin{aligned} \epsilon\delta\, ||\nabla v_j||_{L^2}^2 + \epsilon \sum_{k=1}^{r} ||Z_k v_j||_{L^2}^2 \lesssim_{R,k} ||v_j||_{L^2}^2 + \mathbf{1}_{j\geq 1} \| \langle \nabla \rangle^s v_{j-1} \|_{L^2}^2 \\ + \|f_{\epsilon,\delta}\|_{L^2}^2 + \left| \int v_j \mathcal{C}_j dx \right|. \end{aligned} \tag{A.12}$$

To bound the term involving C_j, we first rewrite it on the Fourier side to obtain

$$\left|\int v_j C_j dx\right| \lesssim \int\int |\hat{v}_j(\xi)| |\langle\xi\rangle^{sj} - \langle\eta\rangle^{sj}| |\widehat{\chi Z_{0,\epsilon}}(\xi-\eta)||\eta||\widehat{\chi_j f_{\epsilon,\delta}}(\eta)|d\eta d\xi, \quad \text{(A.13)}$$

where $\chi \in C_0^\infty(\mathbb{R}^d)$ is a smooth cutoff with $\chi(x) = 1$ for all $|x| \leq R+J+2$. By splitting the integral between the regions $|\xi-\eta| > |\eta|/2$, $|\xi-\eta| \leq |\eta|/2$ and using the mean value theorem in the latter piece to deduce $| \langle\xi\rangle^{sj} - \langle\eta\rangle^{sj} | \lesssim \langle\xi-\eta\rangle \langle\eta\rangle^{sj-1}$ we can show

$$\left|\int v_j \mathcal{C}_j dx\right| \lesssim ||v_j||_{L^2} \left(||v_j||_{L^2} + ||f_{\epsilon,\delta}||_{L^2}\right). \quad \text{(A.14)}$$

Pairing (A.10) with test functions similarly gives

$$||Z_{0,\epsilon} v_j||_{\mathscr{X}_\delta^*} \lesssim ||v_j||_{L^2} + ||f_{\epsilon,\delta}||_{L^2} + \mathbf{1}_{j\geq 1} \| \langle\nabla\rangle^s v_{j-1}\|_{L^2}. \quad \text{(A.15)}$$

Therefore, by Lemma 9.3, we have, independent of δ,

$$\|v_j\|_{H^s} \lesssim \|v_j\|_{L^2} + \mathbf{1}_{j\geq 1}\| \langle\nabla\rangle^s v_{j-1}\|_{L^2} + \|f_{\epsilon,\delta}\|_{L^2} \lesssim \|f_{\epsilon,\delta}\|_{L^2} + \mathbf{1}_{j\geq 1}\|v_{j-1}\|_{H^s}. \quad \text{(A.16)}$$

Iterating gives (A.8). From there, to deduce (A.9) we first use the standard compact embedding theorem to extract a subsequence $\{f_{\epsilon,\delta_n}\}_{n=1}^\infty$ with $\delta_n \to 0$ and a limit $f_{\epsilon,0} \in C^\infty$ with $\lim_{n\to\infty} f_{\epsilon,\delta_n} = f_{\epsilon,0}$ in H^k_{loc} for every k. Clearly, $f_{\epsilon,0} \geq 0$ and $L_\epsilon^* f_{\epsilon,0} = 0$. Moreover by (7.15) we have $\int f_{\epsilon,0} = 1$. Hence, $f_{\epsilon,0} = f_\epsilon$ by uniqueness, which completes the proof. $\square$

We conclude with a qualititative lower bound for f_ϵ that holds for $\epsilon \gtrsim 1$.

Lemma A.3. *Suppose that Assumption 3 holds. Then, for any $R \geq 1$ and $\epsilon_* \in (0,1)$ there exists $C(\epsilon_*, R) > 0$ such that*

$$\inf_{\epsilon \in [\epsilon_*, 1], \delta \in [0,1]} \inf_{|x| \leq R} f_{\epsilon,\delta}(x) \geq C. \tag{A.17}$$

Proof. First, note that $f_{\epsilon,\delta}$ is strictly positive for all $\epsilon \in (0,1)$, $\delta \in [0,1]$. Indeed, $f_\epsilon > 0$ by assumption, and the fact that $f_{\epsilon,\delta} > 0$ when $\delta > 0$ follows from the classical elliptic Harnack inequality. Now, if the claim is false, then using (A.8) and the argument used to prove (A.9) we can obtain $(\epsilon_0, \delta_0, x_0) \in [\epsilon_*, 1] \times [0,1] \times \bar{B}_R$ such that $f_{\epsilon_0,\delta_0}(x_0) = 0$, which contradicts $f_{\epsilon_0,\delta_0} > 0$. $\square$

Bibliography

[1] F. Abedin and G. Tralli. Harnack inequality for a class of Kolmogorov–Fokker–Planck equations in non-divergence form. *Archive for Rational Mechanics and Analysis*, 233(2):867–900, 2019.

[2] A. A. Agrachev and Y. Sachkov. *Control theory from the geometric viewpoint*, volume 87. Springer Science & Business Media, 2013.

[3] F. Alavyoon, D. S. Henningson, and P. H. Alfredsson. Turbulent spots in plane Poiseuille flow–flow visualization. *The Physics of Fluids*, 29(4):1328–1331, 1986.

[4] D. Albritton, R. Beekie, and M. Novack. Enhanced dissipation and Hörmander's hypoellipticity. *arXiv:2105.12308*, 2021.

[5] S. Alinhac. The null condition for quasilinear wave equations in two space dimensions. *Inventiones mathematicae*, 145(3):597–618, 2001.

[6] F. Anceschi, S. Polidoro, and M. A. Ragusa. Moser's estimates for degenerate Kolmogorov equations with non-negative divergence lower order coefficients. *Nonlinear Analysis*, 189(111568), 2019.

[7] S. Armstrong and J.-C. Mourrat. Variational methods for the kinetic Fokker-Planck equation. *arXiv:1902.04037*, 2019.

[8] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Comm. in Part. Diff. Eqns.*, 26(1-2):43–100, 2001.

[9] H. Bahouri, J. Chemin, and R. Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2011.

[10] D. Bakry, F. Barthe, P. Cattiaux, A. Guillin, et al. A simple proof of the Poincaré inequality for a large class of probability measures. *Electronic Communications in Probability*, 13(7):60–66, 2008.

[11] D. Bakry, P. Cattiaux, and A. Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *Journal of Functional Analysis*, 254(3):727–759, 2008.

[12] J. Bedrossian. Nonlinear echoes and landau damping with insufficient regularity. *Tunis. J. Math*, 3(1):121–205, 2016.

[13] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith. Lagrangian chaos and scalar advection in stochastic fluid mechanics. *arXiv:1809.06484*, 2018.

[14] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith. The Batchelor spectrum of passive scalar turbulence in stochastic fluid mechanics at fixed Reynolds number. *arXiv:1911.11014*, 2019.

[15] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith. A regularity method for lower bounds on the Lyapunov exponent for stochastic differential equations. *arXiv:2007.15827*, 2020.

[16] J. Bedrossian, P. Germain, and N. Masmoudi. Dynamics near the subcritical transition of the 3D Couette flow II: Above threshold. *arXiv:1506.03721 (2015), to appear in Mem. of the AMS*.

[17] J. Bedrossian, P. Germain, and N. Masmoudi. On the stability threshold for the 3D Couette flow in Sobolev regularity. *Annals of Mathematics*, 185(2):541–608, 2017.

[18] J. Bedrossian, P. Germain, and N. Masmoudi. Stability of the Couette flow at high Reynolds numbers in two dimensions and three dimensions. *Bull. Amer. Math. Soc.*, 56(3):373–414, 2019.

[19] J. Bedrossian, P. Germain, and N. Masmoudi. Dynamics near the subcritical transition of the 3D Couette flow I: Below threshold case. *Mem. of the AMS*, 266(1294), 2020.

[20] J. Bedrossian and K. Liss. Quantitative spectral gaps and uniform lower bounds in the small noise limit for Markov semigroups generated by hypoelliptic stochastic differential equations. *arXiv:2007.13297 (2020), to appear in Probability and Mathematical Physics*.

[21] J. Bedrossian and N. Masmoudi. Inviscid damping and the asymptotic stability of planar shear flows in the 2D Euler equations. *Publications mathématiques de l'IHÉS*, 122(1):195–300, 2015.

[22] J. Bedrossian, N. Masmoudi, and V. Vicol. Enhanced dissipation and inviscid damping in the inviscid limit of the Navier-Stokes equations near the 2D Couette flow. *Arch. Rat. Mech. Anal.*, 216(3):1087–1159, 2016.

[23] J. Bedrossian, V. Vicol, and F. Wang. The Sobolev stability threshold for 2D shear flows near Couette. *Journal of Nonlinear Science*, 28(6):2051–2075, 2018.

[24] J. Bedrossian and M. C. Zelati. Enhanced dissipation, hypoellipticity, and anomalous small noise inviscid limits in shear flows. *Archive for Rational Mechanics and Analysis*, 224(3):1161–1204, 2017.

[25] J. Bedrossian, M. C. Zelati, S. Punshon-Smith, and F. Weber. A sufficient condition for the Kolmogorov 4/5 law for stationary martingale solutions to the 3D Navier-Stokes equations. *Communications in Mathematical Physics*, 367(3):1045–1075, 2019.

[26] J. Bedrossian, M. C. Zelati, S. Punshon-Smith, and F. Weber. Sufficient conditions for dual cascade flux laws in the stochastic 2d Navier-Stokes equations. *Archive for Rational Mechanics and Analysis*, 237:103–145, 2020.

[27] F. Bouchut. Hypoelliptic regularity in kinetic equations. *Journal de Mathématiques Pures et Appliquées*, 81(11):1135–1159, 2002.

[28] T. Boyd and J. Sanderson. *The physics of plasmas*. Cambridge University Press, 2003.

[29] T. Buckmaster, P. Germain, Z. Hani, and J. Shatah. Onset of the wave turbulence description of the longtime behavior of the nonlinear Schrödinger equation. *Inventiones mathematicae*, 884:1–69, 2021.

[30] O. Butkovsky. Subgeometric rates of convergence of Markov processes in the Wasserstein metric. *Ann. Appl. Probab.*, 24(2):526–552, 2014.

[31] D. R. Carlson, S. E. Widnall, and M. F. Peeters. A flow-visualization study of transition in plane Poiseuille flow. *Journal of Fluid Mechanics*, 121:487–505, 1982.

[32] J. Cassels. *An introduction to Diophantine approximation*. Cambridge University Press, Cambridge, 1957.

[33] S. Chandrasekhar. The stability of viscous flow between rotating cylinders in the presence of a magnetic field. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 216(1126):293–309, 1953.

[34] S. Chandrasekhar. The stability of non-dissipative Couette flow in hydromagnetics. *Proceedings of the National Academy of Sciences of the United States of America*, 46(2):253–257, 1960.

[35] S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability*. Dover Publications, 1981.

[36] C. Cinti, K. Nyström, and S. Polidoro. A note on Harnack inequalities and propagation sets for a class of hypoelliptic operators. *Potential Analysis*, 33(4):341–354, 2010.

[37] G. Da Prato and J. Zabczyk. *Ergodicity for Infinite Dimensional Systems*. Cambridge University Press, 1996.

[38] P. A. Davidson. *An Introduction to Magnetohydrodynamics*. Cambridge University Press, 2001.

[39] W. Deng, J. Wu, and P. Zhang. Stability of Couette flow for 2D Boussinesq system with vertical dissipation. *arXiv:2004.09292*, 2020.

[40] Y. Deng and N. Masmoudi. Long time instability of Couette flow in low Gevrey spaces. *arXiv:1803.01246*, 2018.

[41] P. D. Ditlevsen. *Turbulence and shell models*. Cambridge University Press, 2010.

[42] S. Dong, P. Geng, C. Li, X. Ye, and Z. Wu. Transition in channel flow with streamwise magnetic field. *European Journal of Mechanics-B/Fluids*, 77:79–86, 2019.

[43] R. J. Donnelly and M. Ozima. Experiments on the stability of flow between rotating cylinders in the presence of a magnetic field. *Proceedings of the Royal Society of London A: Mathematical and Physical Sciences*, 266(1325):272–286, 1962.

[44] R. Douc, G. Fort, and A. Guillin. Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stochastic Processes and their Applications*, 119(3):897–923, 2009.

[45] P. G. Drazin and W. H. Reid. *Hydrodynamic Stability*. Cambridge University Press, 1981.

[46] Y. Duguet, A. Monokrousos, L. Brandt, and D. S. Henningson. Minimal transition thresholds in plane Couette flow. *Physics of Fluids*, 25(084103), 2013.

[47] A. Durmus, G. Fort, and E. Moulines. Subgeometric rates of convergence in Wasserstein distance for Markov chains. *Ann. Inst. H. Poincaré Probab. Statist.*, 52(4):1799–1822, 2016.

[48] A. Dymov. Nonequilibrium statistical mechanics of weakly stochastically perturbed system of oscillators. *Annales Henri Poincaré*, 17(7):1825–1882, 2016.

[49] A. Dymov and S. Kuksin. On the Zakharov–L'vov stochastic model for wave turbulence. *Doklady Mathematics*, 101(2):102–109, 2020.

[50] T. Ellingsen and E. Palm. Stability of linear flow. *The Physics of Fluids*, 18(4):487–488, 1975.

[51] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*. Springer, 2001.

[52] F. Flandoli. An introduction to 3D stochastic fluid dynamics. In *SPDE in hydrodynamic: recent progress and prospects*, pages 51–150. Springer, 2008.

[53] F. Flandoli and B. Maslowski. Ergodicity of the 2-D Navier-Stokes equation under random perturbations. *Communications in mathematical physics*, 172(1):119–141, 1995.

[54] J. Foldes, S. Friedlander, N. Glatt-Holtz, and G. Richards. Asymptotic analysis for randomly forced MHD. *SIAM Journal on Mathematical Analysis*, 49(6):4440–4469, 2017.

[55] U. Frisch. *Turbulence: The Legacy of AN Kolmogorov*. Cambridge University Press, 1995.

[56] P. Germain. Global existence for coupled Klein-Gordon equations with different speeds. *Annales de l'Institut Fourier*, 61(6):2463–2506, 2011.

[57] P. Germain. Space-time resonances. *arXiv:1102.1695*, 2011.

[58] P. Germain and N. Masmoudi. Global existence for the Euler-Maxwell system. *Annales Scientifiques de l'Ecole Normale Superieure*, 47(3):469–503, 2014.

[59] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*, volume 224. Springer, 2015.

[60] N. Glatt-Holtz, J. C. Mattingly, and G. Richards. On unique ergodicity in nonlinear stochastic partial differential equations. *Journal of Statistical Physics*, 166(3-4):618–649, 2017.

[61] N. E. Glatt-Holtz, D. P. Herzog, and J. C. Mattingly. Scaling and saturation in infinite-dimensional control problems with applications to stochastic partial differential equations. *Annals of PDE*, 4(2):1–103, 2018.

[62] B. Goldys and B. Maslowski. Exponential ergodicity for stochastic Burgers and 2D Navier–Stokes equations. *Journal of Functional Analysis*, 226(1):230–255, 2005.

[63] B. Goldys and B. Maslowski. Lower estimates of transition densities and bounds on exponential ergodicity for stochastic PDEs. *The Annals of Probability*, 34(4):1451–1496, 2006.

[64] F. Golse, C. Imbert, C. Mouhot, and A. Vasseur. Harnack inequality for kinetic Fokker-Planck equations with rough coefficients and application to the Landau equation. *Scuola Normale Superiore di Pisa, Classe di Scienze*, 19:253–295, 2019.

[65] F. Golse and A. Vasseur. Hölder regularity for hypoelliptic kinetic equations with rough diffusion coefficients. *arXiv:1506.01908*, 2015.

[66] M. Grothaus and F.-Y. Wang. Weak Poincaré inequalities for convergence rate of degenerate diffusion processes. *Ann. Probab.*, 47(5):2930–2952, 2019.

[67] M. Hairer. An introduction to stochastic PDEs. *arXiv:0907.4178*, 2009.

[68] M. Hairer. On Malliavin's proof of Hörmander's theorem. *Bulletin des sciences mathematiques*, 135(6-7):650–666, 2011.

[69] M. Hairer and J. Mattingly. A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs. *Electronic Journal of Probability*, 16(23):658–738, 2011.

[70] M. Hairer, J. Mattingly, and M. Scheutzow. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Prob. Theory Rel. Fields*, 149:223–259, 2011.

[71] M. Hairer and J. C. Mattingly. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Annals of Mathematics*, 164(3):993–1032, 2006.

[72] M. Hairer and J. C. Mattingly. Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations. *Ann. Probab.*, 36(6):2050–2091, 2008.

[73] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 109–117. Springer Basel, 2011.

[74] L.-B. He, L. Xu, and P. Yu. On global dynamics of three dimensional magnetohydrodynamics: nonlinear stability of Alfvén waves. *Annals of PDE*, 4(1):1–105, 2018.

[75] R. Hermann and A. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22(5):728–740, 1977.

[76] D. P. Herzog and J. C. Mattingly. A practical criterion for positivity of transition densities. *Nonlinearity*, 28(8):2823–2845, 2015.

[77] S. Hu and X. Wang. Subexponential decay in kinetic Fokker–Planck equation: Weak hypocoercivity. *Bernoulli*, 25(1):174–188, 2019.

[78] J. Hunt. On the stability of parallel flows with parallel magnetic fields. 293(1434):342–358, 1966.

[79] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[80] C. Imbert and C. Mouhot. Hölder continuity of solutions to hypoelliptic equations with bounded measurable coefficients. *arXiv:1505.04608*, 2015.

[81] A. Ionescu and H. Jia. Inviscid damping near the Couette flow in a channel. *Communications in Mathematical Physics*, 374:2015–2096, 2020.

[82] A. Karimi and M. R. Paul. Extensive chaos in the Lorenz-96 model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 20(043105), 2010.

[83] L. Kelvin. Stability of fluid motion-rectilinear motion of viscous fluid between two parallel plates. *Phil. Mag.*, 24(5):188–196, 1887.

[84] A. E. Kogoj and S. Polidoro. Harnack inequality for hypoelliptic second order partial differential operators. *Potential Anal.*, 45(14):545–555, 2016.

[85] D. Krasnov, M. Rossi, O. Zikanov, and T. Boeck. Optimal growth and transition to turbulence in channel flow with spanwise magnetic field. *Journal of Fluid Mechanics*, 596:73–101, 2008.

[86] S. Kuksin, V. Nersesyan, and A. Shirikyan. Exponential mixing for a class of dissipative PDEs with bounded degenerate noise. *Geometric and Functional Analysis*, 30:126–187, 2020.

[87] S. Kuksin, V. Nersesyan, and A. Shirikyan. Mixing via controllability for randomly forced nonlinear dissipative PDEs. *Journal de l'École polytechnique—Mathématiques*, 7:871–896, 2020.

[88] S. Kuksin and H. Zhang. Exponential mixing for dissipative PDEs with bounded non-degenerate noise. *Stochastic Processes and their Applications*, 130:4721–4745, 2020.

[89] A. Kupiainen. Ergodicity of two dimensional turbulence. *arXiv:1005.0587*, 2010.

[90] A. Lanconelli, A. Pascucci, and S. Polidoro. Gaussian lower bounds for non-homogeneous Kolmogorov equations with measurable coefficients. *Journal of Evolution Equations*, 20:1399–1417, 2020.

[91] T. M. Liggett. L_2 rates of convergence for attractive reversible nearest particle systems: the critical case. *Ann. Probab.*, 19(3):935–959, 1991.

[92] Z. Lin and C. Zeng. Inviscid dynamical structures near Couette flow. *Archive for rational mechanics and analysis*, 200(3):1075–1097, 2011.

[93] K. Liss. On the Sobolev stability threshold of 3D Couette flow in a uniform magnetic field. *Communications in Mathematical Physics*, 377:859–908, 2020.

[94] Y. Liu, Z. H. Chen, H. H. Zhang, and Z. Y. Lin. Physical effects of magnetic fields on the Kelvin-Helmholtz instability in a free shear layer. *Physics of Fluids*, 30(044102), 2018.

[95] R. Lock. The stability of the flow of an electrically conducting fluid between parallel planes under a transverse magnetic field. *Proceedings of the Royal Society of London Series A: Mathematical and Physical Sciences*, 233(1192):105–125, 1955.

[96] E. N. Lorenz. Predictability: A problem partly solved. In *Proc. Seminar on predictability*, volume 1, pages 1–18, 1996.

[97] E. N. Lorenz and K. A. Emanuel. Optimal sites for supplementary weather observations: Simulation with a small model. *Journal of the Atmospheric Sciences*, 55(3):399–414, 1998.

[98] A. Lundbladh, D. S. Henningson, and S. C. Reddy. Threshold amplitudes for transition in channel flows. In *Transition, turbulence and combustion*, pages 309–318. Springer, 1994.

[99] V. S. L'vov, E. Podivilov, A. Pomyalov, I. Procaccia, and D. Vandembroucq. Improved shell model of turbulence. *Physical Review E*, 58(2):1811–1822, 1998.

[100] A. Majda and A. Bertozzi. *Vorticity and Incompressible Flow*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002.

[101] A. J. Majda. *Introduction to turbulent dynamical systems in complex systems*. Springer, 2016.

[102] N. Masmoudi and W. Zhao. Stability threshold of the 2D Couette flow in Sobolev spaces. *arXiv:1908.11042*, 2019.

[103] J. C. Mattingly and É. Pardoux. Malliavin calculus for the stochastic 2d navier—stokes equation. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(12):1742–1790, 2006.

[104] S. Meyn and R. L. Tweedie. Computable bounds for geometric convergence rates of Markov chains. *The Annals of Applied Probability*, 4(4):981–1011, 1994.

[105] S. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, 2nd edition, 2009.

[106] D. H. Michael. Stability of plane parallel flows of electrically conducting fluids. *Mathematical Proceedings of the Cambridge Philosophical Society*, 49(1):166–168, 1953.

[107] C. Mouhot. De Giorgi–Nash–Moser and Hörmander theories: new interplays. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro*, volume 3, pages 2467–2493. World Scientific, 2018.

[108] C. Mouhot and C. Villani. On Landau damping. *Acta mathematica*, 207(1):29–201, 2011.

[109] S. Nazarenko. *Wave turbulence*. Springer Science & Business Media, 2011.

[110] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Hochschultext / Universitext. Springer, 2003.

[111] W. Orr. The stability or instability of steady motions of a perfect liquid and of a viscous liquid, Part I: a perfect liquid. *Proc. Royal Irish Acad. Sec. A: Math. Phys. Sci.*, 27:9–68, 1907.

[112] S. Papathanasiou. Sufficient conditions for local scaling laws for stationary martingale solutions to the 3D Navier–Stokes equations. *Nonlinearity*, 34(5):2937–2969, 2021.

[113] A. Pascucci and S. Polidoro. The Moser's iterative method for a class of ultraparabolic equations. *Communications in Contemporary Mathematics*, 6(3):395–417, 2004.

[114] S. Polidoro. A global lower bound for the fundamental solution of Kolmogorov-Fokker-Planck equations. *Archive for Rational Mechanics and Analysis*, 137(4):321–340, 1997.

[115] S. C. Reddy, P. J. Schmid, J. S. Baggett, and D. S. Henningson. On stability of streamwise streaks and transition thresholds in plane channel flows. *Journal of Fluid Mechanics*, 365:269–303, 1998.

[116] M. Röckner and F.-Y. Wang. Weak Poincaré inequalities and L^2-convergence rates of Markov semigroups. *Journal of Functional Analysis*, 185(2):564 – 603, 2001.

[117] J. Stuart. On the stability of viscous flow between parallel planes in the presence of a co-planar magnetic field. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 221(1145):189–206, 1954.

[118] N. Tillmark and P. H. Alfredsson. Experiments on transition in plane Couette flow. *Journal of Fluid Mechanics*, 235:89–102, 1992.

[119] A. F. Vasseur. The De Giorgi method for elliptic and parabolic equations and some applications. In *Morningside Lectures in Mathematics*, volume 4, pages 2467–2493. International Press, 2016.

[120] E. Velikhov. Stability of an ideally conducting liquid flowing between cylinders rotating in a magnetic field. *Sov. Phys. JETP*, 36(9):995–998, 1959.

[121] C. Villani. *Hypocoercivity*. American Mathematical Society, 2009.

[122] W. Wang and L. Zhang. The C^α regularity of weak solutions of ultraparabolic equations. *Discrete & Continuous Dynamical Systems-A*, 29(3):1261, 2011.

[123] D. Wei and Z. Zhang. Global well-posedness of the MHD equations in a homogeneous magnetic field. *Analysis and PDE*, 10(6):1361–1406, 2017.

[124] D. Wei and Z. Zhang. Transition threshold for the 3D Couette flow in Sobolev space. *Comm. Pure Appl. Math.*, 2018.

[125] D. Wei, Z. Zhang, and W. Zhao. Linear inviscid damping and enhanced dissipation for the Kolmogorov flow. *Advances in Mathematics*, 362(106963), 2020.

[126] E. Weinan and J. C. Mattingly. Ergodicity for the Navier-Stokes equation with degenerate random forcing: finite-dimensional approximation. *Comm. Pure Appl. Math*, 54(11):1386–1402, 2001.

[127] W. Wendong and L. Zhang. C^α regularity of weak solutions of non-homogenous ultraparabolic equations with drift terms. *arXiv:1704.05323*, 2017.

[128] A. Yaglom and U. Frisch. *Hydrodynamic Instability and Transition to Turbulence*. Fluid Mechanics and Its Applications. Springer Netherlands, 2012.

[129] V. E. Zakharov, V. S. L'vov, and G. Falkovich. *Kolmogorov Spectra of Turbulence I: Wave Turbulence*. Springer Science & Business Media, 2012.

[130] M. C. Zelati, M. G. Delgadino, and T. M. Elgindi. On the relation between enhanced dissipation timescales and mixing rates. *Communications on Pure and Applied Mathematics*, 73(6):1205–1244, 2020.

[131] M. C. Zelati, T. M. Elgindi, and K. Widmayer. Enhanced dissipation in the Navier-Stokes equations near the Poiseuille flow. *Communications in Mathematical Physics*, 378(2):987–1010, 2020.

[132] C. Zillinger. On enhanced dissipation for the Boussinesq equations. *Journal of Differential Equations*, 282:407–445, 2021.