VARIATIONAL METHODS FOR THE KINETIC FOKKER-PLANCK EQUATION

D. ALBRITTON, S. ARMSTRONG, J.-C. MOURRAT, AND M. NOVACK

ABSTRACT. We develop a functional analytic approach to the study of the Kramers and kinetic Fokker-Planck equations which parallels the classical H^1 theory of uniformly elliptic equations. In particular, we identify a function space analogous to H^1 and develop a well-posedness theory for weak solutions in this space. In the case of a conservative force, we identify the weak solution as the minimizer of a uniformly convex functional. We prove new functional inequalities of Poincaré and Hörmander type and combine them with basic energy estimates (analogous to the Caccioppoli inequality) in an iteration procedure to obtain the C^∞ regularity of weak solutions. We also use the Poincaré-type inequality to give an elementary proof of the exponential convergence to equilibrium for solutions of the kinetic Fokker-Planck equation which mirrors the classic dissipative estimate for the heat equation. Finally, we prove enhanced dissipation in a weakly collisional limit.

CONTENTS

1. INTRODUCTION

1.1. **Motivation and informal summary of results.** In this paper, we develop a well-posedness and regularity theory for weak solutions of the hypoelliptic equation

$$(1.1) \qquad -\Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f = f^* \quad \text{in } \mathbb{T}^d \times \mathbb{R}^d.$$

The unknown function $f(x, v)$ is a function of the position variable $x \in \mathbb{T}^d$ and the velocity variable $v \in \mathbb{R}^d$. The PDE (1.1) is sometimes called the *Kramers equation*. We also consider the time-dependent version of this equation, namely

$$(1.2) \qquad \partial_t f - \Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f = f^* \quad \text{in } (0, \infty) \times \mathbb{T}^d \times \mathbb{R}^d,$$

which is often called the *kinetic Fokker-Planck equation*.

These equations were first studied by Kolmogorov [34] and were the main motivating examples for the general theory of Hörmander [31] of hypoelliptic equations. They are

of physical interest due to their relation with the *Langevin diffusion process* formally defined by

$$(1.3) \qquad \ddot{X} = \mathbf{b}(X) - \dot{X} + \dot{B},$$

where $\dot{X}$, $\ddot{X}$ stand respectively for the first and second time derivatives of X, a stochastic process taking values in $\mathbb{R}^d$, and $\dot{B}$ denotes a white noise process. Equation (1.3) can be interpreted as Newton's law of motion for a particle subject to the force field $\mathbf{b}(X)$, friction and thermal noise. This process can be recast as a Markovian evolution for the pair (X, V) evolving according to

$$\begin{cases} \dot{X} = -V, \\ \dot{V} = -\mathbf{b}(X) - V - \dot{B}. \end{cases}$$

The infinitesimal generator of this Markov process is the differential operator appearing on the left side of (1.1).

Kolmogorov [34] gave an explicit formula for the fundamental solution of (1.2) in the case $\mathbf{b} = 0$ and $U = \mathbb{R}^d$, which gives the existence of smooth solutions of (1.1) and (1.2) and implies that the operators on the left sides of (1.1) and (1.2) are *hypoelliptic*—that is, if f is a distributional solution of either of these equations and f^* is smooth, then f is also smooth. This result is extended to more general equations in Hörmander's celebrated paper [31], where he gave an essentially complete classification of hypoelliptic operators. In the case of the particular equations (1.1) and (1.2), his arguments yield a more systematic proof of Kolmogorov's results and, in particular, interior regularity estimates.

The study of hypoelliptic equations often falls back on the theory of pseudodifferential operators; see for example Kohn's proof [33] of Hörmander's classical result [31], which Hörmander includes in his monograph [32]. The purpose of this paper is rather to present a functional analytic and variational theory for (1.1) and (1.2) which has strong analogies to the familiar theory of uniformly elliptic equations. In particular, in this paper we:

- identify a function space H^1_{hyp} based on the natural energy estimates and develop a notion of weak solutions in this space;
- prove functional inequalities for H^1_{hyp}, for instance a Poincaré-type inequality, which implies uniform coercivity of our equations and holds not just on the spatial domain $\mathbb{T}^d$ but on any C^1 domain;
- develop a well-posedness theory of weak solutions based on the minimization of a uniformly convex functional;
- develop a regularity theory for weak solutions, based on an iteration of energy estimates, which implies that weak solutions are smooth;
- prove dissipative estimates for solutions of (1.2), using the coercivity of the variational structure, which imply an exponential decay to equilibrium.

Such a theory has until now remained undeveloped, despite the attention these equations have received in the last half century. The definition of the space H^1_{hyp} is not new: it and variants of it have been studied previously in the works [6, 40, 18]. However, the functional inequalities and other key properties which are required to work with this space are established here. A robust notion of weak solutions and corresponding well-posedness theory—besides allowing one to prove classical results for (1.1) and (1.2) in a different way—is important because it provides a natural framework for studying the stability of solutions (i.e., proving that a sequence of approximate solutions converges to a solution). In fact, it is just such an application—namely, developing a theory of homogenization for (1.2)—which motivated the present work. Furthermore, we expect that the theory developed here will provide a closer link between the hypoelliptic equations (1.1) and (1.2)

and the classical theory of uniformly elliptic and parabolic equations, allowing, for example, for a more systematic development of regularity estimates for solutions of the former by analogy to the latter. For instance, it would be interesting to investigate a possible connection between the functional-analytic framework proposed in this paper and the recent works [46, 47, 26, 39] which develop De Giorgi-Nash-type Hölder estimates for generalizations of the kinetic Fokker-Planck equations with measurable coefficients.[1]

In the first part of the paper, we address the well-posedness of (1.1) under a weak formulation based on the Sobolev-type space $H^1_{\mathrm{hyp}}(\mathbb{T}^d)$, defined below in (1.10). In the case in which $\mathbf{b}$ is a potential field, we provide two proofs of well-posedness. The first relies on the abstract Lax-Milgram theorem, while the second identifies a *uniformly convex* functional that has the sought-after weak solution as its unique minimizer. The identification of the correct convex functional is inspired by previous work of Brézis and Ekeland [14, 15] on variational formulations of parabolic equations (see also the more recent works [25, 4] and the references therein). The proof that our functional is coercive relies on a new Poincaré-type inequality for H^1_{hyp}, see Theorem 1.3 below. The Poincaré inequality in fact holds in a much more general setting than the periodic setting in which we consider (1.1). Our convex-analytic arguments for well-posedness can be immediately adapted to cover non-linear equations such as those obtained by replacing $\Delta_v f$ in (1.1) with $\nabla_v \cdot (\mathbf{a}(x, v, \nabla_v f))$, for $p \mapsto \mathbf{a}(x, v, p)$ a Lipschitz and uniformly maximal monotone operator (uniformly over $x \in \mathbb{T}^d$ and $v \in \mathbb{R}^d$).

Roughly speaking, the norm $\|\cdot\|_{H^1_{\mathrm{hyp}}(U)}$ is a measure of the size of the vector fields $\nabla_v f$ and $v \cdot \nabla_x f$, but crucially, the former is measured in a strong $L^2_x L^2_v$-type norm and the latter in a weaker $L^2_x H^{-1}_v$-type norm (see (1.10) below). The importance of measuring the vector fields $\nabla_v f$ and $v \cdot \nabla_x f$ using different norms also features prominently in other works including [13], but only spaces of positive regularity are considered there. Measuring the term $v \cdot \nabla_x f$ in a space of negative regularity in the v-variable is related to the idea of *velocity averaging*, the idea that one should expect better control of the spatial regularity of a solution of (1.1) or (1.2) after averaging in the velocity variable. This concept is therefore wired into the definition of the H^1_{hyp} norm, allowing us to perform velocity averaging in a systematic way. Once we have proved the existence of weak solutions to (1.1) in H^1_{hyp}, we are interested in showing that these solutions are in fact smooth. It is elementary to verify that the differential operators ∇_v and $v \cdot \nabla_x$ satisfy Hörmander's bracket condition, and therefore, as exposed in [31], a control of both $\nabla_v f$ and $v \cdot \nabla_x f$ in $L^2_x L^2_v$ would yield control of the seminorm of the function f in a fractional Sobolev space of positive regularity, namely $H^{1/2}_x L^2_v$. However, since the natural definition of the function space $H^1_{\mathrm{hyp}}(U)$ provides us only with control of $v \cdot \nabla_x f$ in a space of *negative* regularity in v, we are forced to revisit the arguments of [31]. A key step there is an interpolation-type inequality which converts the $L^2_x H^{-1}_v$ control on $v \cdot \nabla_x f$ (i.e., "velocity averaged" regularity) and $L^2_x H^1_v$ regularity on f into $L^2_x L^2_v$ regularity for a type of "fractional derivative" $(v \cdot \nabla_x)^{1/2} f$.[2] With this interpolation in hand, we then prove a functional inequality (see Theorem 1.4 below) which asserts that the $H^1_{\mathrm{hyp}}(U)$ norm controls exactly one-third of a derivative in arbitrary x-directions in the space $L^2_x L^2_v$

---

[1]We refer to works of Guerand and Imbert [27] and Anceschi and Rebucci [3], which appeared after the first version of the present paper.

[2]The analogous estimate for the heat equation is $f \in H^{1/2}_t L^2_x$.

in a weaker (Besov) sense, and almost one-third of a derivative in a stronger (Sobolev) sense. The one-third exponent is identical to that in Hörmander's paper and is sharp. [3]

Once we have proved that an arbitrary H^1_{hyp} function possesses at least a fractional derivative in the x variable, we are in a position to iterate the estimate by repeatedly differentiating the equation a fractional number of times to obtain higher regularity (and eventually smoothness, under appropriate assumptions on $\mathbf{b}$ and f^*) of weak solutions. In order to perform this iteration, we again depart from the original arguments of [31] and subsequent treatments and rely on an appropriate version of the Caccioppoli inequality (i.e., the basic L^2 energy estimate) for the equation (1.1). This avoids any recourse to sophisticated pseudodifferential operators and once again mimics the classical functional analytic arguments in the uniformly elliptic setting.

The developments described above and even the variational structure identified for the equation (1.1) are not restricted to the time-independent setting. Indeed, we show that they can be adapted in a very straightforward way to the kinetic Fokker-Planck equation (1.2), the main difference being that the first-order part in a "sum-of-squares" representation of the differential operator is now $\partial_t + v \cdot \nabla_x$ instead of just $v \cdot \nabla_x$. The adaptation thus consists in replacing the latter by the former throughout; the natural function space associated with equation (1.2), denoted by H^1_{kin}, is defined in (6.2)–(6.3). We also prove a Poincaré inequality for functions in H^1_{kin} which implies the uniform coercivity of the variational structure with respect to the H^1_{kin} norm. This allows us to give a rather direct and natural proof of exponential long-time decay to equilibrium for solutions of (1.2) with constant-in-time right-hand sides. This result (stated in Theorem 1.6 below) can be compared with the celebrated results of exponential convergence to equilibrium for kinetic Fokker-Planck equations on $\mathbb{R}^d$ with confining potentials, see in particular [19, 30, 29, 21, 20, 45, 7], as well as [43, 44] and references therein for a probabilistic approach. Compared to previous approaches, our proof of exponential convergence is once again closer to the classical dissipative argument for the heat equation based on differentiating the square of the spatial L^2 norm of the solution. Informally, our method is based on the idea that hypocoercivity is simply coercivity with respect to the correct norm.

1.2. **Statements of the main results.** We begin by introducing the Sobolev-type function space H^1_{hyp} associated with the equation (1.1). We let $U \subseteq \mathbb{R}^d$ either be a bounded C^1 domain with boundary, or we consider the boundary-less settings of $\mathbb{R}^d$ itself or the torus $\mathbb{T}^d$ with periodic boundary conditions. While we do not prove unique solvability in H^1_{hyp} of the Dirichlet problem in bounded C^1 domains, we nonetheless can prove the Poincaré inequality, so we study the two settings (with and without boundary)

---

[3]When translating Hörmander's work [31] into the present setting, the vector field is $X_0 = \partial_t + v \cdot \nabla_x$, and for simplicity we consider the "flat case" in which $X_1 = \nabla_v$. The regularity along X_0 is of index $1/2$, while the regularity along X_1 is of index 1. Then Hörmander's Theorem 4.3 gives regularity along the commutator $\nabla_x = [X_1, X_0]$ of index $1/3$, since $\frac{1}{1/3} = \frac{1}{1} + \frac{1}{1/2}$. In addition, the exponent $1/3$ arises naturally in the following way: consider $\partial_t f + v \cdot \nabla_x f - \varepsilon \Delta_v f = 0$ on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d$. Dimensionally speaking, $[f] = M$, $[x] = L$, $[v] = L/T$, and $[\varepsilon] = L^2/T^3$. The above PDE has a two-parameter scaling symmetry which keeps ε fixed, namely, $f \to \rho f(\lambda^{2/3} t, \lambda x, \lambda^{1/3} v)$, $\lambda, \rho > 0$. Here, ε is considered "dimensionless": $[\varepsilon] = 1$, that is, we identify $L^2 \sim T^3$. In this convention, the unique exponent α for which $\|(-\Delta)_x^{\alpha/2} f\|_{L^2_{t,x,v}}$ has the same dimensions as $\|\nabla_v f\|_{L^2_{t,x,v}}$ is $\alpha = 1/3$. Furthermore, the "flat case" is the formal limit of (1.2) upon "zooming in."

in tandem. We denote by γ the standard Gaussian measure on $\mathbb{R}^d$, defined by

$$(1.4) \qquad d\gamma(v) := (2\pi)^{-\frac{d}{2}} \exp\left(-\frac{1}{2}|v|^2\right) dv.$$

For each $p \in [1, \infty)$, we denote by $L_\gamma^p := L^p(\mathbb{R}^d, d\gamma)$ the Lebesgue space with norm

$$\|f\|_{L_\gamma^p} := \left(\int_{\mathbb{R}^d} |f(v)|^p \, d\gamma(v)\right)^{1/p},$$

and by H_γ^1 the Banach space with norm

$$\|f\|_{H_\gamma^1} := \left(\|f\|_{L_\gamma^2}^2 + \|\nabla f\|_{L_\gamma^2}^2\right)^{1/2}.$$

The dual space of H_γ^1 is denoted by H_γ^{-1}. By abuse of notation, we typically denote the canonical pairing $\langle \cdot, \cdot \rangle_{H_\gamma^1, H_\gamma^{-1}}$ between $f \in H_\gamma^1$ and $f^* \in H_\gamma^{-1}$ by

$$(1.5) \qquad \int_{\mathbb{R}^d} f f^* \, d\gamma := \langle f, f^* \rangle_{H_\gamma^1, H_\gamma^{-1}}.$$

Concerning the vector field $\mathbf{b}$, we shall often make the following assumption. Throughout the rest of the paper, we shall remind the reader when this assumption is in effect, or when we take more general vector fields $\mathbf{b}$.

Assumption 1.1. *There exists $W \in C^{0,1}(U; \mathbb{R})$ such that $\mathbf{b}(x) = -\nabla W(x)$ for almost every $x \in U$.*

Under the above assumption, we denote by $d\sigma$ the measure on U defined by

$$(1.6) \qquad d\sigma(x) := \exp(-W(x)) \, dx$$

and by dm the measure on $U \times \mathbb{R}^d$ defined by

$$(1.7) \qquad dm(x, v) := d\sigma(x) \, d\gamma(v) = \exp\left(-W(x) - \frac{1}{2}|v|^2\right) dx \, dv.$$

A consequence of this definition and integration by parts is the equality

$$(1.8) \qquad \iint_{\mathbb{T}^d \times \mathbb{R}^d} (v \cdot \nabla_x f(x, v) + \mathbf{b}(x) \cdot \nabla_v f(x, v)) \, dm = 0$$

for all smooth $\mathbb{T}^d$-periodic functions f.

Given $p \in [1, \infty)$, $U \subseteq \mathbb{R}^d$ and an arbitrary Banach space X, we denote by $L^p(U; X)$ the Banach space consisting of measurable functions $f : U \to X$ with norm

$$\|f\|_{L^p(U;X)} := \left(\int_U \|f(x, \cdot)\|_X^p \, dx\right)^{1/p}.$$

It will occasionally be convenient to consider the space $L_\sigma^p(U; X)$, which contains functions for which the norm

$$\|f\|_{L_\sigma^p(U;X)} := \left(\int_U \|f(x, \cdot)\|_X^p \, d\sigma\right)^{1/p}$$

is finite. Notice that, on bounded domains, the above norms induced by dx and $d\sigma$ are equivalent under Assumption 1.1.

We define the space $H_{\text{hyp}}^1(U)$ by

$$(1.9) \qquad H_{\text{hyp}}^1(U) := \left\{f \in L^2\left(U; H_\gamma^1\right) : v \cdot \nabla_x f \in L^2(U; H_\gamma^{-1})\right\}$$

and equip it with the norm

$$(1.10) \qquad \|f\|_{H_{\text{hyp}}^1(U)} := \left(\|f\|_{L^2(U;H_\gamma^1)}^2 + \|v \cdot \nabla_x f\|_{L^2(U;H_\gamma^{-1})}^2\right)^{1/2}.$$

When $\mathbf{b}$ satisfies Assumption 1.1, it is natural to define the H_{hyp}^1 norm with $\|v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f\|_{L_\sigma^2(U;H_\gamma^{-1})}$ replacing $\|v \cdot \nabla_x f\|_{L^2(U;H_\gamma^{-1})}$ in (1.9). The two norms are evidently equivalent on a bounded domain.

Given a bounded domain $U \subseteq \mathbb{R}^d$ and a vector field $\mathbf{b} \in L^\infty(U \times \mathbb{R}^d)^d$, we say that a function $f \in H_{\mathrm{hyp}}^1(U)$ is a *weak solution of* (1.1) *in* $U \times \mathbb{R}^d$ if

$$\forall h \in L^2(U;H_\gamma^1), \qquad \int_{U \times \mathbb{R}^d} \nabla_v h \cdot \nabla_v f \, dx \, d\gamma = \int_{U \times \mathbb{R}^d} h \left(f^* - v \cdot \nabla_x f - \mathbf{b} \cdot \nabla_v f \right) dx \, d\gamma \, .$$

As in (1.5), the precise interpretation of the right side is

$$(1.11) \qquad \int_U \langle h(x,\cdot), (f^* - v \cdot \nabla_x f - \mathbf{b} \cdot \nabla_v f)(x,\cdot) \rangle_{H_\gamma^1, H_\gamma^{-1}} \, dx \, .$$

As mentioned previously, we assume throughout that the domain $U \subseteq \mathbb{R}^d$ is bounded and has a C^1 boundary, or that $U = \mathbb{T}^d$ with periodic boundary conditions or $U = \mathbb{R}^d$. In the case $U \neq \mathbb{T}^d, \mathbb{R}^d$, we denote by $\mathbf{n}_U$ the outward-pointing unit normal to ∂U and define the *hypoelliptic boundary* of U by

$$\partial_{\mathrm{hyp}} U := \left\{ (x,v) \in \partial U \times \mathbb{R}^d \, : \, v \cdot \mathbf{n}_U(x) < 0 \right\} .$$

We denote by $H_{\mathrm{hyp},0}^1(U)$ the closure in $H_{\mathrm{hyp}}^1(U)$ of the set of smooth functions with compact support in $\overline{U} \times \mathbb{R}^d$ which vanish on $\partial_{\mathrm{hyp}} U$.

We give a first demonstration that $H_{\mathrm{hyp}}^1(U)$ is indeed the natural function space on which to build a theory of weak solutions of (1.1) by presenting a well-posedness result for the Kramers equation.

Theorem 1.2 (Well-posedness of the Kramers equation). *Let $\mathbf{b}$ satisfy Assumption 1.1, and let $f^* \in L^2(\mathbb{T}^d; H_\gamma^{-1})$ be such that $\iint_{\mathbb{T}^d \times \mathbb{R}^d} f^*(x,v) \, dm = 0$. Then there exists a unique weak solution $f \in H_{\mathrm{hyp}}^1(\mathbb{T}^d)$ to the Kramers equation*

$$(1.12) \qquad -\Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f = f^* \quad in \ \mathbb{T}^d \times \mathbb{R}^d$$

with $\iint_{\mathbb{T}^d \times \mathbb{R}^d} f(x,v) \, dm = 0$. Furthermore, there exists a constant $C(\mathbf{b}, d) < \infty$ such that f satisfies the estimate

$$(1.13) \qquad \|f\|_{H_{\mathrm{hyp}}^1(\mathbb{T}^d)} \leqslant C \|f^*\|_{L^2(\mathbb{T}^d; H_\gamma^{-1})} \, .$$

We next give an informal discussion regarding how one could naively guess that H_{hyp}^1 is the "correct" space for solving (1.1), and how our proof of Theorem 1.2 will work. We take the simpler case of matrix inversion in finite dimensions as a starting point. Given two matrices A and B with B skew-symmetric and a vector f^*, consider the problem of finding f such that

$$(1.14) \qquad (A^*A + B)f = f^*,$$

where A^* denotes the transpose of A. We propose to approach this problem by looking for a minimizer of the functional

$$f \mapsto \inf \left\{ \frac{1}{2}(Af - \mathbf{g}, Af - \mathbf{g}) \, : \, \mathbf{g} \ \text{such that} \ A^* \mathbf{g} = f^* - Bf \right\},$$

where $(\cdot, \cdot)$ denotes the underlying scalar product. It is clear that the infimum is non-negative, and if f is a solution to (1.14), then choosing $\mathbf{g} = Af$ shows that this infimum is actually zero (null). Moreover, since B is skew-symmetric, whenever $(f, \mathbf{g})$ satisfy the constraint in the infimum above, we have

$$(1.15) \qquad \frac{1}{2}(Af - \mathbf{g}, Af - \mathbf{g}) = \frac{1}{2}(Af, Af) + \frac{1}{2}(\mathbf{g}, \mathbf{g}) - (f, f^*).$$

The latter quantity is clearly a convex function of the pair $(f, \mathbf{g})$. The point is that under very mild assumptions on A and B, it will in fact be *uniformly* convex on the set of pairs $(f, \mathbf{g})$ satisfying the (linear) constraint $A^*\mathbf{g} = f^* - Bf$. Informally, the functional in (1.15) is coercive with respect to the seminorm $(f, \mathbf{g}) \mapsto |Af| + |\mathbf{g}| + |A(A^*A)^{-1}Bf|$.

With this analogy in mind, and assuming that $\mathbf{b}$ vanishes for simplicity, we rewrite the problem of finding a solution to (1.1) (with $\mathbf{b} \equiv 0$) as that of finding a null minimizer of the functional

$$(1.16) \qquad f \mapsto \inf\left\{ \int_{\mathbb{T}^d \times \mathbb{R}^d} \frac{1}{2}|\nabla_v f - \mathbf{g}|^2 \, dx \, d\gamma \ : \ \nabla_v^* \mathbf{g} = f^* - v \cdot \nabla_x f \right\},$$

where $\nabla_v^* F := -\nabla_v \cdot F + v \cdot F$ is the formal adjoint of ∇_v in L_γ^2. It is clear that the infimum above is non-negative, and if we are provided with a solution f to (1.1) (with $\mathbf{b} \equiv 0$), then choosing $\mathbf{g} = \nabla_v f$ reveals that this infimum vanishes at f. This functional gives strong credence to the definition of the space $H_{\mathrm{hyp}}^1(U)$ given in (1.9). Using convex-analytic arguments, we show that the mapping in (1.16) is uniformly convex, and that its infimum is null. This implies the well-posedness of the problem (1.1) with $\mathbf{b} \equiv 0$. The proof of coercivity relies on the following Poincaré-type inequality for $H_{\mathrm{hyp}}^1(U)$.

For every $f \in L^1(U; L_\gamma^1)$, we denote $(f)_U := |U|^{-1} \int_{U \times \mathbb{R}^d} f(x, v) \, d\sigma(x) \, d\gamma(v)$. For the purposes of the Poincaré inequality, we may set $U = \mathbb{T}^d$, or $U \subseteq \mathbb{R}^d$ a general C^1 domain. See Proposition 3.3 and [17] for an extension to the case $U = \mathbb{R}^d$ with a confining potential.

Theorem 1.3 (Poincaré inequality for H_{hyp}^1). *For $U = \mathbb{T}^d$ or $U \subseteq \mathbb{R}^d$ a general bounded C^1 domain, there exists a constant $C(U, d) < \infty$ such that for every $f \in H_{\mathrm{hyp}}^1(U)$, we have*

$$(1.17) \qquad \|f - (f)_U\|_{L^2(U; L_\gamma^2)} \leqslant C\left(\|\nabla_v f\|_{L^2(U; L_\gamma^2)} + \|v \cdot \nabla_x f\|_{L^2(U; H_\gamma^{-1})}\right).$$

Moreover, if in addition $f \in H_{\mathrm{hyp},0}^1(U)$, then we have

$$(1.18) \qquad \|f\|_{L^2(U; L_\gamma^2)} \leqslant C\left(\|\nabla_v f\|_{L^2(U; L_\gamma^2)} + \|v \cdot \nabla_x f\|_{L^2(U; H_\gamma^{-1})}\right).$$

The inequality (1.17) asserts that, up to an additive constant, the full $H_{\mathrm{hyp}}^1(U)$ norm of a function f is controlled by the seminorm

$$[\![f]\!]_{H_{\mathrm{hyp}}^1(U)} := \|\nabla_v f\|_{L^2(U; L_\gamma^2)} + \|v \cdot \nabla_x f\|_{L^2(U; H_\gamma^{-1})}.$$

In particular, any distribution f with $[\![f]\!]_{H_{\mathrm{hyp}}^1(U)} < \infty$ is actually a function, which moreover belongs to $L_x^2 L_\gamma^2$. The inequality (1.18) is a then simple extension which shows that for functions which vanish on the hypoelliptic boundary, the full H_{hyp}^1 norm is controlled by the seminorm.

The proof of Theorem 1.3 thus necessarily uses the Hörmander bracket condition, although in this case the way it is used is rather implicit. If we follow Hörmander's ideas more explicitly, then we obtain more information, namely some positive (fractional) regularity in the x variable. This is encoded in the following functional inequality, which we call the *Hörmander inequality*. The definitions of the fractional Sobolev spaces H^α used in the statement are given in Section 3.2, see (3.30). The Besov space $Q_{\nabla_x}^{1/3}(U)$ is defined in (2.13) in Section 2.3 and measures difference quotients in the spatial variable x of fractional order $1/3$.

Theorem 1.4 (Hörmander inequality for H_{hyp}^1). *Let $\alpha \in [0, \frac{1}{3})$, and let $U = \mathbb{T}^d$ or $U = \mathbb{R}^d$. There exists a constant $C(\alpha, d) < \infty$ such that, for every $f \in H_{\mathrm{hyp}}^1(U)$, we have the*

estimate

$$\|f\|_{H^{\alpha}(U;L^2_{\gamma})} \leqslant C \|f\|_{H^1_{\mathrm{hyp}}(U)} \,. \tag{1.19}$$

For $\alpha = 1/3$, we have the estimate

$$\|f\|_{Q^{1/3}_{\nabla_x}(U)} \leqslant C \|f\|_{H^1_{\mathrm{hyp}}(U)} \,. \tag{1.20}$$

The inequality (1.19) gives control over a norm with non-negative regularity in x and v. The estimate should be considered as an interior estimate in x; in other words, for U a general domain and any $f \in H^1_{\mathrm{hyp}}(U)$, we can apply the inequality (1.19) after multiplying f by a smooth cutoff function which vanishes for x near ∂U.

Our next main result asserts that weak solutions of (1.1) are actually smooth. This is accomplished by an argument which closely parallels the one for obtaining H^k regularity for solutions of uniformly elliptic equations. We first obtain a version of the Caccioppoli inequality, that is, a reverse Poincaré inequality, which states that the H^1_{hyp} seminorm of a solution of (1.1) can be controlled by its L^2 oscillation (see Lemma 5.1 for the precise statement). Combined with Theorem 1.4, this tells us that a fractional spatial derivative of a solution of (1.1) can be controlled by the L^2 oscillation of the function itself. This estimate can then be iterated: we repeatedly differentiate the equation a fractional amount to obtain estimates of the higher derivatives of the solution in the x variable; we then obtain estimates for derivatives in the v variable relatively easily.

Notice that the following statement implies that solutions of (1.1) are C^{∞} in both variables (x, v) provided that the vector field $\mathbf{b}$ is assumed to be smooth. For convenience, in the statement below we use the convention $C^{-1,1} = L^{\infty}$.

Theorem 1.5 (Interior Sobolev regularity for (1.1)). *Let $k \in \mathbb{N}$, $r \in (0, \infty)$ and $\mathbf{b} \in C^{k-1,1}(B_r \times \mathbb{R}^d; \mathbb{R}^d)$. There exists a constant $C < \infty$ depending on*

$$\left(d, k, r, \|\mathbf{b}\|_{C^{k-1,1}(B_r \times \mathbb{R}^d; \mathbb{R}^d)} \right)$$

such that, for every $f \in H^1_{\mathrm{hyp}}(B_r)$ and $f^* \in L^2(B_r; H^{-1}_{\gamma})$ satisfying

$$-\Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f = f^* \quad in \ B_r \times \mathbb{R}^d, \tag{1.21}$$

the following holds: If $\partial^{\alpha} f^* \in L^2(B_r; H^{-1}_{\gamma})$ for all multi-indices $\alpha \in \mathbb{N}^d \times \mathbb{N}^d$ with $|\alpha| \leqslant k$, then we have $\partial^{\alpha} f \in H^1_{\mathrm{hyp}}(B_{r/2})$ and the estimate

$$\|\partial^{\alpha} f\|_{H^1_{\mathrm{hyp}}(B_{r/2})} \leqslant C \left(\|f - (f)_{B_r}\|_{L^2(B_r; L^2_{\gamma})} + \sum_{|\beta| \leqslant k} \|\partial^{\beta} \widetilde{f^*}\|_{L^2(B_r; H^{-1}_{\gamma})} \right)$$

for all multi-indices $\alpha \in \mathbb{N}^d \times \mathbb{N}^d$ with $|\alpha| \leqslant k$.

The results stated above are for the time-independent Kramers equation (1.1). In Section 6, we develop an analogous theory for the time-dependent kinetic Fokker-Planck equation (1.2) with an associated function space H^1_{kin} (defined in (6.2)–(6.3)) in place of H^1_{hyp}. In particular, we obtain analogues of the results above for (1.2) which are stated in Section 6.

The long-time behavior of solutions of (1.2) has been studied by many authors in the last two decades: see the works of Desvillettes and Villani [19], Hérau and Nier [30], Helffer and Nier [29], Eckmann and Hairer [21], Desvillettes and Villani [20] and Villani [45] as well as the references in [45]. Most of these papers consider the case in which $\mathbf{b}(x) = -\nabla W(x)$ for a potential W which has sufficient growth at infinity, in which case dm is an explicit invariant measure, and solutions of (1.2) can be expected to converge exponentially fast to

the constant which is the integral of the initial data with respect to the invariant measure. This setting is in a certain sense easier than the Dirichlet problem, since one does not have to worry about the boundary. While our methods could also handle this setting, we formulate a result for the exponential convergence of a solution of the Cauchy-Dirichlet problem with constant-in-time right-hand side to the solution of the time-independent problem.

Theorem 1.6 (Convergence to equilibrium). *Let $U \subseteq \mathbb{R}^d$ be a C^1 domain and $\mathbf{b} \in L^\infty(U; C^{0,1}(\mathbb{R}^d))^d$. There exists $\lambda\left(\|\mathbf{b}\|_{L^\infty(U \times \mathbb{R}^d)}, U, d\right) > 0$ satisfying the following property. Let $f^* \in L^2(U; H_\gamma^{-1})$. Suppose that $f_\infty \in H_{\mathrm{hyp},0}^1(U)$ solves* (1.12)*, and that for every $T \in (0, \infty)$, $f \in H_{\mathrm{kin}}^1((0,T) \times U)$ solves*

$$
(1.22) \quad
\begin{cases}
\partial_t f - \Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f = f^* & \text{in } (0,T) \times U \times \mathbb{R}^d, \\
f = 0 & \text{on } (0,T) \times \partial_{\mathrm{hyp}} U,
\end{cases}
$$

where the boundary condition is satisfied in the sense that $f \in H_{\mathrm{kin},\|}^1((0,T) \times U)$.[4] Then, for every $t \geqslant 0$, we have

$$
(1.23) \quad \|f(t, \cdot) - f_\infty\|_{L^2(U; L_\gamma^2)} \leqslant 2 \exp(-\lambda t) \|f(0, \cdot) - f_\infty\|_{L^2(U; L_\gamma^2)}.
$$

Notice that interior regularity estimates immediately upgrade the L^2 convergence in (1.23) to convergence in spaces of higher regularity (at least in the interior) with the same exponential rate.

Unlike previous arguments establishing the exponential decay to equilibrium of solutions of (1.2) which are based on differentiation of perhaps non-transparent quantities involving the solution and several (possibly mixed) derivatives in both x and v, the proof of Theorem 1.6 we give here is elementary and close to the classical dissipative estimate for uniformly parabolic equations. The essential idea is to differentiate the square of the L^2 norm of the solution and then apply the Poincaré inequality. We cannot quite perform the computation exactly like this, and so we use a finite difference instead of the time derivative and apply a version of the Poincaré inequality adapted to the kinetic equation in a thin cylinder (see Proposition 6.2). Unlike previous approaches, our method therefore relates the positive constant λ in (1.23) to the optimal constant in a Poincaré-type inequality. One caveat of Theorem 1.6 is that, while we have a hypoelliptic Poincaré inequality in the above setting, we do not yet have a well-posedness theory in H_{kin}^1 except when $U = \mathbb{T}^d$.

Finally, we prove an enhanced dissipation estimate for solutions to the kinetic Fokker-Planck equation on the torus $\mathbb{T}^d$ with no right-hand side and $\mathbf{b} \equiv 0$ in a weakly collisional limit $\varepsilon \to 0^+$. The PDE satisfied by f when initial data f_{in} is given then becomes

$$
(1.24) \quad
\begin{cases}
\partial_t f + v \cdot \nabla_x f = \varepsilon \left(\Delta_v f - v \cdot \nabla_v f\right) & \text{in } (0, \infty) \times \mathbb{T}^d \times \mathbb{R}^d \\
f|_{t=0} = f_{\mathrm{in}}.
\end{cases}
$$

The spatial averages $f_{\mathrm{avg}}(t, v) := \int_{\mathbb{T}^d} f(t, x, v) dx$ satisfy

$$
(1.25) \quad \partial_t f_{\mathrm{avg}} = \varepsilon \left(\Delta_v f_{\mathrm{avg}} - v \cdot \nabla_v f_{\mathrm{avg}}\right)
$$

and decay only on the dissipative timescale $T_{\mathrm{d}} \sim \varepsilon^{-1}$, as can be seen by rescaling t in (1.25). In the setting of (1.24), enhanced dissipation is the observation that $f - f_{\mathrm{avg}}$ decays on the faster timescale $T_{\mathrm{e}} \sim \varepsilon^{-1/3}$:

---

[4]$H_{\mathrm{kin},\|}^1((0,T) \times U)$ is defined to be the closure of test functions $C^\infty([0,T]; U)$ vanishing on the lateral part of the hypoelliptic boundary, see subsection 6.5.

Theorem 1.7 (Enhanced dissipation)**.** *There exist constants $C(d) < \infty$ and $c(d) > 0$ such that for every $\varepsilon \in (0,1]$, initial data $f_{\mathrm{in}} \in L^2(\mathbb{T}^d; L^2_\gamma)$ satisfying*

$$(1.26) \qquad \int_{\mathbb{T}^d} f_{\mathrm{in}}(x,v) dx = 0 \qquad \forall v \in \mathbb{R}^d,$$

and for f the unique solution of (1.24) *constructed in Proposition* 6.10, *we have*

$$(1.27) \qquad \|f(t,\cdot,\cdot)\|_{L^2(\mathbb{T}^d; L^2_\gamma)} \leqslant C \, \|f_{\mathrm{in}}\|_{L^2(\mathbb{T}^d; L^2_\gamma)} \exp\left(-c\varepsilon^{-1/3}t\right).$$

When enhancement cannot be extracted directly from an explicit solution formula, it is often approached by hypocoercivity techniques, which were developed by Villani [45] in the context of kinetic theory; see also work of Guo [28]. These methods were adapted to the context of fluid dynamics in work of Beck and Wayne [8], Gallagher, Gallay, and Nier [24], and Bedrossian and Coti-Zelati [10]. In joint work of the first and last authors with Beekie [2], we demonstrated enhancement for solutions of certain advection-diffusion equations (passive scalars in shear flows) by methods which adhered more closely to Hörmander's original paper [31]. Theorem 1.7, which is inspired by [2], follows from an appropriate time- and ε-dependent version of the Hörmander inequality from Theorem 1.4.

In principle, one may also prove (1.27) with **b** satisfying Assumption 1.1, see Remark 6.14. It would be interesting to understand this method in the context of the Boltzmann and Landau equations.

1.3. **On unique solvability of the Dirichlet problem.** There is a subtle point in the analysis of the Dirichlet problem for (1.1) on general domains U which is due to the fact that we should prescribe the boundary condition only on part of the boundary, namely $\partial_{\mathrm{hyp}} U \coloneqq \{(x,v) \in \partial U \times \mathbb{R}^d : v \cdot \mathbf{n}_U(x) < 0\}$, where $\mathbf{n}_U$ denotes the outer normal to U. There is a difficulty coming from the possibly wild behavior of the trace of an H^1_{hyp} function near the *singular set* $\{(x,v) \in \partial U \times \mathbb{R}^d : v \cdot \mathbf{n}_U(x) = 0\}$, where particle trajectories graze the boundary. The following question remains open:[5]

Question 1.8. Does there exist $C(U,d) < \infty$ such that for every $f \in C^\infty_c(\overline{U} \times \mathbb{R}^d)$,

$$\int_{\partial U \times \mathbb{R}^d} f^2 \, |v \cdot \mathbf{n}_U| \, dx \, d\gamma \leqslant C \|f\|^2_{H^1_{\mathrm{hyp}}(U)} \quad ?$$

In the case of one spatial dimension ($d = 1$), this difficulty has been previously overcome and the well-posedness result was already proved in [6]. A generalization to higher dimensions was announced in [18], but we think that the argument given there is incomplete because the difficulty concerning the boundary behavior was not satisfactorily treated. This is explained in more detail in Appendix A of the original version [5] of the present work. A different way to phrase the main difficulty is discussed in Remark 4.3.

The original version [5] of this paper contained an error in the treatment of the Dirichlet and Cauchy-Dirichlet problems for the Kramers and kinetic Fokker-Planck equations, respectively.[6] We were unable to repair the proof, see Remark 4.3 below. In this version, we only prove unique solvability on the torus. *It remains an interesting open question whether unique solvability holds with boundary in the natural H^1_{hyp} class.*

In the intervening years, we succeeded in improving the results in other ways. Foremost, we sharpen the Hörmander-type inequality from $\alpha = 1/6-$ to $\alpha = 1/3-$ *without* cutoffs in the velocity variable. The second and third authors view this as a significant strengthening

---

[5]It is not difficult to define a pointwise a.e. trace away from the singular set, see Lemma 4.3 in the original version [5] of this paper on arXiv, but apparently this has limited usefulness.

[6]See two equations below (4.20) in the original version on arXiv ("Arguing as in for the last term in (4.19), ...").

of the paper, essentially due to the first and fourth authors. This allows us to prove enhanced relaxation to equilibrium, which was not contained in the first version of the paper. There have also been many works revisiting [31] and at least partially inspired by the first version, see [9, 11, 4, 27, 3, 16, 17, 35].

1.4. **Outline of the paper.** In the next section we present the function space $H^1_{\mathrm{hyp}}(U)$ and its important properties, as well as the Besov spaces used in the Hörmander inequality. In Section 3 we prove the functional inequalities stated in Theorems 1.3 and 1.4 and establish the compactness of the embedding of $H^1_{\mathrm{hyp}}(U)$ into $L^2(U; L^2_\gamma)$. In Section 4 we give two proofs of Theorem 1.2 on the well-posedness of the Dirichlet problem for the Kramers equation. The interior regularity of solutions, and in particular Theorem 1.5, is obtained in Section 5. Finally, in Section 6 we prove the analogous results for the kinetic Fokker-Planck equation (1.2) as well as the exponential decay to equilibrium (Theorem 1.6) and the enhancement estimate (Theorem 1.7).

2. Function space basics

In this section, we establish some basic properties of the function space $H^1_{\mathrm{hyp}}(U)$ defined in (1.9)–(1.10) and introduce several Besov-type spaces which will be necessary for the proof of the Hörmander inequality.

2.1. **Properties of H^1_γ and H^{-1}_γ.** We start by setting up some notation that will be used throughout the paper. We denote the formal adjoint of the operator ∇_v by ∇_v^*; that is, for every $F \in (H^1_\gamma)^d$, we denote

$$(2.1) \qquad \nabla_v^* F := -\nabla_v \cdot F + v \cdot F.$$

This definition can be extended to any $F \in (L^2_\gamma)^d$, in which case $\nabla_v^* F \in H^{-1}_\gamma$ and we have, for every $f \in H^1_\gamma$,

$$\int_{\mathbb{R}^d} f \, \nabla_v^* F \, d\gamma = \int_{\mathbb{R}^d} \nabla_v f \cdot F \, d\gamma \,.$$

Recall that the left side above is shorthand notation for the duality pairing between H^1_γ and H^{-1}_γ. We denote the average of a function $f \in L^1_\gamma$ by

$$(2.2) \qquad \langle f \rangle_\gamma := \int_{\mathbb{R}^d} f \, d\gamma \,.$$

Since $1 \in H^1_\gamma$, the definition of $\langle f \rangle_\gamma$ can be extended to arbitrary $f \in H^{-1}_\gamma$. The Gaussian Poincaré inequality states that, for every $f \in H^1_\gamma$,

$$\| f - \langle f \rangle_\gamma \|_{L^2_\gamma} \leqslant \| \nabla_v f \|_{L^2_\gamma} \,.$$

We can thus replace $\| f \|_{L^2_\gamma}$ by $\left| \langle f \rangle_\gamma \right|$ in the definition of H^1_γ and have an equivalent norm:

$$\left| \langle f \rangle_\gamma \right|^2 + \| \nabla f \|^2_{L^2_\gamma} \leqslant \| f \|^2_{H^1_\gamma} \leqslant 2 \left| \langle f \rangle_\gamma \right|^2 + 3 \| \nabla f \|^2_{L^2_\gamma} \,.$$

This comparison of norms has the following counterpart for the dual space H^{-1}_γ.

Lemma 2.1 (Identification of H^{-1}_γ). *There exists a universal constant $C < \infty$ such that for every $f^* \in H^{-1}_\gamma$,*

$$(2.3) \qquad C^{-1} \| f^* \|_{H^{-1}_\gamma} \leqslant \left| \langle f^* \rangle_\gamma \right| + \inf \left\{ \| \mathbf{h} \|_{L^2_\gamma} \, : \, \nabla_v^* \mathbf{h} = f^* - \langle f^* \rangle_\gamma \right\} \leqslant C \| f^* \|_{H^{-1}_\gamma} \,.$$

Proof. The bilinear form

$$(f, g) \mapsto \langle f \rangle_\gamma \langle g \rangle_\gamma + \int_{\mathbb{R}^d} \nabla_v f \cdot \nabla_v g \, d\gamma$$

is a scalar product for the Hilbert space H^1_γ. By the Riesz representation theorem, for every $f^* \in H^{-1}_\gamma$, there exists $g \in H^1_\gamma$ such that

$$\forall f \in H^1_\gamma \qquad \int_{\mathbb{R}^d} f f^* \, d\gamma = \langle f \rangle_\gamma \langle g \rangle_\gamma + \int_{\mathbb{R}^d} \nabla_v f \cdot \nabla_v g \, d\gamma .$$

(Recall that the integral on the left side is convenient notation for the canonical pairing between H^1_γ and H^{-1}_γ.) We clearly have $\langle g \rangle_\gamma = \langle f^* \rangle_\gamma$, and thus

$$|\langle g \rangle_\gamma|^2 + \int_{\mathbb{R}^d} |\nabla_v g|^2 \, d\gamma \leqslant \|g\|_{H^1_\gamma} \|f^*\|_{H^{-1}_\gamma} .$$

This implies that $\|\nabla_v g\|_{L^2_\gamma} \leqslant C \|f^*\|_{H^{-1}_\gamma}$, and since $\nabla^*_v \nabla_v g = f^* - \langle f^* \rangle_\gamma$, this proves the rightmost inequality in (2.3). Conversely, for any $\mathbf{h} \in L^2_\gamma$, if

$$f^* = \langle f^* \rangle_\gamma + \nabla^*_v \mathbf{h} ,$$

then for every $f \in H^1_\gamma$,

$$\left| \int_{\mathbb{R}^d} f f^* \, d\gamma \right| \leqslant |\langle f \rangle_\gamma| \, |\langle f^* \rangle_\gamma| + \|\nabla f\|_{L^2_\gamma} \|\mathbf{h}\|_{L^2_\gamma} ,$$

and thus the leftmost inequality in (2.3) holds. $\qquad\square$

We often work with the dual pair of Banach spaces $L^2(U; H^1_\gamma)$ and $L^2(U; H^{-1}_\gamma)$. With the identification given by Lemma 2.1, we have

$$(2.4) \qquad \|f^*\|_{L^2(U;H^{-1}_\gamma)} \simeq \left\| \langle f^* \rangle_\gamma \right\|_{L^2(U)} + \inf \left\{ \|\mathbf{g}\|_{L^2(U;L^2_\gamma)} : \nabla^*_v \mathbf{g} = f^* - \langle f^* \rangle_\gamma \right\},$$

in the sense that the norms on each side are equivalent.

For convenience, for every $f \in L^1(U; L^1_\gamma)$, we use the shorthand notation

$$(2.5) \qquad (f)_U := |U|^{-1} \int_{U \times \mathbb{R}^d} f(x, v) \, d\sigma(x) \, d\gamma(v).$$

We will occasionally also use this notation in the case when f depends only on the space variable x, in which case we simply have $(f)_U = |U|^{-1} \int_U f \, d\sigma(x)$.

In the proof of the Hörmander inequality, it will be beneficial to understand which type of finite differences are controlled by $\|f\|_{H^1_\gamma}$. Recall that

$$d\gamma(v) := (2\pi)^{-\frac{d}{2}} \exp\left(-\frac{1}{2}|v|^2\right) dv .$$

The fundamental issue is that $\gamma(\cdot + h)$ is not comparable to γ, above and below, uniformly in v. For instance, while the translation of the measure γ by a fixed vector $y \in \mathbb{R}^d$ is absolutely continuous with respect to γ, the associated Radon-Nikodym derivative is unbounded (unless $y = 0$). This distinguishes Gaussians from $e^{-\langle x \rangle}$, for example, and changes the finite difference characterization of the space

$$\|\nabla_v u\|_{L^2(U;L^2_\gamma)} ,$$

since its finite difference characterization is not in the seminorm

$$\sup_{h>0} h^{-1} \|u(x, v+h) - u(x, v)\|_{L^2(U;L^2_\gamma)} .$$

Towards an appropriate characterization, we first note that a consequence of the logarithmic Sobolev inequality and the Gaussian Poincaré inequality is the estimate

$$(2.6) \qquad \||v|u\|_{L^2(U;L^2_\gamma)} \lesssim \|\nabla_v u\|_{L^2(U;L^2_\gamma)}$$

for functions u satisfying $\langle u \rangle_\gamma = 0$; the reader may consult (3.35) and the ensuing discussion for details. The inequality (2.6), together with the product rule, gives that

$$\left\|\nabla_v(u\gamma^{1/2})\right\|_{L^2(U;L^2(\mathbb{R}^d))} \lesssim \|\nabla_v u\|_{L^2(U;L^2_\gamma(\mathbb{R}^d))},$$

and since the left-hand side has a finite difference characterization, we have
(2.7)
$$\sup_{h \in \mathbb{R}^d \smallsetminus \{0\}} |h|^{-1} \left\| u(x, v+h)\gamma^{1/2}(v+h) - u(x,v)\gamma^{1/2}(v) \right\|_{L^2(U;L^2(\mathbb{R}^d))} \lesssim \|\nabla_v u\|_{L^2(U;L^2_\gamma)}.$$

We refer to Lunardi [36] for further discussion.

2.2. Density of Smooth Functions in H^1_{hyp}.

We show that the set of smooth functions is dense in H^1_{hyp}.

Proposition 2.2. *The set $C^\infty_c(\overline{U} \times \mathbb{R}^d)$ of smooth functions with compact support in $\overline{U} \times \mathbb{R}^d$ is dense in $H^1_{\mathrm{hyp}}(U)$.*

Proof. We focus on the case when $U \subseteq \mathbb{R}^d$ is a bounded C^1 domain. When $U = \mathbb{T}^d$, the proof can be done more simply by cutting off in v and mollifying.

We decompose the proof into three steps.

Step 1. In this step, we show that it suffices to consider the case when U satisfies a convenient quantitative form of the star-shape property. For every $z \in \partial U$, there exist a radius $r > 0$ and a C^1 function $\Psi \in C^1(\mathbb{R}^{d-1}; \mathbb{R})$ such that, up to a relabelling of the axes, we have

$$U \cap B(z,r) = \left\{ x = (x_1, \ldots, x_d) \in B(z,r) \; : \; x_d > \Psi(x_1, \ldots, x_{d-1}) \right\}.$$

Since Ψ is a C^1 function, there exists $\delta > 0$ such that for every $x \in U \cap B(z,r)$, we have the cone containment property

$$(2.8) \qquad \left\{ x + y \; : \; \frac{y_d}{|y|} \geq 1 - \delta \right\} \cap B(z,r) \subseteq U.$$

Setting

$$z' = z + \left(0, \ldots, 0, \frac{r}{2} \right) \in \mathbb{R}^d,$$

and reducing $\delta > 0$ if necessary, we claim that for every $x \in U \cap B(z, \delta^2)$ and $\varepsilon \in (0,1]$, we have

$$(2.9) \qquad B\left(x - \varepsilon(x - z'), \delta^2 \varepsilon \right) \subseteq U.$$

Assuming the contrary, let $y \in \mathbb{R}^d$ be such that

$$x + y \in B\left(x - \varepsilon(x - z'), \delta^2 \varepsilon \right) \smallsetminus U.$$

Then

$$\left| y + \varepsilon(x - z') \right| \leq \delta^2 \varepsilon,$$

and therefore

$$
\begin{aligned}
\left| y - \varepsilon \left(0, \ldots, 0, \frac{r}{2} \right) \right| &\leqslant \left| y + \varepsilon(x - z) - \varepsilon \left(0, \ldots, 0, \frac{r}{2} \right) \right| + \varepsilon |x - z| \\
&\leqslant \left| y + \varepsilon(x - z') \right| + \varepsilon |x - z| \\
&\leqslant 2\delta^2 \varepsilon.
\end{aligned}
$$

Taking $\delta > 0$ sufficiently small, we arrive at a contradiction with the cone property (2.8). Now that (2.9) is proved for every x in a relative neighborhood of z, and up to a further reduction of the value of $\delta > 0$ if necessary, it is not difficult to show that one can find an open set U' containing z and z' and such that (2.9) holds for every $x \in U \cap U'$.

Summarizing, and using the fact that U is a bounded set, we have shown that there exist families of bounded open sets $U_1, \ldots, U_M \subseteq \mathbb{R}^d$, of points $x_1, \ldots, x_M \in \mathbb{R}^d$ and a parameter $r > 0$ such that

$$
U = \bigcup_{k=1}^{M} U_i
$$

and for every $k \in \{1, \ldots, M\}$, $x \in U_k$ and $\varepsilon \in (0, 1]$,

$$
B\left(x - \varepsilon(x - x_k), r\varepsilon \right) \subseteq U_k.
$$

By using a partition of unity, we can reduce our study to the case when this property is satisfied for the domain U itself (in place of each of the U_k's). By translation, we may assume that the reference point x_k is at the origin, and by scaling, we may also assume that this property holds with $r = 1$. That is, from now on, we assume that for every $x \in U$ and $\varepsilon \in (0, 1]$, we have

$$
(2.10) \qquad\qquad B\left((1 - \varepsilon)x, \varepsilon \right) \subseteq U.
$$

Step 2. Let $f \in H^1_{\mathrm{hyp}}(U)$. We aim to show that f belongs to the closure of the set $C_c^\infty(\overline{U} \times \mathbb{R}^d)$ in $H^1_{\mathrm{hyp}}(U)$. Without loss of generality, we may assume that f is compactly supported in $\overline{U} \times \mathbb{R}^d$. Indeed, if $\chi \in C_c^\infty(\mathbb{R}^d; \mathbb{R})$ is a smooth function with compact support and such that $\chi \equiv 1$ in a neighborhood of the origin, then the function $(x, v) \mapsto f(x, v)\chi(v/M)$ belongs to $H^1_{\mathrm{hyp}}(U)$ and converges to f in $H^1_{\mathrm{hyp}}(U)$ as M tends to infinity.

Let $\zeta \in C_c^\infty(\mathbb{R}^d; \mathbb{R})$ be a smooth function with compact support in $B(0, 1)$ and such that $\int_{\mathbb{R}^d} \zeta = 1$. For each $\varepsilon > 0$ and $x \in \mathbb{R}^d$, we write

$$
(2.11) \qquad\qquad \zeta_\varepsilon(x) := \varepsilon^{-d} \zeta(\varepsilon^{-1} x),
$$

and we define, for each $\varepsilon \in \left(0, \frac{1}{2} \right]$, $x \in U$ and $v \in \mathbb{R}^d$,

$$
f_\varepsilon(x, v) := \int_{\mathbb{R}^d} f\left((1 - \varepsilon)x + y, v \right) \zeta_\varepsilon(y) \, dy.
$$

Note that this definition makes sense by the assumption of (2.10). The goal of this step is to show that f belongs to the closure in $H^1_{\mathrm{hyp}}(U)$ of the convex hull of the set $\left\{ f_\varepsilon : \varepsilon \in \left(0, \frac{1}{2} \right] \right\}$. By Mazur's lemma (see [22, page 6]), it suffices to show that f_ε converges weakly to f in $H^1_{\mathrm{hyp}}(U)$. Since it is elementary to show that f_ε converges to f in the sense of distributions, this boils down to checking that f_ε is bounded in $H^1_{\mathrm{hyp}}(U)$. By Jensen's inequality,

$$
\begin{aligned}
\| \nabla_v f_\varepsilon \|^2_{L^2(U; L^2_\gamma)} &\leqslant \int_{U \times \mathbb{R}^d} \int_{\mathbb{R}^d} |\nabla_v f|^2 \left((1 - \varepsilon)x + y, v \right) \zeta_\varepsilon(y) \, dy \, dx \, d\gamma(v) \\
&\leqslant (1 - \varepsilon)^{-1} \| \nabla_v f \|^2_{L^2(U; L^2_\gamma)}.
\end{aligned}
$$

In order to evaluate $\|v \cdot \nabla_x f_\varepsilon\|_{L^2(U;H_\gamma^{-1})}$, we compute, for every $\varphi \in L^2(U;H_\gamma^1)$,

$$\int_{U \times \mathbb{R}^d} v \cdot \nabla_x f_\varepsilon \, \varphi \, dx \, d\gamma$$

$$= (1-\varepsilon) \int_{U \times \mathbb{R}^d} \int_{\mathbb{R}^d} v \cdot \nabla_x f\left((1-\varepsilon)x + y, v\right) \zeta_\varepsilon(y) \varphi(x,v) \, dy \, dx \, d\gamma(v)$$

$$= \int_{U \times \mathbb{R}^d} \int_{\mathbb{R}^d} v \cdot \nabla_x f\left(x + y, v\right) \zeta_\varepsilon(y) \, \varphi\left(\frac{x}{1-\varepsilon}, v\right) dy \, dx \, d\gamma(v)$$

$$= \int_{U \times \mathbb{R}^d} \int_{\mathbb{R}^d} v \cdot \nabla_x f\left(y, v\right) \zeta_\varepsilon(y - x) \, \varphi\left(\frac{x}{1-\varepsilon}, v\right) dy \, dx \, d\gamma(v).$$

Since, by Jensen's inequality,

$$\int_{U \times \mathbb{R}^d} \left| \int_U \zeta_\varepsilon(y-x) \varphi\left(\frac{x}{1-\varepsilon}, v\right) dx \right|^2 dy \, d\gamma(v) \leqslant (1-\varepsilon)^{-1} \|\varphi\|^2_{L^2(U;L_\gamma^2)}$$

as well as

$$\int_{U \times \mathbb{R}^d} \left| \int_U \zeta_\varepsilon(y-x) \nabla_v \varphi\left(\frac{x}{1-\varepsilon}, v\right) dx \right|^2 dy \, d\gamma(v) \leqslant (1-\varepsilon)^{-1} \|\nabla_v \varphi\|^2_{L^2(U;L_\gamma^2)},$$

we deduce that

$$\int_{U \times \mathbb{R}^d} v \cdot \nabla_x f_\varepsilon \, \varphi \, dx \, d\gamma \leqslant (1-\varepsilon)^{-\frac{1}{2}} \|v \cdot \nabla_x f\|_{L^2(U;H_\gamma^{-1})} \|\varphi\|_{L^2(U;H_\gamma^1)},$$

and therefore

$$\|v \cdot \nabla_x f_\varepsilon\|_{L^2(U;H_\gamma^{-1})} \leqslant (1-\varepsilon)^{-\frac{1}{2}} \|v \cdot \nabla_x f\|_{L^2(U;H_\gamma^{-1})}.$$

This completes the proof that the set $\left\{ f_\varepsilon \, : \, \varepsilon \in \left(0, \frac{1}{2}\right] \right\}$ is bounded in $H_{\mathrm{hyp}}^1(U)$, and thus that f belongs to the closed convex hull of this set.

Step 3. It remains to be shown that for each fixed $\varepsilon \in \left(0, \frac{1}{2}\right]$, the function f_ε belongs to the closure in $H_{\mathrm{hyp}}^1(U)$ of the set $C_c^\infty(\overline{U} \times \mathbb{R}^d)$. For every $\eta \in (0,1]$, we define

$$f_{\varepsilon,\eta}(x,v) := \int_{\mathbb{R}^d} f_\varepsilon(x,w) \zeta_\eta(v-w) \, dw$$

$$= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(y,w) \zeta_\varepsilon(y - (1-\varepsilon)x) \zeta_\eta(v-w) \, dy \, dw.$$

From the last expression, we see that $f_{\varepsilon,\eta}$ belongs to $C_c^\infty(\overline{U} \times \mathbb{R}^d)$ (recall that f itself has compact support in $\overline{U} \times \mathbb{R}^d$). Moreover, since $\nabla_v f_\varepsilon \in L^2(U;L_\gamma^2)$ and

$$\nabla_v f_{\varepsilon,\eta}(x,v) = \int_{\mathbb{R}^d} \nabla_v f_\varepsilon(x, v-w) \zeta_\eta(w) \, dw,$$

it is classical to verify that $\nabla_v f_{\varepsilon,\eta}$ converges to $\nabla_v f_\varepsilon$ in $L^2(U;L_\gamma^2)$ as η tends to 0. By the definition of f_ε and the fact that f_ε is compactly supported, we have that $v \cdot \nabla_x f_\varepsilon \in L^2(U;L_\gamma^2)$. The same reasoning as above thus gives that $v \cdot \nabla_x f_{\varepsilon,\eta}$ converges to $v \cdot \nabla_x f_\varepsilon$ in $L^2(U;L_\gamma^2)$, and thus a fortiori in $L^2(U;H_\gamma^{-1})$, as η tends to 0. This shows that

$$\lim_{\eta \to 0} \|f_{\varepsilon,\eta} - f_\varepsilon\|_{H_{\mathrm{hyp}}^1(U)} = 0$$

and thus completes the proof of the proposition. $\qquad\square$

2.3. **Besov Spaces.** We shall use the following Besov-type spaces in the proof of the Hörmander inequality. The first of these spaces measures fractional regularity along the vector field $v \cdot \nabla_x$, while the second measures fractional regularity along ∇_x. As the Hörmander inequality is an interior estimate, we only consider these spaces in the cases that $U = \mathbb{R}^d$ or $U = \mathbb{T}^d$. To lighten the notation, we may frequently write $\|\cdot\|_{Q_{v \cdot \nabla_x}^{1/2}}$ rather than $\|\cdot\|_{Q_{v \cdot \nabla_x}^{1/2}(U)}$, as the choice of $U = \mathbb{R}^d, \mathbb{T}^d$ plays no role in the argument. The Q stands for "quotient."

Definition 2.3. For measurable $f : U \times \mathbb{R}^d \to \mathbb{R}$, we define

$$(2.12) \qquad \|f\|_{Q_{v \cdot \nabla_x}^{1/2}(U)}^2 := \sup_{0 < \eta < \infty} \frac{1}{\eta^2} \iint_{U \times \mathbb{R}^d} \left(f(x + \eta^2 v, v) - f(x, v) \right)^2 d\gamma(v) \, dx.$$

Definition 2.4. For measurable $f : U \times \mathbb{R}^d \to \mathbb{R}$, we define

$$(2.13) \qquad \|f\|_{Q_{\nabla_x}^{1/3}(U)}^2 := \sup_{\substack{0 < \eta < \infty \\ x' \in \mathbb{S}^{d-1}}} \frac{1}{\eta^2} \iint_{U \times \mathbb{R}^d} \left(f(x + \eta^3 x', v) - f(x, v) \right)^2 d\gamma(v) \, dx.$$

3. FUNCTIONAL INEQUALITIES FOR H_{hyp}^1

In this section we present the proofs of Theorems 1.3 and 1.4.

3.1. **The Poincaré inequality for H_{hyp}^1.** We begin with the proof of Theorem 1.3, the Poincaré-type inequality for the space $H_{\text{hyp}}^1(U)$. The proof requires the following fact regarding the equivalence (up to additive constants) of the norms $\|h\|_{L^2(U)}$ and $\|\nabla h\|_{H^{-1}(U)}$.

Lemma 3.1. *Let U be a Lipschitz domain or $U = \mathbb{T}^d$. Then there exists $C(U, d) < \infty$ such that for every $h \in L^2(U)$,*

$$\|h - (h)_U\|_{L^2(U)} \leqslant C \|\nabla h\|_{H^{-1}(U)}.$$

Proof. We begin by considering the case U is a Lipschitz domain. Without loss of generality, we assume that $(h)_U = 0$. We consider the problem

$$(3.1) \qquad \begin{cases} \nabla \cdot \mathbf{f} = h & \text{in } U, \\ \mathbf{f} = 0 & \text{on } \partial U. \end{cases}$$

Bogovskii's operator [12] (see also Galdi's book [23, Section III.3]) guarantees the existence of a solution $\mathbf{f}$ with components in $H_0^1(U)$ satisfying the estimate

$$(3.2) \qquad \|\mathbf{f}\|_{H^1(U)} \leqslant C \|h\|_{L^2(U)}.$$

Then we have

$$\|h\|_{L^2(U)}^2 = \int_U h \, \nabla \cdot \mathbf{f} = -\int_U \nabla h \cdot \mathbf{f} \leqslant \|\nabla h\|_{H^{-1}(U)} \|\mathbf{f}\|_{H^1(U)}.$$

The conclusion then follows by (3.2). In the case $U = \mathbb{T}^d$, the estimate follows from classical Littlewood-Paley estimates, and we omit the details. $\qquad \square$

Proof of Theorem 1.3. Let $f \in H_{\text{hyp}}^1(U)$. In view of Proposition 2.2, we can without loss of generality assume that f is a smooth function. We decompose the proof into five steps.

Step 1. We show that

$$(3.3) \qquad \|f - \langle f \rangle_\gamma\|_{L^2(U; L_\gamma^2)} \leqslant \|\nabla_v f\|_{L^2(U; L_\gamma^2)}.$$

By the Gaussian Poincaré inequality, we have for every $x \in U$ that

$$\|f(x, \cdot) - \langle f \rangle_\gamma(x)\|_{L^2_\gamma} \leqslant \|\nabla_v f(x, \cdot)\|_{L^2_\gamma} .$$

This yields (3.3) after integration over $x \in U$.

Step 2. We show that

$$(3.4) \qquad \|\nabla \langle f \rangle_\gamma\|_{H^{-1}(U)} \leqslant C \left(\|\nabla_v f\|_{L^2(U; L^2_\gamma)} + \|v \cdot \nabla_x f\|_{L^2(U; H^{-1}_\gamma)} \right).$$

We select $\xi_1, \ldots, \xi_d \in C^\infty_c(\mathbb{R}^d)$ satisfying

$$(3.5) \qquad \int_{\mathbb{R}^d} v \xi_i(v) \, d\gamma(v) = e_i,$$

and for each test function $\phi \in H^1_0(U)$ and $i \in \{1, \ldots, d\}$, we compute

$$\int_U \partial_{x_i} \phi(x) \langle f \rangle_\gamma(x) \, dx = \int_{U \times \mathbb{R}^d} v \cdot \nabla_x \phi(x) \langle f \rangle_\gamma(x) \xi_i(v) \, dx \, d\gamma(v)$$
$$= \int_{U \times \mathbb{R}^d} v \cdot \nabla_x \phi(x) f(x, v) \xi_i(v) \, dx \, d\gamma(v)$$
$$+ \int_{U \times \mathbb{R}^d} v \cdot \nabla_x \phi(x) \left(f(x, v) - \langle f \rangle_\gamma(x) \right) \xi_i(v) \, dx \, d\gamma(v).$$

To control the first term on the right side, we perform an integration by parts to obtain

$$\left| \int_{U \times \mathbb{R}^d} v \cdot \nabla_x \phi(x) f(x, v) \xi_i(v) \, dx \, d\gamma(v) \right| = \left| \int_{U \times \mathbb{R}^d} \phi(x) \xi_i(v) \, v \cdot \nabla_x f(x, v) \, dx \, d\gamma(v) \right|$$
$$\leqslant C \|\phi \xi_i\|_{L^2(U; H^1_\gamma)} \|v \cdot \nabla_x f\|_{L^2(U; H^{-1}_\gamma)}$$
$$\leqslant C \|\phi\|_{L^2(U)} \|\xi_i\|_{H^1_\gamma} \|v \cdot \nabla_x f\|_{L^2(U; H^{-1}_\gamma)}$$
$$\leqslant C \|\phi\|_{L^2(U)} \|v \cdot \nabla_x f\|_{L^2(U; H^{-1}_\gamma)} .$$

To control the second term, we use (3.3) and the fact that ξ_i has compact support:

$$\left| \int_{U \times \mathbb{R}^d} v \cdot \nabla_x \phi(x) \left(f(x, v) - \langle f \rangle_\gamma(x) \right) \xi_i(v) \, dx \, d\gamma(v) \right|$$
$$\leqslant C \int_{U \times \mathbb{R}^d} |v| |\xi_i(v)| |\nabla_x \phi(x)| |f(x, v) - \langle f \rangle_\gamma(x)| \, dx \, d\gamma(v)$$
$$\leqslant C \|\phi\|_{H^1(U)} \|\nabla_v f\|_{L^2(U; L^2_\gamma)} .$$

Combining the above displays and taking the supremum over $\phi \in H^1_0(U)$ with $\|\phi\|_{H^1(U)} \leqslant 1$ yields (3.4).

Step 3. We deduce from Lemma 3.1, (3.3) and (3.4) that

$$\|f - (f)_U\|_{L^2(U; L^2_\gamma)} \leqslant \|f - \langle f \rangle_\gamma\|_{L^2(U; L^2_\gamma)} + \|\langle f \rangle_\gamma - (f)_U\|_{L^2(U)}$$
$$\leqslant \|f - \langle f \rangle_\gamma\|_{L^2(U; L^2_\gamma)} + C \|\nabla \langle f \rangle_\gamma\|_{H^{-1}(U)}$$
$$\leqslant C \left(\|\nabla_v f\|_{L^2(U; L^2_\gamma)} + \|v \cdot \nabla_x f\|_{L^2(U; H^{-1}_\gamma)} \right).$$

This completes the proof of (1.17).

Step 4. The remaining steps are specific to the case with boundary. To complete the proof of (1.18), we must show that, under the additional assumption that $U \neq \mathbb{T}^d$ and $f \in H^1_{\mathrm{hyp},0}(U)$, we have

$$(3.6) \qquad |(f)_U| \leqslant C \left(\|\nabla_v f\|_{L^2(U; L^2_\gamma)} + \|v \cdot \nabla_x f\|_{L^2(U; H^{-1}_\gamma)} \right).$$

Let f_1 be a test function belonging to $C_c^\infty\left(\overline{U} \times \mathbb{R}^d\right)$, to be constructed below, which satisfies the following:

$$(3.7) \qquad f_1 = 0 \quad \text{on } (\partial U \times \mathbb{R}^d) \smallsetminus \partial_{\mathrm{hyp}}(U),$$

$$(3.8) \qquad \fint_U \int_{\mathbb{R}^d} v \cdot \nabla_x f_1 \, d\gamma \, dx = 1$$

and, for some constant $C(U, d) < \infty$,

$$(3.9) \qquad \|v \cdot \nabla_x f_1\|_{L^2(U; L_\gamma^2)} \leqslant C.$$

The test function f_1 is constructed in Step 5 below. We first use it to obtain (3.6). We proceed by using (3.8) to split the mean of f as

$$(f)_U = \fint_U \int_{\mathbb{R}^d} f \, v \cdot \nabla_x f_1 \, d\gamma \, dx - \fint_U \int_{\mathbb{R}^d} (f - (f)_U) \, v \cdot \nabla_x f_1 \, d\gamma \, dx$$

and estimate the two terms on the right side separately. For the first term, we have

$$\left| \fint_U \int_{\mathbb{R}^d} f \, v \cdot \nabla_x f_1 \, d\gamma \, dx \right|$$

$$= \left| -\fint_U \int_{\mathbb{R}^d} f_1 \, v \cdot \nabla_x f \, d\gamma \, dx + \frac{1}{|U|} \int_{\partial U} \int_{\mathbb{R}^d} (v \cdot \mathbf{n}_U) f f_1 \, d\gamma \, dx \right|$$

$$= \left| \fint_U \int_{\mathbb{R}^d} f_1 v \cdot \nabla_x f \, d\gamma \, dx \right|,$$

where we used that $(v \cdot \mathbf{n}_U) f f_1$ vanishes on $\partial U \times \mathbb{R}^d$ to remove the boundary integral. (Recall that by the definition of $H_{\mathrm{hyp},0}^1(U)$, we can assume without loss of generality that the function f is smooth, so the justification of the integration by parts above is classical.) We thus obtain that

$$\left| \fint_U \int_{\mathbb{R}^d} f_1 v \cdot \nabla_x f \, d\gamma \, dx \right| \leqslant \frac{1}{|U|} \|f_1\|_{L^2(U; H_\gamma^1)} \|v \cdot \nabla_x f\|_{L^2(U; H_\gamma^{-1})} \, .$$

This completes the estimate for the first term. For the second term, we use (3.9) to get

$$\left| \fint_U \int_{\mathbb{R}^d} (f - (f)_U) \, v \cdot \nabla_x f_1 \, d\gamma \, dx \right| \leqslant \|f - (f)_U\|_{L^2(U; L_\gamma^2)} \|v \cdot \nabla_x f_1\|_{L^2(U; L_\gamma^2)}$$

$$\leqslant C \|f - (f)_U\|_{L^2(U; L_\gamma^2)} \, ,$$

which is estimated using the result of Step 3. Putting these together yields (3.6).

Step 5. We construct the test function $f_1 \in C_c^\infty\left(\overline{U} \times \mathbb{R}^d\right)$ satisfying (3.7), (3.8) and (3.9). Fix $x_0 \in \partial U$ where $\mathbf{n}_U(x_0)$ is well defined. Since the unit normal $\mathbf{n}_U$ is continuous at x_0, there exist $v_0 \in \mathbb{R}^d$ and $r > 0$ such that for every $x, v \in \mathbb{R}^d$ satisfying $(x, v) \in (B_r(x_0) \cap \partial U) \times B_r(v_0)$, we have $v \cdot \mathbf{n}_U(x) > 0$. In other words, every $(x, v) \in (B_r(x_0) \cap \partial U) \times B_r(v_0)$ is such that $(x, v) \in \partial_{\mathrm{hyp}} U$. Observe that, for every $f_1 \in C_c^\infty(\mathbb{R}^d \times \mathbb{R}^d)$, we have

$$\fint_U \int_{\mathbb{R}^d} v \cdot \nabla_x f_1 \, d\gamma \, dx = \frac{1}{|U|} \int_{\partial U} \int_{\mathbb{R}^d} (v \cdot \mathbf{n}_U) f_1 \, d\gamma \, dx.$$

We select a function $f_1 \in C_c^\infty(\mathbb{R}^d \times \mathbb{R}^d)$ with compact support in $B_r(x_0) \times B_r(v_0)$ and such that $f_1 \geqslant 0$ and $f_1(x_0, v_0) = 1$. In this case, the integral on the right side above is nonnegative, since f_1 vanishes whenever $v \cdot \mathbf{n}_U \leqslant 0$. In fact, since f_1 is positive on a set of positive measure on $\partial U \times \mathbb{R}^d$ (in the sense of the product of the $(d-1)$-dimensional Hausdorff and Lebesgue measures), the integral above is positive. Up to multiplying f_1 by a positive scalar if necessary, we can thus ensure that (3.8) holds. It is clear that this construction also ensures that (3.7) and (3.9) hold. $\square$

Remark 3.2. As the argument above reveals, for the inequality (1.18) to hold, the assumption of $f \in H^1_{\mathrm{hyp},0}(U)$ can be weakened: it suffices that f vanishes on a relatively open piece of the boundary $\partial U \times \mathbb{R}^d$. The constant C in (1.18) then depends additionally on the identity of this piece of the boundary where f is assumed to vanish.

3.1.1. *Poincaré inequality with confining potential.* It is also interesting to understand Theorem 1.3 in the global setting with confining potential.[7]

Only in this subsection, we redefine $H^1_{\mathrm{hyp}}(\mathbb{R}^d)$ according to the norm

$$(3.10) \qquad \|f\|_{H^1_{\mathrm{hyp}}(\mathbb{R}^d)} = \|f\|_{L^2_\sigma(\mathbb{R}^d;H^1_\gamma)} + \|v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f\|_{L^2_\sigma(\mathbb{R}^d;H^{-1}_\gamma)},$$

and when $\mathbf{b}$ satisfies Assumption 1.1 with $U = \mathbb{R}^d$, and $f \in L^1_\sigma(\mathbb{R}^d; L^1_\gamma)$, we use the notation

$$(f)_{\mathbb{R}^d} := \int f \, dm.$$

Proposition 3.3 (Poincaré with confining potential). *Suppose that $\mathbf{b}$ satisfies Assumption 1.1 with $U = \mathbb{R}^d$, the potential W satisfies $W \in C^{1,1}(\mathbb{R}^d)$, and that there exists a constant $C_W < \infty$ such that the following weighted Poincaré inequality holds for all $h \in H^1_\sigma(\mathbb{R}^d)$ with $(h)_{\mathbb{R}^d} = 0$:*

$$(3.11) \qquad \int_U |\nabla_x W|^2 |h|^2 \, d\sigma \leqslant C_W \int_U |\nabla_x h|^2 \, d\sigma.$$

Then there exists a constant $C(W,d) < \infty$ such that for all $f \in H^1_{\mathrm{hyp}}(\mathbb{R}^d)$, defined according to (3.10), with $(f)_{\mathbb{R}^d} = 0$,

$$\|f\|_{L^2_\sigma(\mathbb{R}^d;L^2_\gamma)} \leqslant C \left(\|\nabla_v f\|_{L^2_\sigma(\mathbb{R}^d;L^2_\gamma)} + \|v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f\|_{L^2_\sigma(\mathbb{R}^d;H^{-1}_\gamma)} \right).$$

First, we require an analogue of Lemma 3.1.

Lemma 3.4 (Auxiliary lemma). *Under the assumptions of Proposition 3.3, there exists $C(W,d) < \infty$ such that for every $h \in L^2_\sigma$,*

$$\|h - (h)_{\mathbb{R}^d}\|_{L^2_\sigma} \leqslant C \|\nabla_x h\|_{H^{-1}_\sigma}.$$

Proof. Without loss of generality, we assume that $(h)_{\mathbb{R}^d} = 0$. Consider the operators

$$\widetilde{A} = \nabla_x, \quad \widetilde{A}^* = -\operatorname{div}_x - \mathbf{b} \cdot.$$

We consider the problem

$$(3.12) \qquad \widetilde{A}^* \mathbf{g} = h \quad \text{in } \mathbb{R}^d,$$

where we seek $\mathbf{g} \in H^1_\sigma$. The problem can be solved by defining $\mathbf{g} = \widetilde{A} f$ and solving

$$(3.13) \qquad \widetilde{A}^* \widetilde{A} f = h \quad \text{in } \mathbb{R}^d$$

with $(f)_{\mathbb{R}^d} = 0$. By the Lax-Milgram lemma, there exists a solution $f \in H^1_\sigma$ with $(f)_{\mathbb{R}^d} = 0$ and $\|f\|_{H^1_\sigma} \leqslant C \|h\|_{H^{-1}_\sigma} \leqslant C \|h\|_{L^2}$. To demonstrate that $\mathbf{g} \in H^1_\sigma$, we commute a derivative ∂_i through (3.13):

$$(3.14) \qquad \widetilde{A}^* \widetilde{A} \partial_i f = -\Delta_x \partial_i f - \mathbf{b} \cdot \nabla_x \partial_i f = \partial_i h + \partial_i \mathbf{b} \cdot \nabla_x f =: F,$$

---

[7]A proof is also contained in [17] following the methods in the original version of this paper, which only discussed bounded domains.

where F is a forcing term in H_σ^{-1}. Clearly, $\|\partial_i h\|_{H_\sigma^{-1}} \leqslant C\|h\|_{L_\sigma^2}$.[8] For the commutator term, we have

$$\|\partial_i \mathbf{b} \cdot \nabla_x f\|_{L_\sigma^2} \leqslant \|\partial_i \mathbf{b}\|_{L^\infty} \|\nabla_x f\|_{L_\sigma^2} \leqslant C\|h\|_{H_\sigma^{-1}},$$

where C depends on the $C^{1,1}$ regularity of W. By the Lax-Milgram lemma (or energy estimates) applied to (3.14) for each i, we have that

$$(3.15) \qquad \|\nabla_x \mathbf{g}\|_{L_\sigma^2} \leqslant C\|\nabla_x^2 f\|_{L_\sigma^2} \leqslant C\|F\|_{H_\sigma^{-1}} \leqslant C\|h\|_{L_\sigma^2}.$$

While $\mathbf{g}$ may not have zero average, it was already controlled in L_σ^2. Finally, we have

$$\|h\|_{L_\sigma^2}^2 = \int_{\mathbb{R}^d} h \, \nabla \cdot \mathbf{g} \, d\sigma = -\int_{\mathbb{R}^d} \nabla h \cdot \mathbf{g} \, d\sigma \leqslant \|\nabla h\|_{H_\sigma^{-1}} \|\mathbf{g}\|_{H_\sigma^1}.$$

The conclusion then follows by (3.15). $\hfill\square$

Proof of Proposition 3.3. Let $f \in H_{\mathrm{hyp}}^1(\mathbb{R}^d)$, see (3.10). By applying an approximation procedure with smooth cut-off in x and v and mollifying, we can without loss of generality assume that f is a compactly supported, smooth function. Again, we decompose the proof into three steps. Step 1 is identical, so we skip to

Step 2. We show that

$$(3.16) \qquad \|\nabla \langle f \rangle_\gamma\|_{H_\sigma^{-1}(\mathbb{R}^d)} \leqslant C\left(\|\nabla_v f\|_{L_\sigma^2(\mathbb{R}^d;L_\gamma^2)} + \|(v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v)f\|_{L_\sigma^2(\mathbb{R}^d;H_\gamma^{-1})}\right).$$

We select $\xi_1, \ldots, \xi_d \in C_c^\infty(\mathbb{R}^d)$ satisfying

$$\int_{\mathbb{R}^d} v\xi_i(v) \, d\gamma(v) = e_i,$$

and for each test function $\phi \in H_\sigma^1(\mathbb{R}^d)$ and $i \in \{1, \ldots, d\}$, we compute
(3.17)
$$\int \phi \partial_{x_i} \langle f \rangle_\gamma \, dm = \int \phi \xi_i(v) v \cdot \nabla_x \langle f \rangle_\gamma \, dm = -\int \phi \xi_i v \cdot \nabla_x (f - \langle f \rangle_\gamma) \, dm + \int \phi \xi_i v \cdot \nabla_x f \, dm.$$

We expand the second term on the right-hand side as

$$(3.18) \qquad \int \phi \xi_i v \cdot \nabla_x f \, dm = \int \phi \xi_i (v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v) f \, dm - \int \phi \xi_i \mathbf{b} \cdot \nabla_v (f - \langle f \rangle_\gamma) \, dm,$$

where we use that $\mathbf{b} \cdot \nabla_v \langle f \rangle_\gamma = 0$. Combining (3.17) and (3.18), we have

$$\int \phi \partial_{x_i} \langle f \rangle_\gamma \, dm = \int \phi \xi_i (v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v) f \, dm - \int \phi \xi_i (v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v)(f - \langle f \rangle_\gamma) \, dm = \mathrm{I} + \mathrm{II}.$$

For I, we have

$$\left|\int \phi \xi_i (v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v) f \, dm\right| \leqslant C\|\phi \xi_i\|_{L_\sigma^2(\mathbb{R}^d;H_\gamma^1)} \|(v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v)f\|_{L_\sigma^2(\mathbb{R}^d;H_\gamma^{-1})}.$$

For II, we integrate by parts across the measure dm:

$$-\int \phi \xi_i (v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v)(f - \langle f \rangle_\gamma) \, dm = \int v \cdot \nabla_x \phi \xi_i (f - \langle f \rangle_\gamma) \, dm + \int \phi \mathbf{b} \cdot \nabla_v \xi_i (f - \langle f \rangle_\gamma) \, dm = \mathrm{II}_a + \mathrm{II}_b.$$

For II_a, we use

$$\left|\int v \cdot \nabla_x \phi \xi_i (f - \langle f \rangle_\gamma) \, dm\right| \leqslant C\|v\xi_i\|_{L_\gamma^\infty} \|\phi\|_{L_\sigma^2} \|f - \langle f \rangle_\gamma\|_{L_\sigma^2(\mathbb{R}^d;L_\gamma^2)}.$$

For II_b, we use

$$\left|\int \phi \mathbf{b} \cdot \nabla_v \xi_i (f - \langle f \rangle_\gamma) \, dm\right| \leqslant \|\nabla W\| \|\phi\|_{L_\sigma^2} \|\nabla_v \xi_i\|_{L_\gamma^\infty} \|f - \langle f \rangle_\gamma\|_{L_\sigma^2(\mathbb{R}^d;L_\gamma^2)}.$$

---

[8]This follows from integration by parts against a test function $g \in H_\sigma^1$ and the Poincaré inequality in (3.11), which controls the term $\int \partial_i W g h \, d\sigma$ appearing when ∂_i hits the weight.

We use the assumed Poincaré inequality (3.11) to control $\|\nabla W\|\phi\|\|_{L_\sigma^2}$ by $\|\phi\|_{H_\sigma^1}$. Then using (3.3) concludes the proof of (3.16).

Step 3. We deduce from Lemma 3.4, (3.3) and (3.16) that

$$\|f - (f)_U\|_{L_\sigma^2(\mathbb{R}^d; L_\gamma^2)} \leqslant \|f - \langle f \rangle_\gamma\|_{L_\sigma^2(\mathbb{R}^d; L_\gamma^2)} + \|\langle f \rangle_\gamma - (f)_{\mathbb{R}^d}\|_{L_\sigma^2}$$

$$\leqslant \|f - \langle f \rangle_\gamma\|_{L_\sigma^2(\mathbb{R}^d; L_\gamma^2)} + C \|\nabla \langle f \rangle_\gamma\|_{H_\sigma^{-1}}$$

$$\leqslant C \left(\|\nabla_v f\|_{L_\sigma^2(\mathbb{R}^d; L_\gamma^2)} + \|(v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v) f\|_{L_\sigma^2(\mathbb{R}^d; H_\gamma^{-1})} \right).$$

This completes the proof. $\qquad \square$

3.2. Interpolation and Hörmander inequalities for H_{hyp}^1.

In this subsection, we use the Hörmander bracket condition to obtain a functional inequality which provides some interior spatial regularity for general H_{hyp}^1 functions. Both the statement and proof of the inequality follow closely the ideas of Hörmander [31]. Other variants of Hörmander's inequality have been previously obtained, see in particular [13] and [2]. We remind the reader that our initial estimates are phrased in terms of the Besov-type norms defined in subsection 2.3 and are thus valid for $U = \mathbb{T}^d, \mathbb{R}^d$.

Proposition 3.5 (Interpolation). *For every $\delta > 0$, there exists $C(d, \delta) < \infty$ such that for $U = \mathbb{T}^d, \mathbb{R}^d$ and any smooth function $u : U \times \mathbb{R}^d \to \mathbb{R}$, we have*

$$(3.19) \qquad \|u\|_{Q_{v \cdot \nabla_x}^{1/2}(U)} \leqslant C \left(\|u\|_{L^2(U; H_\gamma^1)} + \|v \cdot \nabla_x u\|_{L^2(U; H_\gamma^{-1})} \right) + \delta \|u\|_{Q_{\nabla_x}^{1/3}(U)}.$$

Proof. Step 1. Let $\phi \in C_0^\infty((-1,1)^d)$ be a smooth, positive, radial function with unit L^1 norm. For $t \in (0, \infty)$, we define $\phi_t u(x, v)$ by

$$\phi_t u(x, v) = \int_{\mathbb{R}^d} u(x + t^3 x', v) \phi(x') \, dx',$$

where in the case $U = \mathbb{T}^d$ we have periodicially extended u to a function defined on all of $\mathbb{R}^d$. Using Jensen's inequality, we calculate that

$$\|\phi_t u(x, v) - u(x, v)\|_{L^2(U; L_\gamma^2)}^2 = \iint_{\mathbb{R}^d \times U} \left(\int_{\mathbb{R}^d} \phi(x') \left(u(x + t^3 x', v) - u(x, v) \right) dx' \right)^2 dx \, d\gamma(v)$$

$$\leqslant \iiint_{\mathbb{R}^d \times U \times \mathbb{R}^d} \phi(x') \left(u(x + t^3 x', v) - u(x, v) \right)^2 dx' \, dx \, d\gamma(v)$$

$$= \iiint_{\mathbb{R}^d \times U \times \mathbb{R}^d} \phi(x') t^2 \frac{1}{t^2} \left(u(x + t^3 x', v) - u(x, v) \right)^2 dx' \, dx \, d\gamma(v)$$

$$\leqslant \int_{\mathbb{R}^d} \phi(x') t^2 \|u\|_{Q_{\nabla_x}^{1/3}}^2 \, dx',$$

and thus we see that

$$(3.20) \qquad \|\phi_t u(x, v) - u(x, v)\|_{L^2(U; L_\gamma^2)}^2 \leqslant t^2 \|u\|_{Q_{\nabla_x}^{1/3}(U)}^2.$$

Step 2. Let

$$f(t) = \|u(x + t^2 v, v) - u(x, v)\|_{L^2(U; L_\gamma^2)}^2.$$

For $t \in (0, \infty)$, it will suffice to show that

$$(3.21) \qquad f(t) \leqslant t^2 \left(C \left(\|u\|_{L^2(U; H_\gamma^1)} + \|v \cdot \nabla_x u\|_{L^2(U; H_\gamma^{-1})} \right) + \delta \|u\|_{Q_{\nabla_x}^{1/3}(U)} \right)^2.$$

Moreover, for $t \geqslant 1$, we have the obvious estimate $f(t) \leqslant 4\|u\|_{L^2(U;L^2_\gamma)}^2$, so we consider only $t \in (0,1)$. We may write that

$$
\begin{aligned}
f(t) \leqslant{}& \left\|\phi_{\delta t} u(x + t^2 v, v) - u(x + t^2 v, v)\right\|_{L^2(U;L^2_\gamma)}^2 \\
&+ \left\|\phi_{\delta t} u(x + t^2 v, v) - \phi_{\delta t} u(x, v)\right\|_{L^2(U;L^2_\gamma)}^2 + \left\|\phi_{\delta t} u(x, v) - u(x, v)\right\|_{L^2(U;L^2_\gamma)}^2 .
\end{aligned}
\tag{3.22}
$$

By Step 1, the first and third terms of (3.22) are bounded by

$$
\delta^2 t^2 \|u\|_{Q^{1/3}_{\nabla_x}}^2 .
$$

Step 3. It remains to estimate the second term in (3.22). For $t \in (0,1)$ and $0 \leqslant \tau \leqslant t^2$, consider

$$
F(\tau) = \|\phi_{\delta t} u(x + \tau v, v) - \phi_{\delta t} u(x, v)\|_{L^2(U;L^2_\gamma)}^2 ,
\tag{3.23}
$$

where $F(t^2)$ is precisely the second term in (3.22). Since $F(0) = 0$, it will suffice to show that there exists $C(d, \delta) < \infty$ such that

$$
F'(\tau) \leqslant C^2 \left(\|u\|_{L^2(U;H^1_\gamma)}^2 + \|v \cdot \nabla_x u\|_{L^2(U;H^{-1}_\gamma)}^2 \right) + \delta^2 \|u\|_{Q^{1/3}_{\nabla_x}}^2 .
$$

We have that

$$
\begin{aligned}
F'(\tau) &= \iint_{\mathbb{R}^d \times U} (\phi_{\delta t} u(x + \tau v, v) - \phi_{\delta t} u(x, v)) \, v \cdot \nabla_x (\phi_{\delta t} u) (x + \tau v, v) \, dx \, d\gamma(v) \\
&= \iint_{\mathbb{R}^d \times U} (\phi_{\delta t} u(x, v) - \phi_{\delta t} u(x - \tau v, v)) \, v \cdot \nabla_x (\phi_{\delta t} u) (x, v) \, dx \, d\gamma(v) .
\end{aligned}
$$

Since $[v \cdot \nabla_x, \phi_{\delta t}] u = [\nabla_v, \phi_{\delta t}] u = 0$ and we have a bound on $\|v \cdot \nabla_x u\|_{L^2(U;H^{-1}_\gamma)}$, we will achieve the desired estimate for $F'(\tau)$ if we can bound

$$
(\phi_{\delta t} u(x, v) - \phi_{\delta t} u(x - \tau v, v))
\tag{3.24}
$$

in $L^2(U; H^1_\gamma)$. The only non-trivial estimate comes when the ∇_v lands on the x coordinate of the second term in (3.24), which we may write out as

$$
\begin{aligned}
&\int_{\mathbb{R}^d} -\tau \nabla_x u \left(x + (\delta t)^3 x' - \tau v, v\right) \phi(x') \, dx' \\
&= -\int_{\mathbb{R}^d} \frac{\tau}{(\delta t)^3} \nabla_{x'} u(x + (\delta t)^3 x' - \tau v, v) \phi(x') \, dx' \\
&= \int_{\mathbb{R}^d} \frac{\tau}{(\delta t)^3} u(x + (\delta t)^3 x' - \tau v, v) \nabla_{x'} \phi(x') \, dx' \\
&= \int_{\mathbb{R}^d} \frac{\tau}{(\delta t)^3} \left(u(x + (\delta t)^3 x' - \tau v, v) - u(x - \tau v, v)\right) \nabla_{x'} \phi(x') \, dx' .
\end{aligned}
$$

But by Step 1, this is bounded in $L^2(U; L^2_\gamma)$ by a constant multiple of

$$
\frac{\tau}{\delta t^3} |t| \|u\|_{Q^{1/3}_{\nabla_x}} \leqslant \frac{1}{\delta^3} \|u\|_{Q^{1/3}_{\nabla_x}} ,
$$

where we have used the assumption that $\tau \leqslant t^2$. Note that in order to absorb the $1/\delta^3$ in the denominator, we may appeal to the Cauchy-Schwarz and Young inequalities in front of $\|v \cdot \nabla_x u\|_{L^2_x(U;H^{-1}_\gamma)}$, which leads to the estimate (3.19) after modifying δ to absorb any implicit constants. $\qquad \square$

With Proposition 3.5 in hand, we can now prove a Hörmander inequality which provides regularity in the x variable, measured in the $Q_{\nabla_x}^{1/3}$ space. The H^α estimate in Theorem 1.4 for $\alpha < 1/3$ will be an immediate corollary, and essentially amounts to converting $B_{2,\infty}^{1/3}$-type regularity to $B_{2,2}^\alpha$-type regularity. Following [31], the proof of Theorem 1.4 is based on the splitting of a first-order finite difference in the x variable into finite differences which are either in the v variable, or in the x variable in the direction of v. Explicitly, we have

$$
\begin{aligned}
(3.25) \qquad f(x + t^3 y, v) - f(x, v) = {} & f(x + t^3 y, v) - f(x + t^3 y, v - ty) \\
& + f(x + t^3 y, v - ty) - f(x + t^3 y + t^2(v - ty), v - ty) \\
& + f(x + t^2 v, v - ty) - f(x + tv, v) \\
& + f(x + t^2 v, v) - f(x, v).
\end{aligned}
$$

Notice that the right side consists of four finite differences, two for each of the derivatives ∇_v and $v \cdot \nabla_x$ which we can expect to control by the $L_2(U; H_\gamma^1)$ and $Q_{v \cdot \nabla_x}^{1/2}$ norms, respectively. The fact that the increment on the left is of size t^3 and those on the right side are of size t and t^2 suggests that we may expect to have one-third derivative in the statement of Theorem 1.4, which we are able to obtain in a Besov sense with the $Q_{\nabla_x}^{1/3}$ norm. The exponent $1/3$ is optimal, although it may be possible to improve the endpoint regularity from $B_{2,\infty}^{1/3}$-type to $B_{2,2}^{1/3}$ using more advanced microlocal techniques.

The relation (3.25) is a special case of Hörmander's bracket condition introduced in [31], which for the particular equation we consider here is quite simple to check. Indeed, let $X_1, \ldots, X_d, V_1, \ldots, V_d$ denote the canonical vector fields and X_0 be the vector field $(x, v) \mapsto (v, 0)$. Then the Hörmander bracket condition is implied by the identity

$$
(3.26) \qquad [V_i, X_0] = X_i.
$$

This is a local version of the identity (3.25). More precisely, for every vector field Z, if we denote by $t \mapsto \exp(tZ)$ the flow induced by the vector field Z on $\mathbb{R}^d \times \mathbb{R}^d$, then

$$
\begin{aligned}
(3.27) \quad \exp(-tV_i) \exp(-tX_0) \exp(tV_i) \exp(tX_0)(x, v) & \\
= (x, v) + t^2[V_i, X_0](x, v) + o(t^2) & \qquad (t \to 0).
\end{aligned}
$$

For the vector fields of interest, $Z \in \{X_0, X_1, \ldots, X_d, V_1, \ldots, V_d\}$, the flows take the very simple form

$$
\exp(tZ)(x, v) = (x, v) + tZ(x, v),
$$

the relation (3.27) becomes an identity (that is, the term $o(t^2)$ is actually zero), and loosely, this identity can be rephrased in the form of (3.25). The only difference is that, to exploit that our functions have only $1/2$ derivatives in the $v \cdot \nabla_x$ direction, it is advantageous to flow in the direction $v \cdot \nabla_x$ with speed t rather than unit speed.

Proposition 3.6 (Besov-Type Hörmander Inequality). *There exists a dimensional constant $C(d) < \infty$ such that for $U = \mathbb{T}^d, \mathbb{R}^d$ and any smooth function $u : U \times \mathbb{R}^d \to \mathbb{R}$, we have the estimate*

$$
(3.28) \qquad \|u\|_{Q_{\nabla_x}^{1/3}(U)} \leqslant C \left(\|u\|_{Q_{v \cdot \nabla_x}^{1/2}(U)} + \|u\|_{L_x^2(U; H_\gamma^1)} \right).
$$

Proof of Proposition 3.6. Let $f(x, v) = u(x, v)\gamma^{1/2}(v)$, and choose $\eta \in (0, \infty)$ and $x' \in \mathbb{S}^{d-1}$. Then we may write that

$$
\left\| u(x + \eta^3 x', v) - u(x, v) \right\|_{L^2(U; L_\gamma^2)} = \left\| f(x + \eta^3 x', v) - f(x, v) \right\|_{L^2(U; L^2)},
$$

and

$$
\begin{aligned}
f(x + \eta^3 x', v) - f(x, v) = {}& f(x + \eta^3 x', v) - f(x + \eta^3 x', v - \eta x') \\
& + f(x + \eta^3 x', v - \eta x') - f(x + \eta^3 x' + \eta^2(v - \eta x'), v - \eta x') \\
& + f(x + \eta^2 v, v - \eta x') - f(x + \eta^2 v, v) \\
& + f(x + \eta^2 v, v) - f(x, v) \,.
\end{aligned}
$$
(3.29)

Dividing by η, integrating in $L^2(U; L^2(\mathbb{R}^d))$, and appealing to (2.7) bounds the first term:

$$
\frac{1}{\eta^2} \iint_{\mathbb{R}^d \times U} \left(f(x + \eta^3 x', v) - f(x + \eta^3 x', v - \eta x') \right)^2 dx\, dv \leqslant C \left\| \nabla_v u \right\|_{L^2(U; L^2_\gamma)}^2 \,,
$$

with a similar bound holding for the third term. Dividing again by η and integrating in $L^2(U; L^2(\mathbb{R}^d))$ yields the bound

$$
\frac{1}{\eta^2} \iint_{\mathbb{R}^d \times U} \left(f(x + \eta^3 x', v - \eta x') - f(x + \eta^3 x' + \eta^2(v - \eta x'), v - \eta x') \right)^2 dx\, dv \leqslant \left\| u \right\|_{Q^{1/2}_{v \cdot \nabla_x}(U)}^2 \,,
$$

with a similar bound holding for the fourth term. Appealing to (3.19) with a suitably small choice of δ concludes the proof. $\qquad\square$

To obtain the statements in Theorem 1.4 for $\alpha < \sfrac{1}{3}$, we must work in H_x^α rather than $(B_{2,\infty}^\alpha)_x$ spaces of fractional differentiability, and so we introduce the Banach space-valued fractional Sobolev spaces, defined as follows: for every domain $U \subseteq \mathbb{R}^d$, $\alpha \in (0,1)$, Banach space X with norm $\|\cdot\|_X$ and $u \in L^2(U; X)$, we define the seminorm

$$
[\![u]\!]_{H^\alpha(U; X)} := \left(\int_U \int_U \frac{\| u(x) - u(y) \|_X^2}{|x - y|^{d + 2\alpha}} \, dx\, dy \right)^{1/2}
$$
(3.30)

and the norm

$$
\| u \|_{H^\alpha(U; X)} := \left(\| u \|_{L^2(U; X)}^2 + [\![u]\!]_{H^\alpha(U; X)}^2 \right)^{1/2} .
$$

We then define the fractional Sobolev space

$$
H^\alpha(U; X) := \left\{ u \in L^2(U; X) \,:\, \| u \|_{H^\alpha(U; X)} < \infty \right\}.
$$
(3.31)

The space $H^\alpha(U; X)$ is a Banach space under the norm $\|\cdot\|_{H^\alpha(U; X)}$. We understand that $H^0(U; X) = L^2(U; X)$. We also set

$$
\| u \|_{H^{1+\alpha}(U; X)} := \left(\| u \|_{L^2(U; X)}^2 + \| \nabla u \|_{H^\alpha(U; X)}^2 \right)^{1/2} ,
$$

and define the Banach space $H^{1+\alpha}(U; X)$ as in (3.31). We may now use Proposition 3.6 to prove the non-endpoint estimates from Theorem 1.4.

Proof of Theorem 1.4. We have that for $\alpha < 1/3$,

$$
\begin{aligned}
[\![u]\!]_{H^\alpha(U;L^2_\gamma)}^2 &= \iint_{U\times U} \frac{\|u(x,\cdot) - u(y,\cdot)\|_{L^2_\gamma}^2}{|x-y|^{d+2\alpha}}\, dx\, dy \\
&= \iint_{U\times U} \frac{\|u(x'+y,\cdot) - u(y,\cdot)\|_{L^2_\gamma}^2}{|x'|^{d+2\alpha}}\, dx'\, dy \\
&= \iint_{\{|x'|<1\}\times U} \frac{\|u(x'+y,\cdot) - u(y,\cdot)\|_{L^2_\gamma}^2}{|x'|^{d+2\alpha}}\, dy\, dx' \\
&\quad + \iint_{\{|x'|\geqslant 1\}\times U} \frac{\|u(x'+y,\cdot) - u(y,\cdot)\|_{L^2_\gamma}^2}{|x'|^{d+2\alpha}}\, dy\, dx' \\
&\leqslant \int_{\{|x'|<1\}} \frac{|x'|^{2/3}\|u\|_{Q^{1/3}_{\nabla_x}}^2}{|x'|^{d+2\alpha}}\, dx' \quad + \quad C(\alpha)\|u\|_{L^2(U;L^2_\gamma)}^2 \\
&\leqslant C(\alpha)\left(\|u\|_{L^2(U;H^1_\gamma)}^2 + \|v\cdot\nabla_x u\|_{L^2(U;H^{-1}_\gamma)}^2\right) \leqslant C(\alpha)\|u\|_{H^1_{\mathrm{hyp}}(U)}^2,
\end{aligned}
$$

concluding the proof. Notice that we have restricted the domain to $U = \mathbb{R}^d, \mathbb{T}^d$ in the above estimates. $\square$

For the purposes of interpolation, we also need to consider fractional Sobolev spaces in the velocity variable. As discussed in the arguments leading to (2.7), the relevant spaces are weighted by the measure γ, which is strongly inhomogeneous. Because of this difficulty, we use the following definition. For each $f \in L^2_\gamma$ and $t > 0$, we set

$$
K(t,f) := \inf\left\{\|f_0\|_{L^2_\gamma} + t\|f_1\|_{H^1_\gamma} \ : \ f = f_0 + f_1, \ f_0 \in L^2_\gamma, \ f_1 \in H^1_\gamma\right\},
$$

and, for every $\alpha \in (0,1)$, we define

$$
(3.32) \qquad\qquad \|f\|_{H^\alpha_\gamma} := \left(\int_0^\infty \left(t^{-\alpha} K(f,t)\right)^2 \frac{dt}{t}\right)^{\frac{1}{2}}.
$$

We also define $H^{-\alpha}_\gamma$ to be the space dual to H^α_γ.

We may utilize interpolation to obtain embeddings into other similar spaces of positive regularity in both variables. In particular, appealing to Theorem 1.4 and the interpolation inequality

$$
\|f\|_{H^{\theta\beta}(U;H^{1-\theta}_\gamma)} \lesssim \|f\|_{H^\beta(U;L^2_\gamma)}^\theta \|f\|_{L^2(U;H^1_\gamma)}^{1-\theta} \qquad \theta \in [0,1] \qquad U = \mathbb{T}^d, \mathbb{R}^d,
$$

immediately implies the following estimate.

Corollary 3.7 (Hörmander inequality for H^1_{hyp}). *Let $\alpha \in \left[0, \frac{1}{3}\right)$ and $U = \mathbb{T}^d, \mathbb{R}^d$. There exists a constant $C(\alpha, d) < \infty$ such that for every $\theta \in [0,1]$ and every $f \in H^1_{\mathrm{hyp}}(U)$, we have the estimate*

$$
\|f\|_{H^{\theta\alpha}(U;H^{1-\theta}_\gamma)} \leqslant C\|f\|_{H^1_{\mathrm{hyp}}(U)}.
$$

Observe that, by introducing a cutoff function in the spatial variable, we also obtain analogous embeddings for bounded domains $U \subseteq \mathbb{R}^d$, such as

$$
H^1_{\mathrm{hyp}}(U) \hookrightarrow H^\alpha(U_\delta; L^2_\gamma),
$$

valid for every $\alpha < \frac{1}{3}$ and $\delta > 0$, where $U_\delta := \{x \in U : \mathrm{dist}(x, \partial U) > \delta\}$.

3.3. **Compact embedding of H^1_{hyp} into $L^2(U; L^2_\gamma)$.** Using the results of the previous subsection, we show that the embedding $H^1_{\mathrm{hyp}}(U) \hookrightarrow L^2(U; L^2_\gamma)$ is compact. In this section, we assume that $U \subseteq \mathbb{R}^d$ is a bounded C^1 domain or $\mathbb{T}^d$.

Proposition 3.8 (Compact embedding of $H^1_{\mathrm{hyp}}(U)$ into $L^2(U; L^2_\gamma)$)**.** *The inclusion map* $H^1_{\mathrm{hyp}}(U) \hookrightarrow L^2(U; L^2_\gamma)$ *is compact.*

The proof is straightforward on $\mathbb{T}^d$. First, approximate by functions in $C^\infty_0(\mathbb{T}^d \times \mathbb{R}^d)$. Next, we use the embedding $H^1_{\mathrm{hyp}}(\mathbb{T}^d) \subseteq H^\alpha(\mathbb{T}^d \times B_{v_0})$ for all $v_0 \in [1, +\infty)$. Finally, we apply the standard Rellich compactness theorem. Hence, we focus only on bounded C^1 domains $U \subseteq \mathbb{R}^d$ below.

Before we give the proof of Proposition 3.8, we need to review some basic facts concerning the logarithmic Sobolev inequality and a generalized Hölder inequality for Orlicz norms. The logarithmic Sobolev inequality states that, for some $C < \infty$,

$$(3.33) \quad \int_{\mathbb{R}^d} f^2(v) \log\left(1 + f^2(v)\right) d\gamma(v) \leqslant C \int_{\mathbb{R}^d} |\nabla f|^2 \, d\gamma(v), \quad \forall f \in H^1_\gamma \text{ with } \|f\|_{L^2_\gamma} = 1.$$

Let $F : \mathbb{R} \to [0, \infty)$ denote the (strictly) convex function

$$F(t) := |t| \log\left(1 + |t|\right).$$

Let F^* denote its dual convex conjugate function, defined by

$$F^*(s) := \sup_{t \in \mathbb{R}} \left(st - F(t)\right).$$

Then (F, F^*) is a *Young pair* (see [41]), that is, both F and F^* are nonnegative, even, convex, satisfy $F(0) = F^*(0) = 0$ as well as

$$\lim_{|t| \to \infty} |t|^{-1} F(t) = \lim_{|s| \to \infty} |s|^{-1} F^*(s) = \infty.$$

Moreover, both F and F^* are strictly increasing on $[0, \infty)$ and in particular vanish only at $t = 0$. Given any measure space (X, ω), the *Orcliz spaces* $L_F(X, \omega)$ and $L_{F^*}(X, \omega)$, which are defined by the norms

$$\begin{cases} \|g\|_{L_F(X,\omega)} := \inf\left\{t > 0 : \int_X F\left(t^{-1} g\right) d\omega \leqslant F(1)\right\}, \quad \text{and} \\ \|g\|_{L_{F^*}(X,\omega)} := \inf\left\{t > 0 : \int_X F^*\left(t^{-1} g\right) d\omega \leqslant F^*(1)\right\} \end{cases}$$

are dual Banach spaces and the following generalized version of the Hölder inequality is valid (see [41, Proposition 3.3.1]):

$$\int_X |gg^*| \, d\omega \leqslant \|g\|_{L_F(X,\omega)} \|g^*\|_{L_{F^*}(X,\omega)}, \quad \forall g \in L_F(X, \omega), \, g^* \in L_{F^*}(X, \omega).$$

The logarithmic Sobolev inequality (3.33) may be written in terms of the Orcliz norm as

$$\left\|f^2\right\|_{L_F(\mathbb{R}^d, \gamma)} \leqslant C\left(|\langle f \rangle_\gamma|^2 + \|\nabla f\|^2_{L^2_\gamma}\right), \quad \forall f \in H^1_\gamma.$$

The previous two displays imply that

$$(3.34) \quad \left(\int_{U \times \mathbb{R}^d} g |f|^2 \, dx \, d\gamma(v)\right)^{\frac{1}{2}} \leqslant C \|g\|^{\frac{1}{2}}_{L_{F^*}(U \times \mathbb{R}^d, dx d\gamma)} \|f\|_{L^2(U; H^1_\gamma)}.$$

We do not identify F^* with an explicit formula, although we notice that the inequality

$$s(t + 1) \leqslant \exp(s) + t \log(1 + t), \quad \forall s, t \in (0, \infty)$$

implies that

$$F^*(s) \leqslant \exp(s) - s.$$

This allows us in particular to obtain from (3.34) that

$$(3.35) \qquad \left(\int_{U \times \mathbb{R}^d} |v|^2 |f|^2 \, dx \, d\gamma(v) \right)^{\frac{1}{2}} \leqslant C \, \|f\|_{L^2(U; H^1_\gamma)} \, .$$

We also point out that (3.35) also implies the existence of $C(d, U) < \infty$ such that, for every $f \in L^2(U; L^2_\gamma)$,

$$(3.36) \qquad \|\nabla_v f\|_{L^2(U; H^{-1}_\gamma)} \leqslant C \, \|f\|_{L^2(U; L^2_\gamma)} \, .$$

We now turn to the proof of Proposition 3.8.

Proof of Proposition 3.8. For each $\theta > 0$, we denote

$$(3.37) \qquad U_\theta := \{ x : \mathrm{dist}(x, \partial U) < \theta \} \, .$$

Since U is a C^1 domain, we can extend the outer normal $\mathbf{n}_U$ to a globally C^0 function on $\overline{U}$. We can moreover assume that, for some $\theta_0(U) > 0$, this extension $\mathbf{n}_U$ coincides with the gradient of the mapping $x \mapsto -\mathrm{dist}(x, \partial U)$ in U_{θ_0}.

By Proposition 2.2, we may work under the qualitative assumption that all of our $H^1_{\mathrm{hyp}}(U)$ functions belong to $C^\infty_c(\overline{U} \times \mathbb{R}^d)$. Select $\varepsilon > 0$ and a sequence $\{f_n\}_{n \in \mathbb{N}} \subseteq H^1_{\mathrm{hyp}}(U)$ satisfying

$$\sup_{n \in \mathbb{N}} \|f_n\|_{H^1_{\mathrm{hyp}}(U)} \leqslant 1.$$

We will argue that there exists a subsequence $\{f_{n_k}\}$ such that

$$(3.38) \qquad \limsup_{k \to \infty} \sup_{i, j \geqslant k} \|f_{n_i} - f_{n_j}\|_{L^2(U; L^2_\gamma)} \leqslant \varepsilon \, .$$

The proposition may then be obtained by a diagonalization argument.

Step 1. We claim that there exists $v_0 \in [1, \infty)$ such that, for every $f \in H^1_{\mathrm{hyp}}(U)$,

$$\left(\int_U \int_{\mathbb{R}^d \setminus B_{v_0}} |f(x,v)|^2 \, dx \, d\gamma(v) \right)^{\frac{1}{2}} \leqslant \frac{\varepsilon}{3} \, \|f\|_{H^1_{\mathrm{hyp}}(U)} \, .$$

Indeed, applying (3.34), we find that

$$\left(\int_U \int_{\mathbb{R}^d \setminus B_{v_0}} |f(x,v)|^2 \, dx \, d\gamma(v) \right)^{\frac{1}{2}} \leqslant C \, \left\| \mathbb{1}_{U \times (\mathbb{R}^d \setminus v_0)} \right\|^{\frac{1}{2}}_{L_{F^*}(U \times \mathbb{R}^d, dx d\gamma)} \|f\|_{H^1_{\mathrm{hyp}}(U)} \, .$$

Taking v_0 sufficiently large, depending on ε, ensures that

$$C \, \left\| \mathbb{1}_{U \times (\mathbb{R}^d \setminus v_0)} \right\|^{\frac{1}{2}}_{L_{F^*}(U \times \mathbb{R}^d, dx d\gamma)} \leqslant \frac{\varepsilon}{3}.$$

Step 2. We next claim that there exists $\delta \in \left(0, \frac{1}{2}\right]$ such that, for every $f \in H^1_{\mathrm{hyp}}(U)$,

$$\left(\int_U \int_{\mathbb{R}^d} |f(x,v)|^2 \, \mathbb{1}_{\{|\mathbf{n}_U \cdot v| < \delta\}} \, dx \, d\gamma(v) \right)^{\frac{1}{2}} \leqslant \frac{\varepsilon}{3} \, \|f\|_{H^1_{\mathrm{hyp}}(U)} \, .$$

The argument here is similar to the estimate in Step 1, above. We simply apply (3.34) after choosing δ small enough that

$$C \, \left\| \mathbb{1}_{\{|\mathbf{n}_U \cdot v| < \delta\}} \right\|^{\frac{1}{2}}_{L_{F^*}(U \times \mathbb{R}^d, dx d\gamma)} \leqslant \frac{\varepsilon}{3}.$$

Step 3. We next show that, for every $\delta > 0$, there exists $\theta > 0$ such that, for every function $f \in H^1_{\mathrm{hyp}}(U)$,

$$(3.39) \qquad \left(\int_U \int_{\mathbb{R}^d} |f(x,v)|^2 \, \mathbb{1}_{\{|\mathbf{n}_U \cdot v| \geqslant \delta\}} \mathbb{1}_{\{\mathrm{dist}(x, \partial U) < \theta\}} \, dx \, d\gamma(v) \right)^{\frac{1}{2}} \leqslant \frac{\varepsilon}{3} \, \|f\|_{H^1_{\mathrm{hyp}}(U)} \, .$$

For $\theta \in \left(0, \frac{\theta_0}{2}\right]$ to be taken sufficiently small in terms of $\delta > 0$ in the course of the argument, we let $\varphi \in C^{1,1}(\overline{U})$ be defined by

$$\varphi(x) := -\eta\left(\operatorname{dist}(x, \partial U)\right),$$

where $\eta \in C_c^\infty([0, \infty))$ satisfies

$$0 \leqslant \eta \leqslant 2\theta, \quad 0 \leqslant \eta' \leqslant 1, \quad \eta(x) = x \text{ on } [0, \theta], \quad \eta' = 0, \text{ on } [2\theta, \infty).$$

We have $-2\theta \leqslant \varphi \leqslant 0$. Moreover, by the definition of θ_0 below (3.37), its gradient $\nabla\varphi$ is proportional to $\mathbf{n}_U$ in U, it vanishes outside of $U_{2\theta}$, and $\nabla\varphi = \mathbf{n}_U$ in U_θ. We next select another test function $\chi \in C_c^\infty([0, \infty))$ satisfying

$$0 \leqslant \chi \leqslant 1, \quad \chi \equiv 0 \text{ on } \left[0, \tfrac{1}{2}\delta\right], \quad \chi \equiv 1 \text{ on } [\delta, \infty), \quad |\chi'| \leqslant \delta^{-1},$$

and define

$$\psi_\pm(x, v) := \chi\left((v \cdot \mathbf{n}_U(x))_\pm\right),$$

where for $r \in \mathbb{R}$, we use the notation $r_- := \max(0, -r)$ and $r_+ := \max(0, r)$. Observe that

$$|\nabla_v \psi_\pm(x, v)| = \left|\chi'\left((v \cdot \mathbf{n}_U(x))_\pm\right)\right| |\mathbf{n}_U(x)| \leqslant C\delta^{-1}.$$

Therefore

$$\|\varphi f \psi_\pm\|_{L^2(U; H_\gamma^1)} \leqslant C\left(\|\varphi f \psi_\pm\|_{L^2(U; L_\gamma^2)} + \|\varphi \nabla_v\left(f\psi_\pm\right)\|_{L^2(U; L_\gamma^2)}\right)$$

$$\leqslant C\theta\left(\|f\|_{L^2(U; L_\gamma^2)} + \|\nabla_v f\|_{L^2(U; L_\gamma^2)} + \|f \nabla_v \psi_\pm\|_{L^2(U; L_\gamma^2)}\right)$$

$$\leqslant C\theta\delta^{-1}\|f\|_{L^2(U; H_\gamma^1)},$$

and hence

$$\left|\int_{U \times \mathbb{R}^d} \varphi f \psi_\pm v \cdot \nabla_x f \, dx \, d\gamma(v)\right| \leqslant C\theta\delta^{-1}\|f\|_{H_{\mathrm{hyp}}^1(U)}^2.$$

On the other hand,

$$\int_{U \times \mathbb{R}^d} \varphi f \psi_\pm v \cdot \nabla_x f \, dx \, d\gamma(v)$$

$$= -\frac{1}{2}\int_{U \times \mathbb{R}^d} f^2 v \cdot \nabla_x\left(\varphi \psi_\pm\right) \, dx \, d\gamma(v)$$

$$= -\frac{1}{2}\int_{U \times \mathbb{R}^d} \varphi f^2 v \cdot \nabla_x \psi_\pm \, dx \, d\gamma(v) - \frac{1}{2}\int_{U \times \mathbb{R}^d} \psi_\pm f^2 v \cdot \nabla\varphi \, dx \, d\gamma(v).$$

Since $|v \cdot \nabla_x \psi_\pm(x, v)| \leqslant C\delta^{-1}|v|^2$, we have, by (3.35),

$$\left|\int_{U \times \mathbb{R}^d} \varphi f^2 v \cdot \nabla_x \psi_\pm \, dx \, d\gamma(v)\right| \leqslant C\theta\delta^{-1}\int_{U \times \mathbb{R}^d} |v|^2 f^2 \, dx \, d\gamma(v) \leqslant C\theta\delta^{-1}\|f\|_{H_{\mathrm{hyp}}^1(U)}^2.$$

We deduce that

$$\left|\int_{U \times \mathbb{R}^d} \psi_\pm f^2 v \cdot \nabla\varphi \, dx \, d\gamma(v)\right| \leqslant C\theta\delta^{-1}\|f\|_{H_{\mathrm{hyp}}^1(U)}^2.$$

Finally, we observe from the properties of φ and $\psi_\pm$ that

$$\int_U \int_{\mathbb{R}^d} |f(x, v)|^2 \, \mathbb{1}_{\{|\mathbf{n}_U \cdot v| \geqslant \delta\}} \mathbb{1}_{\{\operatorname{dist}(x, \partial U) < \theta\}} \, dx \, d\gamma(v)$$

$$\leqslant \delta^{-1}\left(\left|\int_{U \times \mathbb{R}^d} \psi_+ f^2 v \cdot \nabla\varphi \, dx \, d\gamma(v)\right| + \left|\int_{U \times \mathbb{R}^d} \psi_- f^2 v \cdot \nabla\varphi \, dx \, d\gamma(v)\right|\right)$$

$$\leqslant C\theta\delta^{-2}\|f\|_{H_{\mathrm{hyp}}^1(U)}^2.$$

Taking $\theta = c\varepsilon^2\delta^2$ for a sufficiently small constant $c > 0$ yields the claimed inequality (3.39).

Step 4. By the results of the previous three steps, to obtain (3.38) it suffices to exhibit a subsequence $\{f_{n_k}\}$ satisfying

$$\limsup_{k\to\infty} \sup_{i,j\geqslant k} \int_{U_\theta\times B_{v_0}} \left|f_{n_i} - f_{n_j}\right|^2 dx\, d\gamma(v) = 0.$$

This is an immediate consequence of Corollary 3.7 and the compactness of the embedding $H^{1/10}\left(U_\theta; H_\gamma^{1/3}\right) \hookrightarrow L^2(U_\theta; L_\gamma^2(B_{v_0}))$ (see for instance [1, Theorem 2.32]). $\square$

4. The Kramers equation

In this section, we present two proofs of the existence of weak solutions in $H^1_{\mathrm{hyp}}(\mathbb{T}^d)$ to the Kramers equation

$$(4.1) \qquad\qquad -\Delta_v f + v\cdot\nabla_v f + v\cdot\nabla_x f + \mathbf{b}\cdot\nabla_v f = g^*,$$

where $g^* \in L^2(\mathbb{T}^d; H_\gamma^{-1})$ satisfies $\iint_{\mathbb{T}^d\times\mathbb{R}^d} g^*\, dm = 0$ (recall that the weighted mean of g^* is well defined by duality since the function 1 belongs to $L^2(\mathbb{T}^d; H_\gamma^1)$). The first proof uses the abstract Lions-Lax-Milgram theorem and a modification of (4.1) with a penalization term νf. The hypoelliptic energy estimates are used in sending the parameter ν to zero. This approach is partly inspired by [18]. The second proof uses a dual variational approach which characterizes the weak solutions of (4.1) as the minimizers of a natural energy under an appropriate constraint, in analogy with the discussion following the statement of Theorem 1.2. In both cases, the Poincaré inequality from Theorem 1.3 provides the necessary coercivity.

Throughout this section, the force field $\mathbf{b}(x) = -\nabla W(x)$ is as in Assumption 1.1. In particular, $\mathbf{b}$ depends only on x and is conservative. Let dm be as defined in (1.7).

4.1. The Lions-Lax-Milgram approach.
We recall the abstract version of Lions' representation theorem from [42, Theorem 3.1, p. 109].

Lemma 4.1 (Lions' representation theorem)**.** *Let H be a Hilbert space and Φ a pre-Hilbert space. Let $E : H\times\Phi\to\mathbb{R}$ be a bilinear form satisfying the continuity criterion*

$$(4.2) \qquad\qquad E(\cdot,\phi) \in H^* \ \text{ for all } \phi\in\Phi.$$

Then the following two properties are equivalent:

- *(Coercivity) We have*

$$(4.3) \qquad\qquad \inf_{\|\phi\|_\Phi=1} \sup_{\|h\|_H\leqslant 1} |E(h,\phi)| \geqslant c > 0.$$

- *(Solvability) For each $L\in\Phi^*$, there exists $f\in H$ such that*

$$(4.4) \qquad\qquad E(f,\phi) = L(\phi), \quad \text{ for all } \phi\in\Phi.$$

Notice that uniqueness and stability estimates are not guaranteed by Lemma 4.1 itself; they are concluded a posteriori.

Proof of Theorem 1.2. We split the argument into steps; in the first step, we solve a penalized problem, and in the second, we send the penalization parameter ν to zero.

Step 1. Consider the penalized problem

$$(4.5) \qquad\qquad (v\cdot\nabla_x + \mathbf{b}\cdot\nabla_v)f + \nu f = g^* + \Delta f - v\cdot\nabla_v f$$

posed on the torus $\mathbb{T}^d$ where $\nu\in(0,1]$. We define the following objects:

(1) the *test function space*
$$\Phi = C_0^\infty(\mathbb{T}^d \times \mathbb{R}^d)$$

with inner product

(4.6) $$\qquad (\phi, \psi) = \iint_{\mathbb{T}^d \times \mathbb{R}^d} \nabla_v \phi \cdot \nabla_v \psi \, dm + \iint_{\mathbb{T}^d \times \mathbb{R}^d} \phi \psi \, dm \, ,$$

(2) the *solution space*
$$H = \{ h \in L_\sigma^2(\mathbb{T}^d; H_\gamma^1) : (h)_{\mathbb{T}^d} = 0 \}$$

with inner product (4.6),

(3) the penalized *bilinear form*
$$E(h, \phi) = \iint_{\mathbb{T}^d \times \mathbb{R}^d} \nabla_v h \cdot \nabla_v \phi \, dm + \nu \iint_{\mathbb{T}^d \times \mathbb{R}^d} h \phi \, dm - \iint_{\mathbb{T}^d \times \mathbb{R}^d} h(v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v) \phi \, dm \, ,$$

(4) and the *linear functional*
$$L = g^* \in L_\sigma^2(\mathbb{T}^d; H_\gamma^{-1}) \text{ with } (g^*)_{\mathbb{T}^d} = 0 \, .$$

It is not difficult to verify that E is continuous (4.2) and coercive (4.3). Indeed, the key features are that (i) the anti-symmetric operator $v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v$ hits the test function ϕ, and (ii) the penalization term $\nu \iint_{\mathbb{T}^d \times \mathbb{R}^d} |\phi|^2 \, dm$ controls the 'lower part' $(L^2(\mathbb{T}^d; L_\gamma^2))$ of the norm after testing with ϕ. Hence, Lemma 4.1 guarantees the existence of a solution $f \in H$ to (4.4), which is the distributional formulation of the penalized equation (4.5).

From the equation itself, we recover that $(v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v) f \in L_\sigma^2(\mathbb{T}^d; H_\gamma^{-1})$, and therefore, $f \in H_{\mathrm{hyp}}^1(\mathbb{T}^d)$ qualitatively. By the density of smooth functions in $H_{\mathrm{hyp}}^1(\mathbb{T}^d)$, this is enough regularity[9] to multiply (4.5) by f and integrate by parts to demonstrate the basic energy estimate:

(4.7) $$\qquad \iint_{\mathbb{T}^d \times \mathbb{R}^d} |\nabla_v f|^2 \, dm + \nu \iint_{\mathbb{T}^d \times \mathbb{R}^d} |f|^2 \, dm \leqslant C \nu^{-1} \|g^*\|_{L_\sigma^2(\mathbb{T}^d; H_\gamma^{-1})}^2 \, ,$$

which guarantees that the solution is unique.[10] From the equation itself, we have

$$\|(v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v) f\|_{L_\sigma^2(\mathbb{T}^d; H_\gamma^{-1})} \leqslant C \|A^* A f\|_{L_\sigma^2(\mathbb{T}^d; H_\gamma^{-1})} + \|g^*\|_{L^2(\mathbb{T}^d; H_\gamma^{-1})} + C\nu \|f\|_{L_\sigma^2(\mathbb{T}^d; H_\gamma^{-1})}$$

(4.8) $$\qquad \overset{(4.7)}{\leqslant} C \|g^*\|_{L_\sigma^2(\mathbb{T}^d; H_\gamma^{-1})} \, ,$$

where the constant C changes from line to line. Then (4.7), (4.8), and the hypoelliptic Poincaré inequality for mean-zero functions imply that

$$\|f\|_{H_{\mathrm{hyp}}^1(\mathbb{T}^d)} \leqslant C \|g^*\|_{L_\sigma^2(\mathbb{T}^d; H_\gamma^{-1})} \, .$$

Step 2. Next, we consider $\nu \to 0^+$. Let f^ν denote the unique solution of the penalized problem (4.5). Subtracting two solutions f^{ν_1} and f^{ν_2}, we have that the difference $\widetilde{f}^{\nu_1, \nu_2}$ solves the equation

(4.9) $$\qquad (v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v) \widetilde{f}^{\nu_1, \nu_2} + (\nu_1 f^{\nu_1} - \nu_2 f^{\nu_2}) = (\Delta - v \cdot \nabla_v) \widetilde{f}^{\nu_1, \nu_2} \, .$$

We may regard $\nu_1 f^{\nu_1} - \nu_2 f^{\nu_2}$ as a forcing term which is $O(\nu_1 + \nu_2)$ in $L_\sigma^2(\mathbb{T}^d; H_\gamma^{-1})$. By the hypoelliptic energy estimates for (4.9), we have

$$\|\widetilde{f}^{\nu_1, \nu_2}\|_{H_{\mathrm{hyp}}^1(\mathbb{T}^d)} = O(\nu_1 + \nu_2) \, .$$

---

[9]To justify this, one may use the density of test functions demonstrated in Proposition 2.2.
[10]The estimate (4.7) can be made more convenient, without the factor ν^{-1}, if $\langle g^* \rangle_\gamma \equiv 0$.

Choosing $\nu = 2^{-k}$, the sequence (f_k) of solutions to (4.5) with penalization $\nu = 2^{-k}$ is Cauchy in $H^1_{\mathrm{hyp}}(\mathbb{T}^d)$ and therefore converges to a solution f in $H^1_{\mathrm{hyp}}(\mathbb{T}^d)$ with $(f)_{\mathbb{T}^d} = 0$. By passing to the distributional limit in each term in (4.5), we find that f solves (4.1) in the sense of distributions. The proof is complete. $\qquad\square$

Remark 4.2 (Role of the penalization). The above proof requires a *coercive* bilinear form E which, in particular, controls the L^2 norm. The a priori estimates for solutions of (4.1) do indeed control the L^2 part of the norm through the hypoelliptic Poincaré inequality, but the control of $\|(v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v)f\|_{L^2(\mathbb{T}^d;H^{-1}_\gamma)}$ is encoded by the PDE itself rather than the bilinear form E, which only encodes the energy estimate. This is why we include the penalization νf. In some sense, control of $\|(v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v)f\|_{L^2(\mathbb{T}^d;H^{-1}_\gamma)}$ is concluded a posteriori.

In the time-dependent case, one can skip the penalization by instead considering the equation satisfied by $e^t f$; see Proposition 6.10.

Remark 4.3 (Difficulty with boundary). Consider (4.1) in a bounded C^1 domain U with force f^* and zero Dirichlet condition on $\partial_{\mathrm{hyp}}U$. What goes wrong with the proof? One can demonstrate that there exists a solution $f^\nu \in H^1_{\mathrm{hyp}}(U)$ of the penalized equations which satisfies $f^\nu|_{\partial_{\mathrm{hyp}}U} = 0$ *away from the singular set*. However, we do not know how to justify that $f^\nu \in H^1_{\mathrm{hyp},0}(U)$. That is, we cannot characterize $H^1_{\mathrm{hyp},0}(U)$ as consisting of $H^1_{\mathrm{hyp}}(U)$ functions which vanish on $\partial_{\mathrm{hyp}}U$ away from the singular set. Consequently, we cannot justify the integration by parts that would generate the energy estimates that would imply uniqueness of f^ν and allow us to send $\nu \to 0^+$.

4.2. The dual variational approach. Define

$$(4.10) \qquad\qquad Bf := v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f.$$

Consider the functional

$$(4.11) \qquad\qquad \mathcal{J}[f,\mathbf{j}] = \iint_{\mathbb{T}^d \times \mathbb{R}^d} \frac{1}{2} |\nabla_v f - \mathbf{j}|^2 \, d\sigma(x) \, d\gamma(v)$$

evaluated at pairs $(f,\mathbf{j}) \in H^1_{\mathrm{hyp}}(\mathbb{T}^d) \times \left(L^2(\mathbb{T}^d; L^2_\gamma) \right)^d$ satisfying

$$(4.12) \qquad \nabla_v^* \mathbf{j} = g^* - Bf = g^* - (v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f), \qquad (f)_{\mathbb{T}^d} = 0.$$

In the remainder of this section, we *always* consider $f \in H^1_{\mathrm{hyp}}(\mathbb{T}^d)$ satisfying the second condition. We seek a null minimizer of $\mathcal{J}$ restricted to such pairs, which, if it exists, will satisfy the implication

$$\nabla_v f = \mathbf{j} \implies \nabla_v^* \mathbf{j} = \nabla_v^* \nabla_v f = g^* - Bf,$$

which is precisely (4.1).

Proposition 4.4 (Solvability of the Kramers equation). *Under Assumption 1.1 and the assumption that*

$$\iint_{\mathbb{T}^d \times \mathbb{R}^d} g^* \, d\gamma(v) \, d\sigma(x) = 0,$$

there exists a unique solution f to (4.1) such that $(f)_{\mathbb{T}^d} = 0$, and f is given as the null minimizer of the functional $\mathcal{J}[f,\mathbf{j}]$ over pairs $(f,\mathbf{j})$ satisfying the constraint (4.12).

Before proving Proposition 4.4, we argue that one may assume that $\langle g^* \rangle_\gamma = 0$ as a function of x. For this, we require

Lemma 4.5. *Let $h \in L^2(\mathbb{T}^d)$ be given with $(h)_{\mathbb{T}^d} := \int_{\mathbb{T}^d} h(x)\, d\sigma(x) = 0$. Then there exists $g \in H^1_{\mathrm{hyp}}(\mathbb{T}^d)$ with $(g)_{\mathbb{T}^d} = 0$ such that*

$$(4.13) \qquad \langle v \cdot \nabla_x g + \mathbf{b}(x) \cdot \nabla_v g \rangle_\gamma(x) = h(x), \qquad \|g\|_{H^1_{\mathrm{hyp}}(\mathbb{T}^d)} \leqslant C \|h\|_{L^2(\mathbb{T}^d)}.$$

Suppose that we can solve (4.1) under the simplification $\langle g^* \rangle_\gamma = 0$. By Lemma 4.5 with $h = \langle g^* \rangle_\gamma$, we can find $g \in H^1_{\mathrm{hyp}}(\mathbb{T}^d)$ such that $\langle v \cdot \nabla_x g + \mathbf{b} \cdot \nabla_v g \rangle_\gamma = h$. Then, since $\langle -\Delta_v g + v \cdot \nabla_v g \rangle_\gamma = 0$, we can solve

$$-\Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f = g^* - \left(-\Delta_v g + v \cdot \nabla_v g + v \cdot \nabla_x g + \mathbf{b} \cdot \nabla_v g \right),$$

so $f + g$ solves (4.1). We now show that such a g exists, and in the argument below we always work under the assumption that $\langle g^* \rangle_\gamma = 0$. We shall occasionally use the notation $g^* \in L^2(\mathbb{T}^d; \dot{H}^{-1}_\gamma)$ to signify that $\langle g^* \rangle_\gamma = 0$.

Proof of Lemma 4.5. Let $\mathbf{f} \in H^1(\mathbb{T}^d; \mathbb{R}^d)$ be a solution to the problem[11]

$$\nabla_x \cdot \mathbf{f}(x) + \mathbf{b}(x) \cdot \mathbf{f}(x) = h(x).$$

Let $\xi(s) : \mathbb{R} \to \mathbb{R}$ be a compactly supported, smooth, odd function of a single variable such that $\int_{\mathbb{R}} \xi(s) s\, ds \neq 0$. Define $\xi_i : \mathbb{R}^d \to \mathbb{R}$ by

$$\xi_i(v) = \xi(v_i) \prod_{i' \neq i} \xi'(v_{i'}),$$

so that ξ_i is odd in v_i and even in all other $v_{i'}$ for $i' \neq i$. Under an appropriate normalization, we find that

$$\int_{\mathbb{R}^d} \partial_{v_j} \xi_i(v)\, d\gamma(v) = \int_{\mathbb{R}^d} v_j \xi_i(v)\, d\gamma(v) = \delta_{ij},$$

since $v_j \xi_i(v) d\gamma(v)$ is odd in v_i unless $i = j$, in which case it is even in all components of v. Define

$$g(x, v) = \mathbf{f}_i(x) \xi_i(v),$$

where we have used the summation convention over repeated indices. By the smoothness of the ξ_is and the $H^1(\mathbb{T}^d)$ regularity of $\mathbf{f}$, it is clear that $g \in H^1_{\mathrm{hyp}}(\mathbb{T}^d)$ with norm controlled by the sum of the respective H^1 norms of $\mathbf{f}$ and ξ. Furthermore, $(g)_{\mathbb{T}^d} = 0$ since for $1 \leqslant i \leqslant d$, ξ_i is odd in v_i. Now we may compute that

$$\begin{aligned}
\langle Bg \rangle_\gamma(x) &= \int_{\mathbb{R}^d} \left(v_j \partial_{x_j} g(x, v) + \mathbf{b}_j(x) \partial_{v_j} g(x, v) \right) d\gamma(v) \\
&= \int_{\mathbb{R}^d} \left(v_j \partial_{x_j} \mathbf{f}_i(x) \xi_i(v) + \mathbf{b}_j(x) \mathbf{f}_i(x) \partial_{v_j} \xi_i(v) \right) d\gamma(v) \\
&= \partial_i \mathbf{f}_i(x) + \mathbf{b}_i(x) \mathbf{f}_i(x) \\
&= h(x).
\end{aligned}$$

The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proof of Proposition 4.4. We split the argument into five steps.

Step 1. In this step, we show that the functional $\mathcal{J}$ is not uniformly equal to $+\infty$ and is uniformly convex on pairs $(f, \mathbf{j})$ satisfying the constraint (4.12). Let us denote the set of pairs satisfying the constraint by

$$\mathcal{A}(g^*) := \left\{ (f, \mathbf{j}) \in H^1_{\mathrm{hyp}}(\mathbb{T}^d) \times (L^2(\mathbb{T}^d; L^2_\gamma))^d : \nabla_v^* \mathbf{j} = g^* - Bf, \ (f)_{\mathbb{T}^d} = 0 \right\}.$$

First, since $g^* \in L^2(\mathbb{T}^d; \dot{H}^{-1}_\gamma)$, there exists $\mathbf{j} \in L^2(\mathbb{T}^d; L^2_\gamma)$ such that $g^* = A^* \mathbf{j}$. The pair $(0, \mathbf{j})$ belongs to $\mathcal{A}(g^*)$, and $\mathcal{J}(0, \mathbf{j}) < +\infty$.

---

[11]For example, one could argue as in the proof of Lemma 3.4 to produce $\mathbf{f}$ via the Lax-Milgram theorem satisfying the bound $\|\mathbf{f}\|_{H^1(\mathbb{T}^d)} \leqslant C \|h\|_{L^2(\mathbb{T}^d)}$.

We now demonstrate uniform convexity. Since for every $(f', \mathbf{j}') \in \mathcal{A}(g^*)$ and $(f, \mathbf{j}) \in \mathcal{A}(0)$,

$$(4.14) \qquad \frac{1}{2}\mathcal{J}[f' + f, \mathbf{j}' + \mathbf{j}] + \frac{1}{2}\mathcal{J}[f' - f, \mathbf{j}' - \mathbf{j}] - \mathcal{J}[f', \mathbf{j}'] = \mathcal{J}[f, \mathbf{j}],$$

it suffices to show that there exists $C(d) < \infty$ such that for every $(f, \mathbf{j}) \in \mathcal{A}(0)$,

$$(4.15) \qquad \mathcal{J}[f, \mathbf{j}] \geqslant C^{-1}\left(\|f\|^2_{H^1_{\mathrm{hyp}}(\mathbb{T}^d)} + \|\mathbf{j}\|^2_{L^2(\mathbb{T}^d; L^2_\gamma)} \right).$$

Expanding the square and using that $\nabla_v^* \mathbf{j} = -Bf$, we find

$$\mathcal{J}[f, \mathbf{j}] = \iint_{\mathbb{T}^d \times \mathbb{R}^d} \left(\frac{1}{2}|\nabla_v f|^2 + \frac{1}{2}|\mathbf{j}|^2 + fBf \right) dm.$$

Moreover, by (1.8), the term $\iint_{\mathbb{T}^d \times \mathbb{R}^d} fBf\, dm$ vanishes. Finally, from $-Bf = \nabla_v^* \mathbf{j}$, we have $\langle Bf \rangle_\gamma = 0$, and thus

$$\|v \cdot \nabla_x f\|_{L^2(\mathbb{T}^d; H_\gamma^{-1})} \leqslant \|Bf\|_{L^2(\mathbb{T}^d; H_\gamma^{-1})} + \|\mathbf{b}(x) \cdot \nabla_v f\|_{L^2(\mathbb{T}^d; H_\gamma^{-1})}$$
$$\leqslant C\|\mathbf{j}\|_{L^2(\mathbb{T}^d; L^2_\gamma)} + C\|\nabla_v f\|_{L^2(\mathbb{T}^d; L^2_\gamma)}.$$

Combining the last displays and Theorem 1.3 yields (4.15), and thus also the uniform convexity of the functional in (4.11).

Step 2. In this step, we rephrase the problem in terms of a perturbed convex minimization problem. Denote by $(f_1, \mathbf{j}_1)$ the unique minimizing pair of the functional $\mathcal{J}$ over $\mathcal{A}(g^*)$. We obviously have

$$\mathcal{J}[f_1, \mathbf{j}_1] \geqslant 0.$$

We now show that there is a one-to-one correspondence between solutions f of the Kramers equation and null minimizers $(f, \mathbf{j})$ of $\mathcal{J}$ satisfying the constraint (4.12): for every $f \in H^1_{\mathrm{hyp}}(\mathbb{T}^d)$ with $(f)_{\mathbb{T}^d} = 0$, we have

$$f \text{ solves } (4.1) \iff \mathcal{J}[f, \mathbf{j}_1] = 0.$$

Indeed, the implication $\implies$ is clear, since if f solves (4.1), then

$$(f, \nabla_v f) \in \mathcal{A}(g^*) \quad \text{and} \quad \mathcal{J}[f, \nabla_v f] = 0.$$

Conversely, if $\mathcal{J}[f_1, \mathbf{j}_1] = 0$, then by convexity we have that $f = f_1$ (assuming the mean-zero constraint from (4.12)), and

$$\nabla_v f_1 = \mathbf{j}_1 \qquad \text{a.e. in } \mathbb{T}^d \times \mathbb{R}^d.$$

Then since $\nabla_v^* \mathbf{j}_1 = g^* - Bf_1$, we recover that $f = f_1$ is indeed a solution of (4.1). In particular, the fact that there is at most one solution to (4.1) is clear.

To complete the proof, it thus remains to show that given the unique minimizing pair $(f_1, \mathbf{j}_1)$, we have that

$$(4.16) \qquad \mathcal{J}[f_1, \mathbf{j}_1] \leqslant 0.$$

We phrase this as a perturbed convex minimization problem for the functional G, which is defined for every $f^* \in L^2(\mathbb{T}^d; H_\gamma^{-1})$ with $(f^*)_{\mathbb{T}^d} = 0$ by

$$G(f^*) := \inf_{\substack{f \in H^1_{\mathrm{hyp}}(\mathbb{T}^d) \\ (f)_{\mathbb{T}^d} = 0}} \left(\iint_{\mathbb{T}^d \times \mathbb{R}^d} ff^*\, dm + \inf_{\substack{\mathbf{j} \in L^2(\mathbb{T}^d) \\ (f, \mathbf{j}) \in \mathcal{A}(f^* + g^*)}} \mathcal{J}[f, \mathbf{j}] \right).$$

To complete the proof, we must show that $G(0) \leqslant 0$. We decompose the argument into the next three steps.

Step 3. In this step, we show that G is convex and reduce the problem to showing that the convex dual of G is nonnegative. For every pair $(f, \mathbf{j})$ satisfying $(f, \mathbf{j}) \in \mathcal{A}(f^* + g^*)$, we have

$$(4.17) \qquad \nabla_v^* \mathbf{j} = f^* + g^* - Bf, \qquad (f)_{\mathbb{T}^d} = 0,$$

and so utilizing (1.8) we find that

$$
\begin{aligned}
\iint_{\mathbb{T}^d \times \mathbb{R}^d} f f^* \, dm + \mathcal{J}[f, \mathbf{j}] &= \iint_{\mathbb{T}^d \times \mathbb{R}^d} f f^* \, dm + \iint_{\mathbb{T}^d \times \mathbb{R}^d} \frac{1}{2} |\nabla_v f - \mathbf{j}|^2 \, dm \\
&= \iint_{\mathbb{T}^d \times \mathbb{R}^d} f f^* \, dm + \iint_{\mathbb{T}^d \times \mathbb{R}^d} \frac{1}{2} |\nabla_v f|^2 + \frac{1}{2} |\mathbf{j}|^2 - f \nabla_v^* \mathbf{j} \, dm \\
&= \iint_{\mathbb{T}^d \times \mathbb{R}^d} f f^* \, dm + \iint_{\mathbb{T}^d \times \mathbb{R}^d} \frac{1}{2} |\nabla_v f|^2 + \frac{1}{2} |\mathbf{j}|^2 - f(f^* + g^* - Bf) \, dm \\
&= \iint_{\mathbb{T}^d \times \mathbb{R}^d} \frac{1}{2} |\nabla_v f|^2 + \frac{1}{2} |\mathbf{j}|^2 - g^* f \, dm.
\end{aligned}
$$

Taking the infimum over all $(f, \mathbf{j})$ satisfying the affine constraint $(f, \mathbf{j}) \in \mathcal{A}(f^* + g^*)$, we obtain the quantity $G(f^*)$. We thus infer that G is convex in the variable f^*. By Lemma 4.5, given $f^* \in L^2(\mathbb{T}^d; H_\gamma^{-1})$ with vanishing mean, we may find $f_0 \in H_{\mathrm{hyp}}^1(\mathbb{T}^d)$ such that $\langle Bf_0 \rangle_\gamma = \langle f^* + g^* \rangle_\gamma = \langle f^* \rangle_\gamma$. Then since $\langle f^* + g^* - Bf_0 \rangle_\gamma = 0$, we may find $\mathbf{j} \in (L^2(\mathbb{T}^d; L_\gamma^2))^d$ such that $\nabla_v^* \mathbf{j} = f^* + g^* - Bf_0$, and we see that the function G is also locally bounded above. These two properties imply that G is lower semi-continuous, see [22, Lemma I.2.1 and Corollary I.2.2]. We denote by G^* the convex dual of G, defined for every $h \in L^2(\mathbb{T}^d; H_\gamma^1)$ with $(h)_{\mathbb{T}^d} = 0$ by

$$G^*(h) := \sup_{\substack{f^* \in L^2(\mathbb{T}^d; H_\gamma^{-1}) \\ (f^*)_{\mathbb{T}^d} = 0}} \left(-G(f^*) + \iint_{\mathbb{T}^d \times \mathbb{R}^d} h f^* \, dm \right),$$

and by G^{**} the bidual of G. Since G is lower semi-continuous, we have that $G^{**} = G$ (see [22, Proposition I.4.1]), and in particular,

$$G(0) = G^{**}(0) = \sup_{\substack{h \in L^2(\mathbb{T}^d; H_\gamma^1) \\ (h)_{\mathbb{T}^d} = 0}} (-G^*(h)).$$

In order to prove that $G(0) \leqslant 0$, it therefore suffices to show that

$$(4.18) \qquad \forall h \in L^2(\mathbb{T}^d; H_\gamma^1) \text{ with } (h)_{\mathbb{T}^d} = 0, \qquad G^*(h) \geqslant 0.$$

Step 4. In this step we show that

$$(4.19) \qquad G^*(h) < +\infty \implies h \in H_{\mathrm{hyp}}^1(\mathbb{T}^d).$$

We rewrite $G^*(h)$ in the form

$$(4.20) \qquad G^*(h) = \sup \left\{ \iint_{\mathbb{T}^d \times \mathbb{R}^d} \left(-\frac{1}{2} |\nabla_v f - \mathbf{j}|^2 - f f^* + h f^* \right) dm \right\},$$

where the supremum is over every $f \in H_{\mathrm{hyp}}^1(\mathbb{T}^d)$, $\mathbf{j} \in L^2(\mathbb{T}^d; L_\gamma^2)^d$ and $f^* \in L^2(\mathbb{T}^d; H_\gamma^{-1})$ satisfying the constraint (4.17). Given f with $(f)_{\mathbb{T}^d} = 0$, we choose to restrict the supremum above to $f^* := Bf$ and $\mathbf{j} = \mathbf{j}_0$ the solution of $\nabla_v^* \mathbf{j}_0 = g^*$. Recall that such a

$\mathbf{j}_0 \in L^2(\mathbb{T}^d; L^2_\gamma)^d$ exists since $\langle g^* \rangle_\gamma = 0$. With such choices of f^* and $\mathbf{j}$, the constraint (4.17) is satisfied, and we obtain that

$$G^*(h) \geqslant \sup \left\{ \iint_{\mathbb{T}^d \times \mathbb{R}^d} \left(-\frac{1}{2} |\nabla_v f - \mathbf{j}_0|^2 - fBf + hBf \right) dm \, : \, f \in H^1_{\mathrm{hyp}}(\mathbb{T}^d) \, , \, (f)_{\mathbb{T}^d} = 0 \right\} .$$

Recalling that $\iint fBf \, dm = 0$, and using that $C_0^\infty(\mathbb{T}^d \times \mathbb{R}^d)$ is dense in $H^1_{\mathrm{hyp}}(\mathbb{T}^d)$, we deduce that

$$G^*(h) \quad \geqslant \quad \sup \left\{ \iint_{\mathbb{T}^d \times \mathbb{R}^d} \left(-\frac{1}{2} |\nabla_v f - \mathbf{j}_0|^2 + hBf \right) dm \, : \, f \in C_c^\infty(\mathbb{T}^d \times \mathbb{R}^d) \, , \, (f)_{\mathbb{T}^d} = 0 \right\} .$$

Then the assumption of $G^*(h) < \infty$ implies that

$$\sup \left\{ \iint_{\mathbb{T}^d \times \mathbb{R}^d} hBf \, dm \, : \, f \in C_c^\infty(\mathbb{T}^d \times \mathbb{R}^d) \, , \, (f)_{\mathbb{T}^d} = 0 \, , \, \|f\|_{L^2(\mathbb{T}^d; H^1_\gamma)} \leqslant 1 \right\} < \infty \, .$$

This then shows that the distribution Bh belongs to the dual of $L^2(\mathbb{T}^d; H^1_\gamma)$, which is $L^2(\mathbb{T}^d; H^{-1}_\gamma)$. Since

$$v \cdot \nabla_x h = Bh - \mathbf{b} \cdot \nabla_v h \, ,$$

the proof of (4.19) is complete.

Step 5. In place of (4.18), we have left to show that

$$(4.21) \qquad\qquad \forall h \in H^1_{\mathrm{hyp}}(\mathbb{T}^d) \text{ with } (h)_{\mathbb{T}^d} = 0 \, , \qquad G^*(h) \geqslant 0 \, .$$

Since $Bf \in L^2(\mathbb{T}^d; H^{-1}_\gamma)$, we may replace f^* by $f^* + Bf$ in the variational formula (4.20) for G^* to get that

$$(4.22) \qquad G^*(h) = \sup \left\{ \iint_{\mathbb{T}^d \times \mathbb{R}^d} \left(-\frac{1}{2} |\nabla_v f - \mathbf{j}|^2 + (h - f)(f^* + Bf) \right) dm \right\} ,$$

where the supremum is now over every $f \in H^1_{\mathrm{hyp}}(\mathbb{T}^d)$, $\mathbf{j} \in L^2(\mathbb{T}^d; L^2_\gamma)^d$ and $f^* \in L^2(\mathbb{T}^d; H^{-1}_\gamma)$ satisfying the constraint

$$(4.23) \qquad\qquad \nabla_v^* \mathbf{j} = f^* + g^* \, , \qquad (f)_{\mathbb{T}^d} = 0 \, .$$

Setting $f = h$ in (4.22), we find that

$$G^*(h) \geqslant \sup \left\{ \iint_{\mathbb{T}^d \times \mathbb{R}^d} -\frac{1}{2} |\nabla_v h - \mathbf{j}|^2 \, dm \right\} ,$$

with the supremum ranging over all $f^* \in L^2(\mathbb{T}^d; H^{-1}_\gamma)$ and $\mathbf{j} \in L^2(\mathbb{T}^d; L^2_\gamma)^d$ satisfying the constraint (4.23). We now simply select $\mathbf{j} = \nabla_v h \in L^2(\mathbb{T}^d; L^2_\gamma)^d$ and

$$f^* = \nabla_v^* \mathbf{j} - g^* \in L^2(\mathbb{T}^d; H^{-1}_\gamma) \, ,$$

at which point we conclude that $G^*(h) \geqslant 0$. $\qquad\qquad\square$

5. Interior regularity of solutions

In this subsection, we use energy methods to obtain interior regularity estimates for solutions of the equation

$$(5.1) \qquad\qquad -\Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f + cf = f^* \, .$$

In analogy to the classical theory for uniformly elliptic equations (such as the Laplace or Poisson equations), we obtain an appropriate version of the Caccioppoli inequality, apply it iteratively to obtain H^1_{hyp} estimates on all spatial derivatives of the solution, and then apply the Hörmander and Sobolev inequalities to obtain pointwise estimates. In particular, we obtain higher regularity estimates—strong enough to imply that our

weak solutions are C^∞—without resorting to sophisticated theory for pseudodifferential operators.

We begin with a version of the Caccioppoli inequality for the equation (5.1).

Lemma 5.1 (Caccioppoli inequality)**.** *Suppose that* $r > 0$, $\mathbf{b} \in L^\infty(B_r; L^\infty(\mathbb{R}^d; \mathbb{R}^d))$, $c \in L^\infty(B_r; L^\infty(\mathbb{R}^d))$, *and the pair* $(f, f^*) \in L^2(B_r; H^1_\gamma) \times L^2(B_r; H^{-1}_\gamma)$ *satisfies the equation*

$$(5.2) \qquad -\Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f + c f = f^* \quad in \ B_r \times \mathbb{R}^d.$$

Then $f \in H^1_{\mathrm{hyp}}(B_r)$, *and there exists* $C\left(d, r, \|\mathbf{b}\|_{L^\infty(B_r; L^\infty(\mathbb{R}^d))}, \|c\|_{L^\infty(B_r; L^\infty(\mathbb{R}^d))}\right) < \infty$ *such that*

$$(5.3) \qquad \|\nabla_v f\|_{L^2(B_{r/2}; L^2_\gamma)} + \|v \cdot \nabla_x f\|_{L^2(B_{r/2}; H^{-1}_\gamma)} \leqslant C \|f\|_{L^2(B_r; L^2_\gamma)} + C \|f^*\|_{L^2(B_r; H^{-1}_\gamma)}.$$

Proof. The PDE (5.2) guarantees that $f \in L^2(B_r; H^1_\gamma)$ belongs qualitatively to $H^1_{\mathrm{hyp}}(B_r)$.

Step 1. We show that there exists $C(d) < \infty$ such that

$$(5.4) \qquad \|\nabla_v f\|_{L^2(B_{r/2}; L^2_\gamma)} \leqslant C\left(\frac{1}{r} + \|\mathbf{b}\|_{L^\infty(B_r \times \mathbb{R}^d)} + \|c\|^{1/2}_{L^\infty(B_r \times \mathbb{R}^d)}\right) \|f\|_{L^2(B_r; L^2_\gamma)}$$
$$+ C(1 + r) \|f^*\|_{L^2(B_r; H^{-1}_\gamma)}.$$

Select a smooth cutoff function $\phi \in C^\infty_c(B_r)$ which is compactly supported in B_r and satisfies $0 \leqslant \phi \leqslant 1$ in B_r, $\phi \equiv 1$ on $B_{r/2}$ and $\|\nabla\phi\|_{L^\infty(B_r)} \leqslant 8r^{-1}$. Testing the equation (5.2) with $(x, v) \mapsto \phi^2(x) f(x, v)$ yields

$$(5.5) \qquad \int_{B_r \times \mathbb{R}^d} \phi^2 |\nabla_v f|^2 \, dx \, d\gamma = \int_{B_r \times \mathbb{R}^d} \phi^2 f f^* \, dx \, d\gamma - \int_{B_r \times \mathbb{R}^d} \phi^2 f v \cdot \nabla_x f \, dx \, d\gamma$$

$$(5.6) \qquad - \int_{B_r \times \mathbb{R}^d} \phi^2 f \mathbf{b} \cdot \nabla_v f \, dx \, d\gamma - \int_{B_r \times \mathbb{R}^d} \phi^2 c f^2 \, dx \, d\gamma.$$

We estimate each of the terms on the right-hand side of (5.5) separately.

For the first term on the right side of (5.5), we use

$$\left| \int_{B_r \times \mathbb{R}^d} \phi^2 f f^* \, dx \, d\gamma \right| \leqslant \|\phi^2 f\|_{L^2(B_r; H^1_\gamma)} \|f^*\|_{L^2(B_r; H^{-1}_\gamma)}$$
$$\leqslant \left(\|\phi^2 \nabla_v f\|_{L^2(B_r; L^2_\gamma)} + \|f\|_{L^2(B_r; L^2_\gamma)}\right) \|f^*\|_{L^2(B_r; H^{-1}_\gamma)}$$

and then apply Young's inequality to obtain

$$(5.7) \quad \left| \int_{B_r \times \mathbb{R}^d} \phi^2 f f^* \, dx \, d\gamma \right|$$
$$\leqslant \frac{1}{6} \int_{B_r \times \mathbb{R}^d} \phi^2 |\nabla_v f|^2 \, dx \, d\gamma + \frac{C}{r^2} \int_{B_r \times \mathbb{R}^d} f^2 \, dx \, d\gamma + C(1 + r^2) \|f^*\|^2_{L^2(B_r; H^{-1}_\gamma)}.$$

For the second term on the right side of (5.5), we integrate by parts to find

$$-\int_{B_r \times \mathbb{R}^d} \phi^2 f v \cdot \nabla_x f \, dx \, d\gamma = -\int_{B_r \times \mathbb{R}^d} \phi^2 v \cdot \nabla_x \left(\frac{1}{2} f^2\right) dx \, d\gamma$$
$$= \int_{B_r \times \mathbb{R}^d} \phi \nabla_x \phi \cdot v f^2 \, dx \, d\gamma$$
$$= \int_{B_r \times \mathbb{R}^d} \phi(x) \nabla_x \phi(x) \cdot v \exp\left(-\frac{|v|^2}{2}\right) f^2(x, v) \, dx \, dv$$
$$= -\int_{B_r \times \mathbb{R}^d} 2 f \phi \nabla_x \phi \cdot \nabla_v f \, dx \, d\gamma.$$

Thus, by Young's inequality,

$$(5.8) \quad \left| \int_{B_r \times \mathbb{R}^d} \phi^2 f v \cdot \nabla_x f \, dx \, d\gamma \right| \leqslant \frac{1}{6} \int_{B_r \times \mathbb{R}^d} \phi^2 \left| \nabla_v f \right|^2 \, dx \, d\gamma + C \int_{B_r \times \mathbb{R}^d} f^2 \left| \nabla_x \phi \right|^2 \, dx \, d\gamma$$

$$\leqslant \frac{1}{6} \int_{B_r \times \mathbb{R}^d} \phi^2 \left| \nabla_v f \right|^2 \, dx \, d\gamma + \frac{C}{r^2} \int_{B_r \times \mathbb{R}^d} f^2 \, dx \, d\gamma \,.$$

For the third term on the right side of (5.5), we use Young's inequality to obtain

$$(5.9) \quad \left| \int_{B_r \times \mathbb{R}^d} \phi^2 f \mathbf{b} \cdot \nabla_v f \right| \leqslant \frac{1}{6} \int_{B_r \times \mathbb{R}^d} \phi^2 \left| \nabla_v f \right|^2 \, dx \, d\gamma + C \int_{B_r \times \mathbb{R}^d} \phi^2 f^2 \left| \mathbf{b} \right|^2 \, dx \, d\gamma$$

$$\leqslant \frac{1}{6} \int_{B_r \times \mathbb{R}^d} \phi^2 \left| \nabla_v f \right|^2 \, dx \, d\gamma + C \left\| \mathbf{b} \right\|_{L^\infty(B_r \times \mathbb{R}^d)}^2 \int_{B_r \times \mathbb{R}^d} f^2 \, dx \, d\gamma \,.$$

To conclude, we combine (5.5)-(5.9) and the obvious estimate on the final term to obtain

$$\int_{B_r \times \mathbb{R}^d} \phi^2 \left| \nabla_v f \right|^2 \, dx \, d\gamma \leqslant \frac{2}{3} \int_{B_r \times \mathbb{R}^d} \phi^2 \left| \nabla_v f \right|^2 \, dx \, d\gamma + \frac{C}{r^2} \int_{B_r \times \mathbb{R}^d} f^2 \, dx \, d\gamma$$

$$+ C(1 + r^2) \left\| f^* \right\|_{L^2(B_r; H_\gamma^{-1})}^2$$

$$+ C \left(\left\| \mathbf{b} \right\|_{L^\infty(B_r \times \mathbb{R}^d)}^2 + \left\| c \right\|_{L^\infty(B_r \times \mathbb{R}^d)} \right) \int_{B_r \times \mathbb{R}^d} f^2 \, dx \, d\gamma.$$

The first term on the right may now be reabsorbed on the left. Using that $\phi = 1$ on $B_{r/2}$, we thus obtain (5.4). The analysis in Step 1 is enough to conclude that $f \in H_{\mathrm{hyp}}^1(B_{r/2})$ and the gradient bound in (5.3).

Step 2. We show that there exists $C(d) < \infty$ such that

$$(5.10) \quad \left\| v \cdot \nabla_x f \right\|_{L^2(B_{r/2}; H_\gamma^{-1})} \leqslant C \left(1 + \left\| \mathbf{b} \right\|_{L^\infty(B_{r/2} \times \mathbb{R}^d)} \right) \left\| \nabla_v f \right\|_{L^2(B_{r/2}; L_\gamma^2)}$$

$$+ C \left\| c \right\|_{L^\infty(B_{r/2} \times \mathbb{R}^d)} \left\| f \right\|_{L^2(B_{r/2}; L_\gamma^2)} + C \left\| f^* \right\|_{L^2(B_{r/2}; H_\gamma^{-1})} \,.$$

This estimate may be combined with (5.4) to obtain the bound for the second term in (5.3), which completes the proof of the lemma.

To obtain (5.10), we test the equation (5.2) with $w \in L^2(B_{r/2}; H_\gamma^1)$ to find that

$$\int_{B_r \times \mathbb{R}^d} w \left(v \cdot \nabla_x f \right) dx \, d\gamma = - \int_{B_r \times \mathbb{R}^d} \nabla_v f \cdot (\nabla_v w + w \mathbf{b}) + \int_{B_r \times \mathbb{R}^d} w f^* \, dx \, d\gamma - \int_{B_r \times \mathbb{R}^d} c w f \, dx \, d\gamma \,.$$

We deduce that

$$\left| \int_{B_r \times \mathbb{R}^d} w \left(v \cdot \nabla_x f \right) dx \, d\gamma \right|$$

$$\leqslant \left\| \nabla_v f \right\|_{L^2(B_{r/2}; L_\gamma^2)} \left(\left\| \nabla_v w \right\|_{L^2(B_{r/2}; L_\gamma^2)} + \left\| \mathbf{b} \right\|_{L^\infty(B_{r/2} \times \mathbb{R}^d)} \left\| w \right\|_{L^2(B_{r/2}; L_\gamma^2)} \right)$$

$$+ \left\| w \right\|_{L^2(B_{r/2}; H_\gamma^1)} \left\| f^* \right\|_{L^2(B_{r/2}; H_\gamma^{-1})} + \left\| c \right\|_{L^\infty(B_r \times \mathbb{R}^d)} \left\| f \right\|_{L^2(B_r; L_\gamma^2)} \left\| w \right\|_{L^2(B_r; L_\gamma^2)}.$$

Taking the supremum over $w \in L^2(B_{r/2}; H_\gamma^1)$ with $\left\| w \right\|_{L^2(B_{r/2}; H_\gamma^1)} \leqslant 1$ yields (5.10).

The combination of (5.4) and (5.10) yields (5.3). $\qquad \square$

In the next lemma, under appropriate regularity conditions on the coefficients, we differentiate the equation (5.1) with respect to x_i to obtain an equation for $\partial_{x_i} f$, and then apply the previous lemma to obtain an interior H_{hyp}^1 estimate for $\partial_{x_i} f$. We need to essentially differentiate the equation a fractional number of times (cf. [37, 38]).

Lemma 5.2 (Differentiating in x). *Fix $r \in (0, \infty)$ and coefficients $\mathbf{b} \in C^{0,1}(B_r \times \mathbb{R}^d; \mathbb{R}^d)$, $c \in C^{0,1}(B_r \times \mathbb{R}^d; \mathbb{R})$. Suppose that $f^* \in H^1(B_r; H_\gamma^{-1})$ and $f \in H_{\mathrm{hyp}}^1(B_r)$ satisfy*

$$(5.11) \quad -\Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f + c f = f^* \quad in \ B_r \times \mathbb{R}^d$$

Then, for each $i \in \{1, \ldots, d\}$, the function $h := \partial_{x_i} f$ belongs to $H^1_{\mathrm{hyp}}(B_{r'})$ for all $r' \in (0, r)$ and satisfies

$$(5.12) \quad -\Delta_v h + v \cdot \nabla_v h + v \cdot \nabla_x h + \mathbf{b} \cdot \nabla_v h + ch = \partial_{x_i} f^* - \partial_{x_i} \mathbf{b} \cdot \nabla_v f - \partial_{x_i} c\, f \quad in \ B_{r'} \times \mathbb{R}^d.$$

Moreover, there exists $C\left(d, r, \|\mathbf{b}\|_{C^{0,1}(B_r \times \mathbb{R}^d)}, \|c\|_{C^{0,1}(B_r \times \mathbb{R}^d)}\right) < \infty$ such that

$$(5.13) \qquad\qquad \|\partial_{x_i} f\|_{H^1_{\mathrm{hyp}}(B_{r/2})} \leqslant C \|f\|_{L^2(B_r; L^2_\gamma)} + C \|f^*\|_{H^1(B_r; H^{-1}_\gamma)}.$$

Proof. The argument is by induction on the fractional exponent of differentiability of f in the spatial variable x. Essentially, we want to differentiate the equation a fractional amount (almost $1/3$ times), apply the Caccioppoli inequality to the fractional derivative, and then iterate until we have one full spatial derivative.

Step 1. We first prove that, for every $(f, f^*) \in H^1_{\mathrm{hyp}}(B_r) \times H^1(B_r, H^{-1}_\gamma)$ satisfying (5.11), there exists $C\left(d, r, \|\mathbf{b}\|_{C^{0,1}(B_r \times \mathbb{R}^d)}, \|c\|_{C^{0,1}(B_r \times \mathbb{R}^d)}\right) < \infty$ such that f belongs to $H^1(B_{r/2}; H^1_\gamma)$ and satisfies the estimate

$$(5.14) \qquad\qquad \|\nabla_x f\|_{L^2(B_{r/2}; H^1_\gamma)} \leqslant C \|f\|_{L^2(B_r; L^2_\gamma)} + C \|f^*\|_{H^1(B_r; H^{-1}_\gamma)}.$$

Suppose that $\alpha_0 \in [0, 1)$ is such that the following statement is valid: For every $\alpha \in [0, \alpha_0]$, $r > 0$, and pair $(f, f^*) \in H^1_{\mathrm{hyp}}(B_r) \times H^\alpha(B_r, H^{-1}_\gamma)$ satisfying (5.11), we have $f \in H^\alpha(B_{r/2}; H^1_\gamma)$ and, for $C\left(d, r, \|\mathbf{b}\|_{C^{0,1}(B_r \times \mathbb{R}^d)}, \|c\|_{C^{0,1}(B_r \times \mathbb{R}^d)}, \alpha\right) < \infty$, the estimate

$$(5.15) \qquad\qquad \|f\|_{H^\alpha(B_{r/2}; H^1_\gamma)} \leqslant C \|f\|_{L^2(B_r; L^2_\gamma)} + C \|f^*\|_{H^\alpha(B_r; H^{-1}_\gamma)}.$$

We argue that the statement is also valid for $\min(\alpha_0 + 1/3 - \delta, 1)$ in place of α_0 for all $\delta \in (0, 1/3)$. Note that this statement is clearly valid for $\alpha_0 = 0$ by the Caccioppoli inequality (Lemma 5.1).

Fix $\alpha \in [0, \alpha_0]$ and a pair

$$(f, f^*) \in H^1_{\mathrm{hyp}}(B_r) \times H^\alpha(B_r, H^{-1}_\gamma)$$

satisfying (5.11), an index $i \in \{1, \ldots, d\}$, and a cutoff function $\phi \in C_c^\infty(B_{r/2})$ with $0 \leqslant \phi \leqslant 1$ and $\phi \equiv 1$ on $B_{r/4}$. Define the functions

$$\widetilde{f} := \phi^2 f,$$
$$\widetilde{f}^* := \phi^2 f^* + 2f\phi\, v \cdot \nabla_x \phi.$$

Observe that $\widetilde{f} \in H^1_{\mathrm{hyp}}(\mathbb{R}^d)$ and $\widetilde{f}^* \in H^\alpha(\mathbb{R}^d; H^{-1}_\gamma)$ are compactly supported in B_r and satisfy

$$\|\widetilde{f}\|_{H^\alpha(\mathbb{R}^d; L^2_\gamma)} \leqslant C \|f\|_{H^\alpha(B_r; L^2_\gamma)},$$
$$\|\widetilde{f}^*\|_{H^\alpha(\mathbb{R}^d; H^{-1}_\gamma)} \leqslant C\left(\|f^*\|_{H^\alpha(B_r; H^{-1}_\gamma)} + \|f\|_{H^\alpha(B_r; L^2_\gamma)}\right),$$

and the PDE (5.1) in $\mathbb{R}^d \times \mathbb{R}^d$.

Next, we mollify. This step ensures that the function qualitatively belongs to good enough spaces to justify the computations (the analogous step in Nirenberg's method is finite differences). Define

$$\overline{f} = \widetilde{f} *_x \psi^\varepsilon,$$
$$\overline{f}^* = \widetilde{f}^* *_x \psi^\varepsilon - [\psi^\varepsilon *_x, \mathbf{b} \cdot] \nabla_v \widetilde{f} - [\psi^\varepsilon *_x, c] \widetilde{f},$$

where ψ^ε is an appropriate mollification at scale ε. Then $(\overline{f}, \overline{f}^*)$ satisfies the PDE (5.1) in $\mathbb{R}^d \times \mathbb{R}^d$. We have that

$$\left\| (1 - \Delta_x)^{\alpha/2} \overline{f} \right\|_{L^2(\mathbb{R}^d; L^2_\gamma)} \leqslant C \left\| \widetilde{f} \right\|_{H^\alpha(\mathbb{R}^d; L^2_\gamma)} \tag{5.16}$$

$$\left\| (1 - \Delta_x)^{\alpha/2} \overline{f}^* \right\|_{L^2(\mathbb{R}^d; H^{-1}_\gamma)} \leqslant C \left\| \widetilde{f}^* \right\|_{H^\alpha(\mathbb{R}^d; H^{-1}_\gamma)} + C \left\| \widetilde{f} \right\|_{H^\alpha(\mathbb{R}^d; L^2_\gamma)}, \tag{5.17}$$

since $[\psi^\varepsilon *_x, \mathbf{b}\cdot]$ and $[\psi^\varepsilon *_x, c]$ are $H^\alpha(\mathbb{R}^d; H^{-1}_\gamma)$-bounded for all $\alpha \in [0,1]$ while $\mathbf{b}$ and c are Lipschitz. We apply $(1 - \Delta_x)^{\alpha/2}$ to the PDE (5.1) satisfied by $(\overline{f}, \overline{f}^*)$ and denote $f_\alpha = (1 - \Delta_x)^{\alpha/2} \overline{f}$. We have that f_α satisfies the equation

$$-\Delta_v f_\alpha + v \cdot \nabla_v f_\alpha + v \cdot \nabla_x f_\alpha + \mathbf{b} \cdot \nabla_v f_\alpha + c f_\alpha$$
$$= (1 - \Delta)^{\alpha/2} \overline{f}^* - [(1 - \Delta)^{\alpha/2}, \mathbf{b}\cdot] \nabla_v \overline{f} - [(1 - \Delta)^{\alpha/2}, c] \overline{f}$$

in $\mathbb{R}^d \times \mathbb{R}^d$. The Cacciopoli inequality for $f_\alpha \in L^2(\mathbb{R}^d; H^1_\gamma)$, the Hörmander inequality, and (5.16)-(5.17) give

$$\|f_\alpha\|_{H^{1/3-\delta}(\mathbb{R}^d; L^2_\gamma)} + \|f_\alpha\|_{L^2(\mathbb{R}^d; H^1_\gamma)} \leqslant C \|\widetilde{f}\|_{H^\alpha(\mathbb{R}^d; L^2_\gamma)} + C \|\widetilde{f}^*\|_{H^\alpha(\mathbb{R}^d; H^{-1}_\gamma)} \tag{5.18}$$

for all $\delta \in (0, 1/3)$, where C depends on δ. Sending the mollification parameter $\varepsilon \to 0^+$ completes the induction and the proof. We emphasize that this induction demonstrates that $\partial_{x_i} f \in L^2(B_{r'}; H^1_\gamma)$ for all $r' < r$, where f is a function satisfying the hypotheses of Lemma 5.2. Once this is known, one may plainly differentiate the equation in ∂_{x_i} and apply Caccioppoli's inequality to conclude. $\qquad\square$

Lemma 5.3 (Differentiating in v). *Fix $r \in (0, \infty)$ and coefficients $\mathbf{b} \in C^{0,1}(B_r \times \mathbb{R}^d; \mathbb{R}^d)$, $c \in C^{0,1}(B_r \times \mathbb{R}^d; \mathbb{R})$. Suppose that $f^* \in H^1(B_r; L^2_\gamma)$ and $f \in H^1_{\mathrm{hyp}}(B_r)$ satisfy*

$$- \Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f + c f = f^* \quad in \ B_r \times \mathbb{R}^d. \tag{5.19}$$

Then, for each $i \in \{1, \ldots, d\}$, the function $h := \partial_{v_i} f$ belongs to $H^1_{\mathrm{hyp}}(B_{r'})$ for all $r' \in (0, r)$ and satisfies

$$- \Delta h + v \cdot \nabla_v h + v \cdot \nabla_x h + \mathbf{b} \cdot \nabla_v h + (c + 1) h = h^* \quad in \ B_{r'} \times \mathbb{R}^d, \tag{5.20}$$

where

$$h^* := \partial_{v_i} f^* - \partial_{x_i} f - (\partial_{v_i} \mathbf{b}) \cdot \nabla_v f - (\partial_{v_i} c) f. \tag{5.21}$$

Moreover, there exists $C \left(d, r, \|\mathbf{b}\|_{C^{0,1}(B_r \times \mathbb{R}^d)}, \|c\|_{C^{0,1}(B_r \times \mathbb{R}^d)} \right) < \infty$ such that

$$\|\partial_{v_i} f\|_{H^1_{\mathrm{hyp}}(B_{r/2})} \leqslant C \|f\|_{L^2(B_r; L^2_\gamma)} + C \|f^*\|_{H^1(B_r; L^2_\gamma)}. \tag{5.22}$$

Proof. The standard procedure is to differentiate the equation and apply Caccioppoli's inequality. This introduces a forcing term h^*, defined in (5.21), which contains $\partial_{x_i} f$, and this is why we improve the spatial regularity beforehand in Lemma 5.2. That is, we already know

$$\|f\|_{H^1_{\mathrm{hyp}}(B_{r'})} + \|\partial_{x_i} f\|_{L^2(B_{r'}; H^1_\gamma)} \leqslant C \left(\|f\|_{L^2(B_r; L^2_\gamma)} + \|f^*\|_{H^1(B_r; H^{-1}_\gamma)} \right),$$

as in Lemma 5.2, where $r' = 7r/8$. In addition to this observation, we require a cut-off and mollification procedure to compensate for the fact that we did not assume *qualitatively* that $\partial_{v_i} f \in L^2(B_r; H^1_\gamma)$, which would be enough to make the energy estimate rigorous.

For $\ell \geqslant 1$, consider a standard cut-off function φ^ℓ in v at scale ℓ. Define

$$\widetilde{f} = \varphi^\ell f$$
$$\widetilde{f}^* = \varphi^\ell f^* - 2 \nabla_v f \cdot \nabla_v \varphi^\ell - f \Delta_v \varphi^\ell + f(v \cdot \nabla_v \varphi^\ell + \mathbf{b} \cdot \nabla_v \varphi^\ell),$$

where we suppress the dependence on ℓ in the notation. Then $(\widetilde{f}, \widetilde{f}^*)$ solves (5.1) in $B_{r'} \times \mathbb{R}^d$, and it is not difficult to verify that

$$\|\widetilde{f}\|_{L^2(B_{r'};H^1_\gamma)} \leqslant C\|f\|_{L^2(B_{r'};H^1_\gamma)},$$

$$\|\partial_{x_i}\widetilde{f}\|_{L^2(B_{r'};L^2_\gamma)} \leqslant \|\partial_{x_i}f\|_{L^2(B_{r'};L^2_\gamma)},$$

and

$$\|\widetilde{f}^*\|_{L^2(B_{r'};L^2_\gamma)} \leqslant C\left(\|f\|_{L^2(B_{r'};H^1_\gamma)} + \|f^*\|_{L^2(B_{r'};L^2_\gamma)}\right).$$

Next, we mollify. Let ψ^ε be a standard mollification function in v at scale $0 < \varepsilon \ll 1$. Define

$$\overline{f} = \psi^\varepsilon *_v \widetilde{f}$$

(5.23) $$\overline{f}^* = \psi^\varepsilon *_v \widetilde{f}^* - [\psi^\varepsilon_v *_v, v\cdot](\nabla_v \widetilde{f} + \nabla_x \widetilde{f}) - ([\psi^\varepsilon *_v, \mathbf{b}\cdot]\nabla_v \widetilde{f}) - [\psi^\varepsilon *_v, c]\widetilde{f},$$

where again we suppress the dependence on ℓ, ε in the notation. Then $(\overline{f}, \overline{f}^*)$ is well defined in $B_{r'} \times \mathbb{R}^d$ and solves (5.1) there.

We highlight a few features of the cut-off and mollification procedure. Translations of L^2_γ functions may not belong to L^2_γ, due to the superexponential nature of the weight (compare with exponential weights $e^{-c\langle v \rangle}$). Hence, mollification is not well behaved on L^2_γ. The velocity cut-off φ^ℓ tames this issue. This cut-off has the additional benefit of taming commutators with v which occur naturally in the force term $\overline{f}^*$.

We claim

(5.24) $$\limsup_{\varepsilon \to 0^+} \|\overline{f}\|_{L^2(B_{r'};H^1_\gamma)} \leqslant \|\widetilde{f}\|_{L^2(B_{r'};H^1_\gamma)},$$

$$\limsup_{\varepsilon \to 0^+} \|\partial_{x_i}\overline{f}\|_{L^2(B_{r'};L^2_\gamma)} \leqslant \|\partial_{x_i}\widetilde{f}\|_{L^2(B_{r'};L^2_\gamma)},$$

and, more subtly,

(5.25) $$\limsup_{\varepsilon \to 0^+} \|\overline{f}^*\|_{L^2(B_{r'};L^2_\gamma)} \leqslant C\left(\|\widetilde{f}\|_{L^2(B_{r'};H^1_\gamma)} + \|\widetilde{f}^*\|_{L^2(B_{r'};L^2_\gamma)}\right),$$

where (5.24) and (5.25) are for fixed ℓ. Both estimates in (5.24) are evident due to the support properties of $\widetilde{f}$, so we focus on (5.25). For each fixed ℓ, we have

$$\left\|([\psi^\varepsilon *_v, \mathbf{b}\cdot]\nabla_v \widetilde{f}) + [\psi^\varepsilon *_v, c]\widetilde{f}\right\|_{L^2(B_{r'};L^2_\gamma)} \to 0.$$

as $\varepsilon \to 0^+$.[12] Here, we use that the coefficients are Lipschitz and $\widetilde{f}$ is compactly supported. It remains to analyze the second term in (5.23). From the compact support, we may replace v by $\varphi^{2\ell}v$. Then

(5.26) $$\left\|[\psi^\varepsilon_v *_v, (\varphi^{2\ell}v)\cdot](\nabla_v \widetilde{f})\right\|_{L^2(B_{r'};L^2_\gamma)} \to 0$$

(5.27) $$\left\|[\psi^\varepsilon_v *_v, (\varphi^{2\ell}v)\cdot](\nabla_x \widetilde{f})\right\|_{L^2(B_{r'};L^2_\gamma)} \to 0$$

as $\varepsilon \to 0^+$ for fixed ℓ.

Finally, we define $\overline{h} = \partial_{v_i}\overline{f}$ and

(5.28) $$\overline{h}^* := \partial_{v_i}\overline{f}^* - \partial_{x_i}\overline{f} - (\partial_{v_i}\mathbf{b}) \cdot \nabla_v\overline{f} - (\partial_{v_i}c)\overline{f},$$

which solve (5.20) in $B_{r'} \times \mathbb{R}^d$ and satisfy

$$\|\overline{h}\|_{L^2(B_{r'};L^2_\gamma)} \leqslant \|\overline{f}\|_{L^2(B_{r'};H^1_\gamma)}$$

---

[12]One may verify this by writing out the commutator explicitly and using the fundamental theorem of calculus for the difference terms that arise, such as $c(x, v - v') - c(x, v)$ if the mollification variable is v'.

and
$$\|\overline{h}^*\|_{L^2(B_{r'};H_\gamma^{-1})} \leqslant C\left(\|\overline{f}^*\|_{L^2(B_{r'};L_\gamma^2)} + \|\partial_{x_i}\overline{f}\|_{L^2(B_{r'};L_\gamma^2)} + \|\overline{f}\|_{L^2(B_{r'};H_\gamma^1)}\right).$$

These, in turn, are estimated by the aforementioned inequalities for $\overline{f}$, $\widetilde{f}$, and f. Applying Caccioppoli's inequality and sending $\varepsilon \to 0^+$ and $\ell \to +\infty$ completes the proof. $\qquad\square$

Theorem 1.5 concerning the interior regularity, jointly in the variables x and v, is obtained by differentiating the equation and repeatedly applying Lemma 5.2 and Lemma 5.3, and we omit the details.

6. The kinetic Fokker-Planck equation

In this last section, we study the time-dependent kinetic Fokker-Planck equation

$$(6.1) \qquad \partial_t f - \varepsilon\left(\Delta_v f - v \cdot \nabla_v f\right) + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f = f^*.$$

The parameter ε is only relevant for the enhancement estimate, and one may imagine that $\varepsilon = 1$ until the final subsection. As with the Kramers equation, we prove a Poincaré inequality for bounded domains $V \subseteq \mathbb{R} \times \mathbb{R}^d$ which are either C^1 or cylindrical products $I \times U$ where $I \subseteq \mathbb{R}$ is a bounded interval and U is a bounded C^1 domain, but we consider the initial value problem only for $U = \mathbb{T}^d$.

6.1. **Function spaces.** We define the function space

$$(6.2) \qquad H_{\mathrm{kin}}^1(V) := \left\{f \in L^2(V; H_\gamma^1) : \partial_t f + v \cdot \nabla_x f \in L^2(V; H_\gamma^{-1})\right\},$$

equipped with the norm

$$(6.3) \qquad \|f\|_{H_{\mathrm{kin}}^1(V)} := \|f\|_{L^2(V;H_\gamma^1)} + \|\partial_t f + v \cdot \nabla_x f\|_{L^2(V;H_\gamma^{-1})}.$$

We denote the unit exterior normal to V by $\mathbf{n}_V \in L^\infty(\partial V; \mathbb{R}^{d+1})$. If V is a C^1 domain, then $\mathbf{n}_V(t,x)$ is well defined for every $(t,x) \in \partial V$; if V is of the form $I \times U$, then $\mathbf{n}_V(t,x)$ is well defined unless $(t,x) \in \partial I \times \partial U$, in which case we take the convention that $\mathbf{n}_V(t,x) = 0$. We define the hypoelliptic boundary of $V \subseteq \mathbb{R} \times \mathbb{R}^d$ as

$$\partial_{\mathrm{kin}}(V) := \left\{((t,x),v) \in \partial V \times \mathbb{R}^d : \begin{pmatrix} 1 \\ v \end{pmatrix} \cdot \mathbf{n}_V(t,x) < 0\right\},$$

We denote by $H_{\mathrm{kin},0}^1(V)$ the closure in $H_{\mathrm{kin}}^1(V)$ of the set of smooth functions which vanish on $\partial_{\mathrm{kin}}V$.

Proposition 6.1 (Density of smooth functions). *Let $V \subseteq \mathbb{R} \times \mathbb{R}^d$ be a bounded C^1 domain or cylindrical product $I \times U$, where U is a bounded C^1 domain. The set $C_c^\infty(\overline{V} \times \mathbb{R}^d)$ of smooth functions with compact support in $\overline{V} \times \mathbb{R}^d$ is dense in $H_{\mathrm{kin}}^1(V)$.*

Proof. Mimicking the first step of the proof of Proposition 2.2, which only uses that the domain is Lipschitz, we see that we can assume without loss of generality that for every $z \in V$ and $\varepsilon \in (0,1]$, we have

$$B((1-\varepsilon)z, \varepsilon) \subseteq V.$$

Here we use z to denote a generic variable in $\mathbb{R} \times \mathbb{R}^d$; in standard notation, $z = (t,x)$. Let ζ_ε be a $(1+d)$-dimensional version of the mollifier defined in (2.11), and let $f \in H_{\mathrm{kin}}^1(V)$. We define, for every $\varepsilon \in \left(0, \frac{1}{2}\right]$, $z \in V$ and $v \in \mathbb{R}^d$,

$$f_\varepsilon(z,v) := \int_{\mathbb{R}^{1+d}} f((1-\varepsilon)z + z', v)\zeta_\varepsilon(z')\,dz'.$$

We then show as in Step 2 of the proof of Proposition 2.2 that f belongs to the closed convex hull of the set $\left\{f_\varepsilon : \varepsilon \in \left(0, \frac{1}{2}\right]\right\}$, and then, as in Step 3 of this proof, that for each $\varepsilon > 0$, we have that f_ε belongs to the closure of the set $C_c^\infty(\overline{V} \times \mathbb{R}^d)$. $\qquad\square$

6.2. **Functional inequalities for H^1_{kin}.** We next show a Poincaré inequality for $H^1_{\mathrm{kin}}(V)$. For the sake of generality, we allow for more flexible boundary conditions than in Theorem 1.3, in the spirit of Remark 3.2.

Proposition 6.2 (Poincaré inequality). *Let $V \subseteq \mathbb{R} \times \mathbb{R}^d$ be a bounded C^1 domain or a cylindrical product $I \times U$ where U is a bounded C^1 domain.*

(1) There exists a constant $C(V, d) < \infty$ such that for every $f \in H^1_{\mathrm{kin}}(V)$, we have

$$\|f - (f)_V\|_{L^2(V;L^2_\gamma)} \leqslant C \left(\|\nabla_v f\|_{L^2(V;L^2_\gamma)} + \|v \cdot \nabla_x f + \partial_t f\|_{L^2(V;H^{-1}_\gamma)} \right).$$

(2) Let W be a relatively open subset of $\partial V \times \mathbb{R}^d$. There exists a constant $C(V, W, d) < \infty$ such that for every $f \in C^\infty_c(\overline{V} \times \mathbb{R}^d)$ that vanishes on W, we have

$$\|f\|_{L^2(V;L^2_\gamma)} \leqslant C \left(\|\nabla_v f\|_{L^2(V;L^2_\gamma)} + \|v \cdot \nabla_x f + \partial_t f\|_{L^2(V;H^{-1}_\gamma)} \right).$$

Proof of Proposition 6.2. The proof is similar to that of Theorem 1.3. By Proposition 6.1, we can assume that $f \in C^\infty_c(\overline{W} \times \mathbb{R}^d)$. We start by using the Gaussian Poincaré inequality to assert that

$$\|f - \langle f \rangle_\gamma\|_{L^2(V;L^2_\gamma)} \leqslant \|\nabla_v f\|_{L^2(V;L^2_\gamma)}.$$

Paralleling the second step of the proof of Theorem 1.3, we then aim to gain control on a negative Sobolev norm of the derivatives of $\langle f \rangle_\gamma$. Here we treat the time and space variables on an equal footing, and thus are interested in controlling $\partial_t \langle f \rangle_\gamma$ and $\nabla \langle f \rangle_\gamma$ in the $H^{-1}(V)$ norm. The precise claim is that there exists $C(d, V) < \infty$ such that for every test function $\phi \in C^\infty_c(V)$ satisfying

$$(6.4) \qquad \|\phi\|_{L^2(V)} + \|\nabla \phi\|_{L^2(V)} + \|\partial_t \phi\|_{L^2(V)} \leqslant 1,$$

we have

$$(6.5) \quad \left| \int_V \phi \, \partial_t \langle f \rangle_\gamma \right| + \sum_{i=1}^d \left| \int_V \phi \, \partial_{x_i} \langle f \rangle_\gamma \right| \leqslant C \left(\|\nabla_v f\|_{L^2(V;L^2_\gamma)} + \|v \cdot \nabla_x f + \partial_t f\|_{L^2(V;H^{-1}_\gamma)} \right).$$

We start by showing that the first term on the left side of (6.5), which refers to the time derivative of $\langle f \rangle_\gamma$, is estimated by the right side of (6.5). We select a smooth function $\xi_0 \in C^\infty_c(\mathbb{R}^d)$ such that

$$(6.6) \qquad \int_{\mathbb{R}^d} \xi_0(v) \, d\gamma(v) = 1 \quad \text{and} \quad \int_{\mathbb{R}^d} v \xi_0(v) \, d\gamma(v) = 0,$$

and observe that, using these properties of ξ_0, we can write

$$\int_V \partial_t \phi(t, x) \, \langle f \rangle_\gamma(t, x) \, dt \, dx$$

$$= \int_{V \times \mathbb{R}^d} \xi_0(v) \left(\partial_t \phi(t, x) + v \cdot \nabla_x \phi(t, x) \right) \langle f \rangle_\gamma(t, x) \, dt \, dx \, d\gamma(v)$$

$$= \int_{V \times \mathbb{R}^d} \xi_0(v) \left(\partial_t + v \cdot \nabla_x \right) \phi(t, x) \, f(t, x, v) \, dt \, dx \, d\gamma(v)$$

$$\qquad + \int_{V \times \mathbb{R}^d} \xi_0(v) \left(\partial_t + v \cdot \nabla_x \right) \phi(t, x) \left(\langle f \rangle_\gamma(t, x) - f(t, x, v) \right) dt \, dx \, d\gamma(v).$$

Using (6.4) and the fact that ξ_0 has compact support, we can bound the second integral above by

$$C \|f - \langle f \rangle_\gamma\|_{L^2(V;L^2_\gamma)} \leqslant C \|\nabla_v f\|_{L^2(V;L^2_\gamma)}.$$

By integration by parts, the absolute value of the first integral is equal to

$$\left| \int_{V \times \mathbb{R}^d} \xi_0(v) \phi(t, x) \left(v \cdot \nabla_x + \partial_t \right) f(t, x, v) \, dt \, dx \, d\gamma(v) \right| \leqslant C \|v \cdot \nabla_x f + \partial_t f\|_{L^2(V;H^{-1}_\gamma)}.$$

This completes the proof of the estimate in (6.5) involving the time derivative. To estimate the terms involving the space derivatives, we fix $i \in \{1, \ldots, d\}$ and use a smooth function $\xi_i \in C_c^\infty(\mathbb{R}^d)$ satisfying

$$\int_{\mathbb{R}^d} \xi_i(v) \, d\gamma(v) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} v\xi_i(v) \, d\gamma(v) = e_i$$

to get that

$$\int_V \partial_{x_i}\phi(t,x) \, \langle f \rangle_\gamma(t,x) \, dt \, dx$$
$$= \int_{V \times \mathbb{R}^d} \xi_i(v) \left(v \cdot \nabla_x \phi(t,x) + \partial_t \phi(t,x) \right) \langle f \rangle_\gamma(t,x) \, dt \, dx \, d\gamma(v).$$

The rest of the argument is then identical to the estimate involving the time derivative, and thus (6.5) is proved. The remainder of the proof is then identical to that for Theorem 1.3. Note that we need to invoke Lemma 3.1, which allows Lipschitz regularity, for the domain V. $\square$

6.3. The Hörmander inequality for H^1_{kin}.

For the Hörmander inequality, we recall the parameter ε from (6.1) and assume that the spatial/temporal domain is $V = [0, \varepsilon^{-1/3}] \times \mathbb{T}^d$, although a similar estimate would hold for $V = [0, \varepsilon^{-1/3}] \times \mathbb{R}^d$. We emphasize that we have included this particular factor of ε due to the fact that the a priori estimates for (6.1) control only $\varepsilon^{1/2}\nabla_v f$, and also due to the scaling between the regularity exponent we shall be able to obtain for $\nabla_x f$ and the a priori estimate. This inequality for $H^1_{\text{kin}}(V)$ is proved in an almost identical way to the one for $H^1_{\text{hyp}}(\mathbb{T}^d)$; the only difference is that the time variable is *not* periodic as is the space variable. So a bit of care must be taken with the finite differences corresponding to the vector field $\partial_t + v \cdot \nabla_x$. We track the parameter ε throughout the proof for the purposes of the enhancement estimate later on. The version of (3.25) we use here is

(6.7)
$$f(t, x + \eta^3 \varepsilon^{1/2} x', v) - f(t,x,v)$$
$$= f(t, x + \eta^3 \varepsilon^{1/2} x', v) - f(t, x + \eta^3 \varepsilon^{1/2} x', v - \eta\varepsilon^{1/2} x')$$
$$+ f(t, x + \eta^3 \varepsilon^{1/2} x', v - \eta\varepsilon^{1/2} x') - f(t + \eta^2, x + \eta^3 \varepsilon^{1/2} x' + \eta^2(v - \varepsilon^{1/2}\eta x'), v - \eta\varepsilon^{1/2} x')$$
$$+ f(t + \eta^2, x + \eta^2 v, v - \eta\varepsilon^{1/2} x') - f(t + \eta^2, x + \eta^2 v, v)$$
$$+ f(t + \eta^2, x + \eta^2 v, v) - f(t,x,v).$$

As before, we must define the following Besov spaces based on finite differences in the ∇_x and $D_t = \partial_t + v \cdot \nabla_x$ directions. The Besov space measuring fractional regularity in the x variable now depends *fundamentally* on ε and t, and so we denote this space $Q^{1/3,\varepsilon}_{\nabla_x}$. To lighten the notation, in the context of proofs in which ε is always fixed, we sometimes shall substitute the notation $Q^{1/3}_{\nabla_x}$ instead of the more cumbersome $Q^{1/3,\varepsilon}_{\nabla_x}$, and similarly for $Q^{1/2,\varepsilon}_{D_t}$.

Definition 6.3. For measurable $u : (0, \varepsilon^{-1/3}) \times \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$, we define

$$\|u\|^2_{Q^{1/2,\varepsilon}_{D_t}} := \sup_{0 < \eta \leqslant \sqrt{\frac{\varepsilon^{-1/3}}{2}}} \frac{1}{\eta^2} \Bigg(\iiint_{\left(0, \frac{\varepsilon^{-1/3}}{2}\right) \times \mathbb{R}^d \times \mathbb{T}^d} \left(u(t + \eta^2, x + \eta^2 v, v) - u(t,x,v) \right)^2 dx \, d\gamma(v) \, dt$$

(6.8)
$$+ \iiint_{\left(\frac{\varepsilon^{-1/3}}{2}, \varepsilon^{-1/3}\right) \times \mathbb{R}^d \times \mathbb{T}^d} \left(u(t - \eta^2, x - \eta^2 v, v) - u(t,x,v) \right)^2 dx \, d\gamma(v) \, dt \Bigg).$$

We define
(6.9)
$$\|u\|^2_{Q^{1/3,\varepsilon}_{\nabla_x}} := \sup_{\substack{0<\eta\leqslant\sqrt{\frac{\varepsilon^{-1/3}}{2}} \\ x'\in\mathbb{S}^{d-1}}} \frac{1}{\eta^2} \iiint_{(0,\varepsilon^{-1/3})\times\mathbb{R}^d\times\mathbb{T}^d} \left(u(t,x+\varepsilon^{1/2}\eta^3 x',v)-u(t,x,v)\right)^2 dx\,d\gamma(v)\,dt\,.$$

Notice that the quantity $\varepsilon^{1/2}\eta^3$ is of order one if η^2 takes its maximum value of $\frac{\varepsilon^{-1/3}}{2}$. Then by iterating the finite differences, the norm in (6.9) is equivalent to one in which the supremum is taken over values of η at least as large as the diameter of $\mathbb{T}^d$, at which point the norm is equivalent to one including all positive values of η.

To streamline the proof of the enhancement estimate later, we assume in the following proposition that $\langle\partial_t u+v\cdot\nabla_x u\rangle_\gamma\equiv 0$ (a condition which will be satisfied in the enhancement context). Then from Lemma 2.1, the $L^2_{t,x}H^{-1}_\gamma$ norm of $\partial_t u+v\cdot\nabla_x u$ may be obtained via duality against the gradients (in v) of $L^2_{t,x}H^1_\gamma$ functions which have vanishing means $\langle\cdot\rangle_\gamma$. Thus the inequality (6.10) does not require the $L^2_{t,x}L^2_\gamma$ norm of u on the right-hand side; one could easily adjust the statement in the case that $\langle\partial_t u+v\cdot\nabla_x u\rangle_\gamma\neq 0$ by including the necessary term.

Lemma 6.4 (Interpolation). *For every $\delta>0$, there exists a constant $C(\delta,d)<\infty$ (not depending on ε) such that for any smooth function u satisfying $\langle\partial_t u+v\cdot\nabla_x u\rangle_\gamma\equiv 0$,*

(6.10)
$$\|u\|^2_{Q^{1/2,\varepsilon}_{D_t}} \leqslant \delta\|u\|^2_{Q^{1/3,\varepsilon}_{\nabla_x}}$$
$$+ C(\delta)\left(\varepsilon\|\nabla_v u\|^2_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)}+\varepsilon^{-1}\|\partial_t u+v\cdot\nabla_x u\|^2_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;H^{-1}_\gamma)}\right).$$

Remark 6.5. The factors of ε ensure that the right-hand side remains of order 1 as $\varepsilon\to 0$ and arise naturally when deriving the a priori estimates for solutions to (6.1); see section 6.6 for more details.

Proof. The proof is similar for both halves of (6.8), i.e. the forward and backward differences, and so we focus on the case of the forward difference.

Step 1. Let $\phi\in C^\infty_0((-1,1)^d)$ be a smooth, positive, radial function with unit L^1 norm. For $\zeta>0$, we define $\phi_\zeta u(t,x,v)$ by

$$\phi_\zeta u(t,x,v) = \int_{\mathbb{R}^d} u(t,x+\zeta^3\varepsilon^{1/2}x',v)\phi(x')\,dx'\,.$$

Analogously to Step 1 from the proof of Theorem 1.4, we have that

(6.11)
$$\left\|\phi_\zeta u(t,x,v)-u(t,x,v)\right\|^2_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)} \leqslant \zeta^2\|u\|^2_{Q^{1/3}_{\nabla_x}}\,.$$

Step 2. Let

$$f(\eta) = \left\|u(t+\eta^2,x+\eta^2 v,v)-u(t,x,v)\right\|^2_{L^2\left((0,\frac{\varepsilon^{-1/3}}{2})\times\mathbb{T}^d;L^2_\gamma\right)}\,.$$

We may write that

$$f(\eta) \lesssim \left\|\phi_{\delta\eta}u(t+\eta^2,x+\eta^2 v,v)-u(t+\eta^2,x+\eta^2 v,v)\right\|^2_{L^2\left((0,\frac{\varepsilon^{-1/3}}{2})\times\mathbb{T}^d;L^2_\gamma\right)}$$

$$+ \left\|\phi_{\delta\eta}u(t+\eta^2,x+\eta^2 v,v)-\phi_{\delta\eta}u(t,x,v)\right\|^2_{L^2\left((0,\frac{\varepsilon^{-1/3}}{2})\times\mathbb{T}^d;L^2_\gamma\right)}$$

(6.12)
$$+ \left\|\phi_{\delta\eta}u(t,x,v)-u(t,x,v)\right\|^2_{L^2\left((0,\frac{\varepsilon^{-1/3}}{2})\times\mathbb{T}^d;L^2_\gamma\right)}\,,$$

where the implicit constant is independent of η, δ, and u. By Step 1 with $\zeta = \delta\eta$, the first and third terms are bounded by

$$\delta^2\eta^2 \left\|u\right\|^2_{Q^{1/3}_{\nabla_x}} .$$

Step 3. It remains to estimate the second term in (6.12). For $\eta \in \left(0, \sqrt{\frac{\varepsilon^{-1/3}}{2}}\right)$ and $0 \leqslant \tau \leqslant \eta^2$, consider

$$(6.13) \qquad F(\tau) = \left\|\phi_{\delta\eta}u(t + \tau, x + \tau v, v) - \phi_{\delta\eta}u(t, x, v)\right\|^2_{L^2\left((0,\frac{\varepsilon^{-1/3}}{2})\times\mathbb{T}^d;L^2_\gamma\right)} .$$

The term in question is $F(\eta^2)$. Since $F(0) = 0$, it suffices to estimate $F'(\tau)$. We have that

$$F'(\tau) = 2 \iiint_{\left(0,\frac{\varepsilon^{-1/3}}{2}\right)\times\mathbb{R}^d\times\mathbb{T}^d} \left(\phi_{\delta\eta}u(t + \tau, x + \tau v, v) - \phi_{\delta\eta}u(t, x, v)\right)$$

$$\cdot D_t\left(\phi_{\delta\eta}u\right)(t + \tau, x + \tau v, v)\, dx\, d\gamma(v)\, dt$$

$$= 2 \iiint_{\left(\tau,\frac{\varepsilon^{-1/3}}{2}+\tau\right)\times\mathbb{R}^d\times\mathbb{T}^d} \left(\phi_{\delta\eta}u(t, x, v) - \phi_{\delta\eta}u(t - \tau, x - \tau v, v)\right)$$

$$(6.14) \qquad\qquad\qquad \cdot D_t\left(\phi_{\delta\eta}u\right)(t, x, v)\, dx\, d\gamma(v)\, dt .$$

From $[D_t, \phi_{\delta\eta}]u = [\nabla_v, \phi_{\delta\eta}]u = 0$, the assumption $\langle\partial_t u + v \cdot \nabla_x u\rangle_\gamma \equiv 0$, and our control of

$$\left\|\partial_t u + v \cdot \nabla_x u\right\|_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;H^{-1}_\gamma)} ,$$

we will achieve the desired estimate for $F'(\tau)$ if we can bound

$$\nabla_v\left(\phi_{\delta\eta}u(t, x, v) - \phi_{\delta\eta}u(t - \tau, x - \tau v, v)\right)$$

in $L^2((\tau, \frac{\varepsilon^{-1/3}}{2} + \tau) \times \mathbb{T}^d; L^2_\gamma)$. Notice that after obtaining these bounds, we apply the Cauchy-Schwarz inequality with a prefactor of ε in front of one term and ε^{-1} in front of the other in order to obtain (6.10). The only non-trivial estimate comes when the ∇_v lands on the x coordinate of the second term, which we may write out as

$$\int_{\mathbb{T}^d} - \tau\nabla_x u\left(t - \tau, x + (\delta\eta)^3\varepsilon^{1/2}x' - \tau v, v\right)\phi(x')\, dx'$$

$$= -\int_{\mathbb{T}^d} \frac{\tau}{(\delta\eta)^3\varepsilon^{1/2}}\nabla_{x'}u(t - \tau, x + (\delta\eta)^3\varepsilon^{1/2}x' - \tau v, v)\phi(x')\, dx'$$

$$= \int_{\mathbb{T}^d} \frac{\tau}{(\delta\eta)^3\varepsilon^{1/2}}u(t - \tau, x + (\delta\eta)^3\varepsilon^{1/2}x' - \tau v, v)\nabla_{x'}\phi(x')\, dx'$$

$$= \int_{\mathbb{T}^d} \frac{\tau}{(\delta\eta)^3\varepsilon^{1/2}}\left(u(t - \tau, x + (\delta\eta)^3\varepsilon^{1/2}x' - \tau v, v) - u(t - \tau, x - \tau v, v)\right)\nabla_{x'}\phi(x')\, dx' .$$

But slight adjustments to the argument from Step 1 show that this is bounded in $L^2\left((\tau, \frac{\varepsilon^{1/2}}{2} + \tau) \times \mathbb{T}^d; L^2_\gamma\right)$ by a constant independent of δ times

$$\frac{\tau}{(\delta\eta)^3\varepsilon^{1/2}}\delta\eta \left\|u\right\|_{Q^{1/3}_{\nabla_x}} \leqslant \frac{1}{\delta^2\varepsilon^{1/2}}\left\|u\right\|_{Q^{1/3}_{\nabla_x}} ,$$

where here we have used the assumption that $\tau \leqslant \eta^2$. Using the Cauchy-Schwarz and Young inequalities to absorb the negative powers of ε and δ with the $L^2_{t,x}H^{-1}_\gamma$ norm concludes the proof. $\qquad\square$

We may now state and prove the following proposition. As with the interpolation, in the case that $\langle\partial_t u + v \cdot \nabla_x u\rangle_\gamma \neq 0$, one could adjust the statement of the second inequality to include the necessary $L^2_{t,x}L^2_\gamma$ norm of u.

Proposition 6.6 (Hörmander inequality). *There exists $C(d) < \infty$ (not depending on ε) such that for every smooth function u satisfying $\langle \partial_t u + v \cdot \nabla_x u \rangle_\gamma \equiv 0$, we have*

$$\|u\|_{Q_{\nabla_x}^{1/3,\varepsilon}} \leqslant C\left(\varepsilon^{1/2}\|\nabla_v u\|_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)} + \|u\|_{Q_{D_t}^{1/2,\varepsilon}}\right)$$

$$(6.15) \qquad \leqslant C\left(\varepsilon^{1/2}\|\nabla_v u\|_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)} + \varepsilon^{-1/2}\|\partial_t u + v\cdot\nabla_x u\|_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;H^{-1}_\gamma)}\right).$$

Proof of Proposition 6.6. Set $g(t,x,v) = f(t,x,v)\gamma^{1/2}(v)$, and choose $\eta^2 \in (0,\varepsilon^{-1/3}]$ and $x' \in \mathbb{S}^{d-1}$. Then we may write that

$$\left\|f(t,x+\varepsilon^{1/2}\eta^3 x',v) - f(t,x,v)\right\|_{L^2\left((0,\frac{\varepsilon^{-1/3}}{2})\times\mathbb{T}^d;L^2_\gamma\right)}$$

$$= \left\|g(t,x+\varepsilon^{1/2}\eta^3 x',v) - g(t,x,v)\right\|_{L^2\left((0,\frac{\varepsilon^{-1/3}}{2})\times\mathbb{T}^d;L^2(\mathbb{R}^d)\right)},$$

and

$$(6.16) \qquad g(t,x+\eta^3\varepsilon^{1/2}x',v) - g(t,x,v)$$

$$= g(t,x+\eta^3\varepsilon^{1/2}x',v) - g(t,x+\eta^3\varepsilon^{1/2}x',v-\eta\varepsilon^{1/2}x')$$

$$+ g(t,x+\eta^3\varepsilon^{1/2}x',v-\eta\varepsilon^{1/2}x')$$

$$- g(t+\eta^2,x+\eta^3\varepsilon^{1/2}x'+\eta^2(v-\varepsilon^{1/2}\eta x'),v-\eta\varepsilon^{1/2}x')$$

$$+ g(t+\eta^2,x+\eta^2 v,v-\eta\varepsilon^{1/2}x') - g(t+\eta^2,x+\eta^2 v,v)$$

$$+ g(t+\eta^2,x+\eta^2 v,v) - g(t,x,v).$$

Dividing by η, integrating in $L^2\left((0,\frac{\varepsilon^{-1/3}}{2})\times\mathbb{T}^d;L^2(\mathbb{R}^d)\right)$, and appealing to (2.7) as in the time-independent case yields that

$$\frac{1}{\eta}\left\|f(t,x+\varepsilon^{1/2}\eta^3 x',v) - f(t,x,v)\right\|_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)}$$

$$\lesssim \varepsilon^{1/2}\|\nabla_v f\|_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)} + \|f\|_{Q_{D_t}^{1/2}}.$$

For the other half of the time interval, it is easy to rewrite (6.16) with a backwards difference in the $\partial_t + v \cdot \nabla_x$ direction by first adding $\eta\varepsilon^{1/2}x'$ in the v variable and then subtracting η^2 in the t variable and $\eta^2(v+\varepsilon^{1/2}\eta x')$ in the x variable. Arguing as for the forward differences produces an identical estimate. Then using Lemma 6.4 and absorbing the $\|f\|^2_{Q_{\nabla_x}^{1/3}}$ factor required to bound $\|f\|_{Q_{D_t}^{1/2}}$ from the right-hand side onto the left-hand side gives the result. $\qquad\square$

Remark 6.7. From the embedding $Q_{\nabla_x}^{1/3} \hookrightarrow L^2\left((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma\right)$ for functions with vanishing x-mean $\langle u\rangle(t,v) = \int_{\mathbb{T}^d} u(t,x,v)\,dx$ (see, for example, [2]), we obtain the following ε-dependent Poincaré inequality:

$$(6.17) \qquad\qquad \|u\|_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)} \leqslant C\varepsilon^{-1/6}\|u\|_{Q_{\nabla_x}^{1/3}}.$$

Note that to obtain this inequality, we have rescaled out the factors of ε used in the finite differences of the $Q_{\nabla_x}^{1/3}$ norm and then appealed to an ε-independent function space embedding.

Remark 6.8 (Regularity in time). By an interpolation argument, the result of Proposition 6.6 implies some time regularity for a function $f \in H^1_{\mathrm{kin}}(V)$ for $V = (0,\varepsilon^{-1/3})\times\mathbb{T}^d$. Indeed, by the definition of the norm $\|\cdot\|_{H^1_{\mathrm{kin}}}$, we have that

$$\|f\|_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;H^1_\gamma)} \leqslant \|f\|_{H^1_{\mathrm{kin}}((0,\varepsilon^{-1/3})\times\mathbb{T}^d)}.$$

By interpolation and (6.15), for every $\theta \in [0,1]$ and $\alpha \in \left[0, \frac{1}{3}\right)$,

$$\|f\|_{L^2\left((0,\varepsilon^{-1/3});H^{\theta\alpha}(\mathbb{T}^d;H_\gamma^{1-2\theta})\right)} \leqslant C \|f\|_{H^1_{\mathrm{kin}}\left((0,\varepsilon^{-1/3})\times\mathbb{T}^d\right)} .$$

We also have, by (6.15), for any $\alpha \in \left[0, \frac{1}{3}\right)$,

$$\begin{aligned}
\|f\|_{H^1\left((0,\varepsilon^{-1/3});H^{\alpha-1}(\mathbb{T}^d;H_\gamma^{-1})\right)} & \\
&\leqslant \|f\|_{L^2\left((0,\varepsilon^{-1/3});H^{\alpha-1}(\mathbb{T}^d;H_\gamma^{-1})\right)} + \|\partial_t f\|_{L^2\left((0,\varepsilon^{-1/3});H^{\alpha-1}(\mathbb{T}^d;H_\gamma^{-1})\right)} \\
&\leqslant \|f\|_{L^2\left((0,\varepsilon^{-1/3});L^2(\mathbb{T}^d;H_\gamma^{-1})\right)} + \|\partial_t f - v\cdot\nabla_x f\|_{L^2\left((0,\varepsilon^{-1/3});L^2(\mathbb{T}^d;H_\gamma^{-1})\right)} \\
&\qquad + \|v\cdot\nabla_x f\|_{L^2\left((0,\varepsilon^{-1/3});H^{\alpha-1}(\mathbb{T}^d;H_\gamma^{-1})\right)} \\
&\leqslant C \|f\|_{H^1_{\mathrm{kin}}\left((0,\varepsilon^{-1/3})\times\mathbb{T}^d\right)} .
\end{aligned}$$

By interpolation of the previous two displays, we obtain, for any $\theta, \sigma \in [0,1]$ and $\alpha \in \left[0, \frac{1}{3}\right)$,

$$(6.18) \qquad \|f\|_{H^\sigma\left((0,\varepsilon^{-1/3});H^{\theta\alpha-\sigma(1-\alpha+\theta\alpha)}(\mathbb{T}^d;H_\gamma^{1-2(\theta+\sigma-\theta\sigma)})\right)} \leqslant C \|f\|_{H^1_{\mathrm{kin}}\left((0,\varepsilon^{-1/3})\times\mathbb{T}^d\right)} .$$

Each of the constants C above depends only on (α, d). Note that all three exponents can be made simultaneously positive, for example taking $\alpha = \theta = \frac{1}{4}$ and $\sigma = \frac{1}{32}$ yields

$$(6.19) \qquad \|f\|_{H^{1/32}\left((0,\varepsilon^{-1/3});H^{1/32}(\mathbb{T}^d;H_\gamma^{7/16})\right)} \leqslant C \|f\|_{H^1_{\mathrm{kin}}\left((0,\varepsilon^{-1/3})\times\mathbb{T}^d\right)} .$$

By (6.19) and an argument very similar to the proof of Proposition 3.8, which we omit, we obtain the following compact embedding statement.

Proposition 6.9 (Compact embedding of H^1_{kin} into L^2). *For any bounded C^1 domain $V \subseteq \mathbb{R} \times \mathbb{R}^d$ or cylindrical product $I \times U$ where U is a bounded C^1 domain, the inclusion map $H^1_{\mathrm{kin}}(V) \hookrightarrow L^2(V; L^2_\gamma)$ is compact.*

6.4. Well-posedness of the Cauchy problem.

Proposition 6.10 (Solvability of the kinetic Fokker-Planck equation). *Let $T \in (0, +\infty]$, $f_{\mathrm{in}} \in L^2_m$, and $g^* \in L^2(\mathbb{T}^d \times (0,T); H_\gamma^{-1})$. Under Assumption 1.1, there exists a unique solution*

$$(6.20) \qquad f \in C([0,T]; L^2_m(\mathbb{T}^d \times \mathbb{R}^d)) \cap H^1_{\mathrm{kin}}((0,T) \times \mathbb{T}^d)$$

to the kinetic Fokker-Planck equation (6.1) with initial data f_{in} and forcing term g^*.

Proof. Let $T \in (0, +\infty]$. Let $f_{\mathrm{in}} \in L^2_m$ and $g^* \in L^2\left((0,T); L^2_\sigma(\mathbb{T}^d; H_\gamma^{-1})\right)$. A function g solves the kinetic Fokker-Planck equation if and only if $f(t,x,v) = g(t,x,v)e^t$ solves

$$(6.21) \qquad \partial_t f + (v\cdot\nabla_x + \mathbf{b}\cdot\nabla_v)f + f = f^* + \varepsilon(\Delta f - v\cdot\nabla_v f) ,$$

where $f^* = e^t g^*$. We solve (6.21) on $(0,T) \times \mathbb{T}^d \times \mathbb{R}^d$ by applying Lemma 4.1 with an appropriate functional setup:

(1) the *test function space*

$$(6.22) \qquad \Phi = C_0^\infty(\mathbb{T}^d \times \mathbb{R}^d \times [0,T))$$

with inner product

$$(6.23) \qquad (\phi, \psi) = \int_0^T \int_{\mathbb{T}^d\times\mathbb{R}^d} \nabla_v\phi\cdot\nabla_v\psi \, dm \, dt + \int_0^T \int_{\mathbb{T}^d\times\mathbb{R}^d} \phi\psi \, dm \, dt ,$$

(2) the *solution space*

$$H = L^2(0,T; L^2_\sigma(\mathbb{T}^d; H_\gamma^1))$$

with inner product (6.23),

(3) the *bilinear form*

$$E(h, \phi) = \varepsilon \int_0^T \int_{\mathbb{T}^d \times \mathbb{R}^d} \nabla_v h \cdot \nabla_v \phi \, dm \, dt + \int_0^T \int_{\mathbb{T}^d \times \mathbb{R}^d} h\phi \, dm \, dt$$
$$- \int_0^T \int_{\mathbb{T}^d \times \mathbb{R}^d} h(\partial_t + v \cdot \nabla_x + \mathbf{b} \cdot \nabla_v)\phi \, dm \, dt \,,$$

(4) and the *linear functional*

$$L\phi = \int_{\mathbb{T}^d \times \mathbb{R}^d} f_{\mathrm{in}}\phi(x, v, 0) \, dm + g^*(\phi) \,.$$

As before, in the Kramers equation, one may verify that E is continuous (4.2) on H for each fixed $\phi \in \Phi$. We now verify coercivity (4.3) and mention two essential new features: (i) the initial data f_{in} is built into the linear function L, and (ii) test functions $\phi \in \Phi$ vanish at $t = T$ but are not required to vanish at $t = 0$ (which is necessary for them to 'detect' the initial data). After integrating by parts in all variables, we have

$$E(\phi, \phi) = \varepsilon \int_0^T \int_{\mathbb{T}^d \times \mathbb{R}^d} |\nabla_v \phi|^2 \, dm \, dt + \int_0^T \int_{\mathbb{T}^d \times \mathbb{R}^d} |\phi|^2 \, dm \, dt + \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} |\phi(x, v, 0)|^2 \, dm$$
$$\geqslant \varepsilon(\psi, \psi)_H \,.$$

Lemma 4.1 generates a weak solution $f \in H$ to $E(f, \phi) = L\phi$ for all $\phi \in \Phi$. In particular, choosing $\phi \in \Phi$ that additionally vanish near $t = 0$ guarantees that the PDE (6.21) is satisfied in the sense of distributions. From the PDE itself, we recover that $f \in H^1_{\mathrm{kin}}(\mathbb{T}^d \times (0, T))$ and, in particular, $f \in C([0, T]; L^2_\sigma(\mathbb{T}^d; L^2_\gamma))$; see Lemma 6.12. This is enough regularity to justify that the initial data is f_{in} and the basic energy estimate which guarantees uniqueness. $\qquad\square$

We do not include a proof of the following statement in this paper, since the argument is a close adaptation of the one of Theorem 1.5. We denote $V_r := (-r, r) \times B_r$ and by $\nabla_{t,x}$ the full gradient in t and x, that is, $\nabla_{t,x} = (\partial_t, \nabla_x)$.

Proposition 6.11 (Interior regularity, kinetic Fokker-Planck)**.** *Let* $k \in \mathbb{N}$, $r \in (0, \infty)$, *and* $\mathbf{b} \in C^{k-1,1}(V_r \times \mathbb{R}^d; \mathbb{R}^d)$. *There exists a constant* $C < \infty$ *depending on*

$$\left(d, k, r, \|\mathbf{b}\|_{C^{k-1,1}(V_r \times \mathbb{R}^d; \mathbb{R}^d)}\right)$$

such that, for every $f \in H^1_{\mathrm{kin}}(V_r)$ *and* $f^* \in L^2(V_r; H^{-1}_\gamma)$ *satisfying*

(6.24) $$\partial_t f - \Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f = f^* \quad \text{in } V_r \times \mathbb{R}^d \,,$$

the following holds: If $\partial^\alpha f^* \in L^2(B_r; H^{-1}_\gamma)$ *for all multi-indices* $\alpha \in \mathbb{N} \times \mathbb{N}^d \times \mathbb{N}^d$ *satisfying* $|\alpha| \leqslant k$, *then we have* $\partial^\alpha f \in H^1_{\mathrm{kin}}(V_{r/2})$ *and the estimate*

$$\|\partial^\alpha f\|_{H^1_{\mathrm{kin}}(V_{r/2})} \leqslant C\left(\|f - (f)_{V_r}\|_{L^2(V_r; L^2_\gamma)} + \sum_{|\beta| \leqslant k} \|\partial^\beta \widetilde{f}^*\|_{L^2(V_r; H^{-1}_\gamma)}\right).$$

for all multi-indices $\alpha \in \mathbb{N} \times \mathbb{N}^d \times \mathbb{N}^d$ *satisfying* $|\alpha| \leqslant k$.

6.5. **Exponential decay in time.** For each bounded interval $I = (I_-, I_+) \subseteq \mathbb{R}$ and bounded C^1 domain U, we denote by $H^1_{\mathrm{kin}, \|}(I \times U)$ the closure in $H^1_{\mathrm{kin}}(I \times U)$ of the set of smooth functions which vanish on $I \times \partial_{\mathrm{hyp}} U$. Note that in particular, we allow the trace of $f \in H^1_{\mathrm{kin}, \|}(I \times U)$ on the initial time slice $\{I_-\} \times U$ to be non-zero. In this section, we show that a solution to the kinetic Fokker-Planck equation with zero right-hand side and belonging to $H^1_{\mathrm{kin}, \|}(I \times U)$ decays to zero exponentially fast in time. We start with a preliminary classical lemma.

Lemma 6.12 (continuity in L^2). *Every function in $H^1_{\mathrm{kin},\|}(I \times U)$ can be identified (up to a set of null measure) with an element of $C(\overline{I}; L^2(U; L^2_\gamma))$.*

Proof. If f is a smooth function which vanishes on $I \times \partial_{\mathrm{hyp}} U$, then for every $t \in I$, we have

$$\partial_t \|f(t, \cdot)\|^2_{L^2(U; L^2_\gamma)} + \int_{\partial U \times \mathbb{R}^d} f^2(t, x, v)(v \cdot \mathbf{n}_U(x))_+ \, dx \, d\gamma(v)$$
$$= 2 \int_{U \times \mathbb{R}^d} (f(\partial_t f + v \cdot \nabla_x f))(t, x, v) \, dx \, d\gamma(v),$$

where we recall that $(r)_+ := \max(0, r)$. Since the second integral on the left side is nonnegative, we deduce that for every $s, t \in I$,

$$\left| \|f(t, \cdot)\|^2_{L^2(U; L^2_\gamma)} - \|f(s, \cdot)\|^2_{L^2(U; L^2_\gamma)} \right| \leqslant 2 \|f\|_{L^2((s,t) \times U; H^1_\gamma)} \|\partial_t f + v \cdot \nabla_x f\|_{L^2((s,t) \times U; H^{-1}_\gamma)},$$

and thus, for a constant $C(I) < \infty$,

$$\sup_{t \in \overline{I}} \|f(t, \cdot)\|_{L^2(U; L^2_\gamma)} \leqslant C \|f\|_{H^1_{\mathrm{kin}}(I \times U)}.$$

For a general $f \in H^1_{\mathrm{kin},\|}(I \times U)$, there exists a sequence (f_n) of smooth functions which vanish on $I \times \partial_{\mathrm{hyp}} U$ and such that f_n converges to f in $H^1_{\mathrm{kin}}(I \times U)$. It follows from the inequality above that f_n converges to f with respect to the $L^\infty(I; L^2(U; L^2_\gamma))$ norm; in particular, $f \in C(\overline{I}; L^2(U; L^2_\gamma))$. $\qquad\square$

We finally turn to the proof of Theorem 1.6, which is restated in the following proposition. Notice that, by linearity, it suffices to prove the theorem in the case $f^* = 0$ and $f_\infty = 0$.

Proposition 6.13 (Exponential decay to equilibrium). *Let $U \subseteq \mathbb{R}^d$ be a bounded C^1 domain and $\mathbf{b} \in L^\infty(U \times \mathbb{R}^d)^d$. There exists $\lambda(\|\mathbf{b}\|_{L^\infty(U \times \mathbb{R}^d)}, U, d) > 0$ such that, for every $T \in (0, \infty)$ and $f \in H^1_{\mathrm{kin},\|}((0,T) \times U)$ satisfying*

$$\partial_t f - \Delta_v f + v \cdot \nabla_v f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f = 0 \qquad \text{in } (0, T) \times U \times \mathbb{R}^d,$$

we have, for every $t \in (0, T)$,

$$\|f(t, \cdot)\|_{L^2(U; L^2_\gamma)} \leqslant 2 \exp(-\lambda t) \|f(0, \cdot)\|_{L^2(U; L^2_\gamma)}.$$

Proof. For every $0 \leqslant s < t$, we compute

$$\frac{1}{2} \left(\|f(t, \cdot)\|^2_{L^2(U; L^2_\gamma)} - \|f(s, \cdot)\|^2_{L^2(U; L^2_\gamma)} \right) \leqslant - \|\nabla_v f\|^2_{L^2((s,t) \times U; L^2_\gamma)}.$$

In particular,

(6.25)   the mapping $t \mapsto \|f(t, \cdot)\|_{L^2(U; L^2_\gamma)}$ is nonincreasing.

Since

$$-\nabla_v^* \nabla_v f = \partial_t f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f,$$

we have

$$\|\partial_t f + v \cdot \nabla_x f\|_{L^2((s,t) \times U; H^{-1}_\gamma)}$$
$$\leqslant \|\partial_t f + v \cdot \nabla_x f + \mathbf{b} \cdot \nabla_v f\|_{L^2((s,t) \times U; H^{-1}_\gamma)} + \|\mathbf{b} \cdot \nabla_v f\|_{L^2((s,t) \times U; H^{-1}_\gamma)}$$
$$\leqslant C \|\nabla_v f\|_{L^2((s,t) \times U; L^2_\gamma)},$$

and thus

(6.26)
$$-\left(\|f(t,\cdot)\|^2_{L^2(U;L^2_\gamma)} - \|f(s,\cdot)\|^2_{L^2(U;L^2_\gamma)} \right)$$
$$\geqslant \frac{1}{C} \left(\|\nabla_v f\|^2_{L^2((s,t)\times U;L^2_\gamma)} + \|\partial_t f + v \cdot \nabla_x f\|^2_{L^2((s,t)\times U;H^{-1}_\gamma)} \right).$$

We aim to appeal to Proposition 6.2 to conclude. We define

(6.27)
$$V := [0,1] \times U.$$

For every $t \geqslant 0$, we write

$$V_t := (t,0) + V = \{(t+s,x) \in \mathbb{R} \times \mathbb{R}^d \ : \ (s,x) \in V\}.$$

Inequality (6.26) implies that, for every $t \geqslant 0$,

$$-\left(\|f(t+1,\cdot)\|^2_{L^2(U;L^2_\gamma)} - \|f(t,\cdot)\|^2_{L^2(U;L^2_\gamma)} \right)$$
$$\geqslant \frac{1}{C} \left(\|\nabla_v f\|^2_{L^2(V_t;L^2_\gamma)} + \|\partial_t f - v \cdot \nabla_x f\|^2_{L^2(V_t;H^{-1}_\gamma)} \right).$$

Proposition 6.2 yields that

$$-\left(\|f(t+1,\cdot)\|^2_{L^2(U;L^2_\gamma)} - \|f(t,\cdot)\|^2_{L^2(U;L^2_\gamma)} \right) \geqslant \frac{1}{C} \|f\|^2_{L^2(V_t;L^2_\gamma)}.$$

Using (6.25) and (6.27), we deduce that

$$-\left(\|f(t+1,\cdot)\|^2_{L^2(U;L^2_\gamma)} - \|f(t,\cdot)\|^2_{L^2(U;L^2_\gamma)} \right) \geqslant \frac{1}{C} \|f(t+1,\cdot)\|^2_{L^2(U;L^2_\gamma)}.$$

This implies exponential decay of the mapping $t \mapsto \|f(t,\cdot)\|_{L^2(U;L^2_\gamma)}$ along integer values of t, and we then obtain the conclusion of the proposition by using (6.25) once more. $\qquad\square$

6.6. **Enhancement.** Finally, we prove Theorem 1.7. Recall that f is assumed to be a solution to

(6.28)
$$\partial_t f + v \cdot \nabla_x f = \varepsilon \left(\Delta_v f - v \cdot \nabla_v f \right) \text{ in } (0,\infty) \times \mathbb{T}^d \times \mathbb{R}^d.$$

Proof of Theorem 1.7. After multiplying (6.28) by f and integrating over $(0,\varepsilon^{-1/3}) \times \mathbb{T}^d \times \mathbb{R}^d$, we obtain the a priori estimates

$$\varepsilon \|\nabla_v f\|^2_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d\times\mathbb{R}^d)} \leqslant \|f_{\text{in}}\|^2_{L^2(\mathbb{T}^d;L^2_\gamma)} - \left\| f(\varepsilon^{-1/3},\cdot,\cdot) \right\|^2_{L^2(\mathbb{T}^d;L^2_\gamma)}$$

$$\varepsilon^{-1} \|\partial_t f + v \cdot \nabla_x f\|^2_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d(H^{-1}_\gamma))} \lesssim \|f_{\text{in}}\|^2_{L^2(\mathbb{T}^d;L^2_\gamma)} - \left\| f(\varepsilon^{-1/3},\cdot,\cdot) \right\|^2_{L^2(\mathbb{T}^d;L^2_\gamma)}.$$

Applying the inequality in (6.15) from Proposition 6.6, which is justified since $\langle \partial_t f + v \cdot \nabla_x f \rangle_\gamma = \varepsilon \langle \Delta_v f - v \cdot \nabla_v f \rangle_\gamma \equiv 0$, we obtain that

$$\|f\|^2_{Q^{1/3}_{\nabla_x}} \lesssim \varepsilon \|\nabla_v f\|^2_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)} \lesssim \|f_{\text{in}}\|^2_{L^2(\mathbb{T}^d;L^2_\gamma)} - \left\| f(\varepsilon^{-1/3},\cdot,\cdot) \right\|^2_{L^2(\mathbb{T}^d;L^2_\gamma)}.$$

From (6.17) and the observation that the mean-zero in x condition from (1.26) is propagated forward in time, we then obtain that

$$\|f\|^2_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)} \lesssim \varepsilon^{-1/3} \|f\|^2_{Q^{1/3}_{\nabla_x}}$$
$$\lesssim \varepsilon^{2/3} \|\nabla_v f\|^2_{L^2((0,\varepsilon^{-1/3})\times\mathbb{T}^d;L^2_\gamma)}$$
$$\lesssim \varepsilon^{-1/3} \left(\|f_{\text{in}}\|^2_{L^2(\mathbb{T}^d;L^2_\gamma)} - \left\| f(\varepsilon^{-1/3},\cdot,\cdot) \right\|^2_{L^2(\mathbb{T}^d;L^2_\gamma)} \right).$$

Translating in time and iterating this procedure yields exponential decay with rate $\exp(-c\varepsilon^{-1/3}t)$ along integer multiples of $\varepsilon^{-1/3}$, similarly to the proof of Proposition 6.13. Applying (6.25), which holds as well for solutions to (6.28), we obtain (1.27). $\qquad\square$

Remark 6.14. In principle, one can also incorporate a conservative $\mathbf{b}$ satisfying Assumption 1.1 into the enhancement estimate, since $[\mathbf{b}(x) \cdot \nabla_v, \partial_{v_i}] = 0$ for all $i = 1, \cdots, d$.

REFERENCES

[1] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[2] D. Albritton, R. Beekie, and M. Novack. Enhanced dissipation and Hörmander's hypoellipticity, 2021. arxiv.2105.12308.

[3] F. Anceschi and A. Rebucci. A note on the weak regularity theory for degenerate kolmogorov equations, 2021.

[4] S. Armstrong, A. Bordas, and J.-C. Mourrat. Quantitative stochastic homogenization and regularity theory of parabolic equations. *Anal. PDE*, 11(8):1945–2014, 2018.

[5] S. Armstrong and J.-C. Mourrat. Variational methods for the kinetic Fokker-Planck equation. *arXiv preprint arXiv:1902.04037, v1*, 2019.

[6] M. S. Baouendi and P. Grisvard. Sur une équation d'évolution changeant de type. *J. Functional Analysis*, 2:352–367, 1968.

[7] F. Baudoin. Bakry-Émery meet Villani. *J. Funct. Anal.*, 273(7):2275–2291, 2017.

[8] M. Beck and C. E. Wayne. Metastability and rapid convergence to quasi-stationary bar states for the two-dimensional Navier-Stokes equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 143(5):905–927, 2013.

[9] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith. A regularity method for lower bounds on the Lyapunov exponent for stochastic differential equations. *Inventiones mathematicae*, Sept. 2021.

[10] J. Bedrossian and M. Coti Zelati. Enhanced dissipation, hypoellipticity, and anomalous small noise inviscid limits in shear flows. *Arch. Ration. Mech. Anal.*, 224(3):1161–1204, 2017.

[11] 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. *To appear in Probab. Math. Phys.*, 2021.

[12] M. E. Bogovskiĭ. Solutions of some problems of vector analysis, associated with the operators div and grad. In *Theory of cubature formulas and the application of functional analysis to problems of mathematical physics*, volume 1980 of *Trudy Sem. S. L. Soboleva, No. 1*, pages 5–40, 149. Akad. Nauk SSSR Sibirsk. Otdel., Inst. Mat., Novosibirsk, 1980.

[13] F. Bouchut. Hypoelliptic regularity in kinetic equations. *J. Math. Pures Appl. (9)*, 81(11):1135–1159, 2002.

[14] H. Brézis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17):Aii, A971–A974, 1976.

[15] H. Brézis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(20):Ai, A1197–A1198, 1976.

[16] G. Brigati. Time averages for kinetic Fokker-Planck equations. *arXiv preprint arXiv:2106.12801*, 2021.

[17] Y. Cao, J. Lu, and L. Wang. On explicit l^2-convergence rate estimate for underdamped Langevin dynamics. *arXiv preprint arXiv:1908.04746*, 2019.

[18] J. A. Carrillo. Global weak solutions for the initial-boundary-value problems to the Vlasov-Poisson-Fokker-Planck system. *Math. Methods Appl. Sci.*, 21(10):907–938, 1998.

[19] L. Desvillettes and C. Villani. On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: the linear Fokker-Planck equation. *Comm. Pure Appl. Math.*, 54(1):1–42, 2001.

[20] L. Desvillettes and C. Villani. On the trend to global equilibrium for spatially inhomogeneous kinetic systems: the Boltzmann equation. *Invent. Math.*, 159(2):245–316, 2005.

[21] J.-P. Eckmann and M. Hairer. Spectral properties of hypoelliptic operators. *Comm. Math. Phys.*, 235(2):233–253, 2003.

[22] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1976.

[23] G. P. Galdi. *An introduction to the mathematical theory of the Navier-Stokes equations*. Springer Monographs in Mathematics. Springer, New York, second edition, 2011. Steady-state problems.

[24] I. Gallagher, T. Gallay, and F. Nier. Spectral asymptotics for large skew-symmetric perturbations of the harmonic oscillator. *Int. Math. Res. Not. IMRN*, (12):2147–2199, 2009.

[25] N. Ghoussoub. *Self-dual partial differential systems and their variational principles*. Springer Monographs in Mathematics. Springer, New York, 2009.

[26] 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. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2019.

[27] J. Guerand and C. Imbert. Log-transform and the weak harnack inequality for kinetic fokker-planck equations, 2021.

[28] Y. Guo. The Landau equation in a periodic box. *Comm. Math. Phys.*, 231(3):391–434, 2002.

[29] B. Helffer and F. Nier. *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, volume 1862 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2005.

[30] F. Hérau and F. Nier. Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.*, 171(2):151–218, 2004.

[31] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[32] L. Hörmander. *The analysis of linear partial differential operators. III*. Classics in Mathematics. Springer, Berlin, 2007. Pseudo-differential operators, Reprint of the 1994 edition.

[33] J. J. Kohn. Pseudo-differential operators and hypoellipticity. In *Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971)*, pages 61–69, 1973.

[34] A. Kolmogoroff. Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math. (2)*, 35(1):116–117, 1934.

[35] J. Lu and L. Wang. On explicit l^2-convergence rate estimate for piecewise deterministic Markov processes. *arXiv preprint arXiv:2007.14927*, 2020.

[36] A. Lunardi. *Interpolation theory*, volume 16 of *Appunti. Scuola Normale Superiore di Pisa*. Edizioni della Normale, Pisa, 2018.

[37] G. Mingione. Calderón-Zygmund estimates for measure data problems. *C. R. Math. Acad. Sci. Paris*, 344(7):437–442, 2007.

[38] G. Mingione. Gradient potential estimates. *J. Eur. Math. Soc. (JEMS)*, 13(2):459–486, 2011.

[39] C. Mouhot. De Giorgi-Nash-Moser and Hörmander theories: new interplay. *Proc. ICM*, 2018.

[40] G. Papanicolaou and S. R. S. Varadhan. Ornstein-Uhlenbeck process in a random potential. *Comm. Pure Appl. Math.*, 38(6):819–834, 1985.

[41] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1991.

[42] R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.

[43] D. Talay. Approximation of invariant measures of nonlinear Hamiltonian and dissipative stochastic differential equations. In *Progress in stochastic structural dynamics*, volume 152, pages 139–169. Publ. LMA-CNRS, 1999.

[44] D. Talay. Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields*, 8(2):163–198, 2002.

[45] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950):iv+141, 2009.

[46] W. Wang and L. Zhang. The C^α regularity of a class of non-homogeneous ultraparabolic equations. *Sci. China Ser. A*, 52(8):1589–1606, 2009.

[47] W. Wang and L. Zhang. The C^α regularity of weak solutions of ultraparabolic equations. *Discrete Contin. Dyn. Syst.*, 29(3):1261–1275, 2011.

(D. Albritton) Courant Institute of Mathematical Sciences, New York University, 251 Mercer St., New York, NY 10012

Current address: School of Mathematics, Institute for Advanced Study, 1 Einstein Dr., Princeton, NJ 08540, USA

E-mail address: `dallas.albritton@ias.edu`

(S. Armstrong) Courant Institute of Mathematical Sciences, New York University, 251 Mercer St., New York, NY 10012

E-mail address: `scotta@cims.nyu.edu`

(J.-C. Mourrat) ENS Lyon, CNRS, 46 allée d'Italie, 69007 Lyon, France; Courant Institute of Mathematical Sciences, New York University, 251 Mercer St., New York, NY 10012

E-mail address: `jean-christophe.mourrat@ens-lyon.fr`

(M. Novack) Courant Institute of Mathematical Sciences, New York University, 251 Mercer St., New York, NY 10012

Current address: School of Mathematics, Institute for Advanced Study, 1 Einstein Dr., Princeton, NJ 08540, USA

E-mail address: `mdn@ias.edu`