ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

MATEMATICA

Ciclo XXXVII

Settore Concorsuale: 01/A3 ANALISI MATEMATICA, PROBABILITÀ E STATISTICA MATEMATICA
Settore Scientifico Disciplinare: MAT/06 PROBABILITÀ E STATISTICA MATEMATICA

ON SOME REGULARITY PROBLEMS FOR KINETIC-TYPE EQUATIONS

Presentata da: *Giacomo Lucertini*

Coordinatore Dottorato

Chiar.mo Prof. Giovanni Mongardi

Supervisore

Chiar.mo Prof. Andrea Pascucci

Co-supervisore:

Prof. Stefano Pagliarani

Esame finale anno 2024

Abstract

In this thesis we study some regularity problems for kinetic-type partial differential equations. These equations are characterized by the fact that their second order part is fully degenerate, but the presence of a first order operator restores good properties for the solution. In the first part of the thesis we consider a class of Backward Kolmogorov equations with rough coefficients, namely coefficients that are measurable in time and Hölder continuous in space. We prove optimal regularity for the fundamental solution and Schauder estimates for the Cauchy problem. In the second part we study boundary regularity for a kinetic Fokker-Planck equation with constant coefficients. We also prove a Nash inequality for kinetic Sobolev spaces.

This thesis is based on the results presented in [71], [65] and [42].

Contents

Introduction and motivations

In the last decades the Theory of Partial Differential Equations has seen an increasing interest in *kinetic-type equations*. These equations are mathematically characterized by two parts: the first is elliptic on a strict sub-vector space of the spacial domain only, and therefore the equations are fully degenerate, the second is a first order vector field, often referred to as *transport term*, that allows to derive good regularity and asymptotic properties of the solutions. This structural assumptions make the treatment on this subject challenging from a mathematical point of view and one can confront many different problems. In fact, these equations can be easily compared to classical parabolic equations: for this reason, kinetic-type equations are sometimes referred as *ultraparabolic equations* (see [60]). This analogy suggests that results holding for parabolic equations should have an ultraparabolic counterpart: in recent literature we find many results along these lines, together with open problems. Concurrently, the interest for this class of equations is also motivated by applications in different fields, from the more practical to the more theoretical ones. Kinetic-type equations naturally arise in modeling many natural phenomena as well as in studying pure mathematical results. We now give a brief overview of some of the most common applications.

Examples of applications and literature

In physics, kinetic-type equations arise in models from multiple brunches of the subject, especially in evolution problems with interactions. Classical examples are Boltzmann-type equations. Some classical monographs are [14], [15], [28], while a more recent survey on the topic can be found in [85].

In statistical mechanics we find models with a similar structure for anharmonic chains of oscillators forced by heat baths at different temperatures; we cite [29] as a reference for interested reader. In general, kinetic-type equations appear in stochastic Hamiltonian systems or in some geophysics models (see, for example, [83] and [19]).

In mathematical finance, kinetic-type equations are used for modeling different objects (see [75] for an introduction to the subject). Applying the classical Black-Scholes theory, we can derive the equation for the value of an Asian option, that is a kinetic-type Kolmogorov equation (see, for example, [7], or, for a very recent survey, [62]). A similar class of equations can be obtained in the study of volatility models with path-dependent coefficients as the Hobson-Rogers model (see [45] or [22]).

From the point of view of stochastic analysis, these equations are strongly related to a class of degenerate *stochastic differential equations* (SDEs). In particular, the *characteristic operator* (or *generator*) of a kinetic SDEs is a kinetic-type operator (see [34] for a general introduction). These types of stochastic equations were initially studied by Kolmogorov in [56], starting from the Langevin model (see also Hörmander [46]). The theory around these equations has been developed to the present day: for example, strong uniqueness for a general class of kinetic SDEs can be found in [18] or [41] (see also [4] for a recent review on kinetic Kolmogorov equations).

Overview of this work

The main subject of this thesis will be the study of regularity for kinetic-type equations. We will focus on two kind of kinetic-type equations in order to give precise estimates on the solution in different settings: in one case, we consider equations with rough coefficients, in the other one, we investigate boundary regularity for a kinetic constant coefficient equation.

In the first part of the thesis we consider a class of Backward Kolmogorov equations with non-constant coefficients that are measurable in time and Hölder continuous in space: together with the degeneracy of the equation, the challenge here is to deal with the low regularity of the coefficients. There is a long stream of research that investigate Kolmogorov equations in these settings: two of the main branches were introduced by Da Prato and Lunardi in [20] and Lanconelli and Polidoro in [60] (see also [79]). In the present study, we focus our attention on equations posed on a strip $[0, T] \times \mathbb{R}^N$, therefore on a limited time domain and in the whole space.

In particular, in Chapter 1 we fix the settings for the results in Chapters 2 and 3. We define a general class of Backward Kolmogorov operators with its structural assumptions.

Then we define an anisotropic norm that captures the spatial multiscale nature of the equation. The crucial definitions of this chapter are those of the functional spaces that will be used in the two chapters that follow: anisotropic Hölder spaces describe the regularity of the coefficients of the equation, while intrinsic Hölder spaces describe the optimal regularity of the solution of the equation. Understanding these spaces is important to appreciate the results of Chapters 2 and 3 and to understand the proofs. Chapter 2 is devoted to the construction of the fundamental solution for the same class of Kolmogorov operators and to study its optimal regularity (in term of intrinsic Hölder spaces). We use a time-dependent parametrix technique that is computationally heavy method, but strong in the results. Exploiting these results, in Chapter 3 we solve the associated Cauchy problem. In particular, through Schauder estimates for the equation we give optimal regularity of the solution. Via a Duhamel principle we represent the solution of the Cauchy problem using the fundamental solution obtained in Chapter 2; using the regularity results that we have proved we are able to obtain the desired estimates.

In the second part of the thesis, we study a kinetic Fokker-Planck equation with constant coefficients. In this setting we are interested in facing the regularity problem on a domain with (nonempty) boundary. From a physical point of view, the solution of this equation is the density of particles in a point of the phase space at a certain time: the dynamics depends on the initial distribution and the boundary condition assumed, together with the equation. In this framework, interior regularity have been widely studied, leaving open problems on the regularity of the solution close to the boundary. This is the direction taken in Chapter 4, where we study a kinetic homogeneous Fokker-Planck equation on a half space domain (i.e. $\mathbb{R}_+ \times \mathbb{H}^d \times \mathbb{R}^d$, where $\mathbb{H}^d = \big\{(x_1, \dots, x_d) \in \mathbb{R}^d : x_1 > 0\big\}$). Precisely, we prove a Nash-type inequality for functions in a kinetic Sobolev space with absorbing boundary condition and use it to prove sharp regularity for the solution of the kinetic Fokker-Planck equation.

Future developments

We end this introduction pointing at future developments that might follow these results. As it was already mentioned, Kolmogorov equations have a huge importance in the study of SDEs. Results from the first part of the thesis can be applied to prove the uniqueness for solutions of kinetic stochastic equations: it is well known that, in non-degenerate settings, good properties for the solution of the Kolmogorov equation allow to prove the uniqueness for the solution of the SDE (see [31] for an introduction to the subject). This explains a phenomenon that is often called *regularization by noise*: if we consider a ordinary differential equation with non-Lipschitz continuous coefficients (for example, Hölder continuous), we do not have uniqueness for the solution, but, as soon as we add a non-degenerate stochastic noise, the obtained SDE has a path-by-path unique solution. In recent years, this phenomenon was also studied in kinetic settings (see, for example, [18]), but open problems are left.

In the same spirit, another possible development is to extend the technique in order to face different type of stochastic kinetic equations. For example, in the last decades there was a growing interest for McKean-Vlasov (MKV) equations: in this case, the associated Backward Kolmogorov operator also present a measure-argument derivative, making the study even more challenging (see, for example, [16] and the references within). Recently, these results have been investigated for kinetic MKV equations (see, for some recent results, [55] and [40]). Our technique can be extended in order to prove results in this direction. Another type of stochastic kinetic equations that can be studied following our approach are equations driven by a Lévy alpha-stable process (instead of classical Brownian motion), where the generators are non-local operators. Also in this case, the study at the PDE level is useful as much as it is challenging (for a recent example, see [41]).

Finally, we refer to the introduction of Chapter 4 for extension and generalization of results on the boundary regularity for the kinetic Fokker-Planck equation. We just add here that in the literature for kinetic-type equations we found different definition of kinetic Sobolev spaces (see, for example, [2] and [78]). We believe that it would be interesting to find precise inclusion relations between these different spaces: this study would allow to better understand the nature of these spaces and to extend some functional inequalities.

Acknowledgments

I would like to thank my supervisors Andrea Pascucci and Stefano Pagliarani for their guidance and their always being ready to help.

I would also like to thank Christopher Henderson for inviting me at the University of Arizona and treating me as if I were one of his students, and Weinan Wang for patiently collaborating during my exchange period.

I would like to thank all the professors and colleagues from the Department of Mathematics on the University of Bologna that have shared with me some of their mathematical experience with me.

Finally, I would like to thank all my family and friends for supporting me in these three years.

Chapter 1

A class of degenerate Kolmogorov operator and induced intrinsic geometry

In this chapter we set the class of Kolmogorov operators that will be studied in Chapters 2 and 3. Together with that, we introduce the intrinsic geometry that these operators naturally induce on $\mathbb{R}^{N+1}$. Based on these notions, we define the functional spaces that will be used throughout Chapters 2 and 3.

1.1 Class of operators

For fixed $d \le N$ and $T > 0$, we consider the second order operator on $\mathcal{S}_T := [0, T] \times \mathbb{R}^N$ in non-divergence form

$$\mathcal{L} := \mathcal{A} + Y \tag{1.1}$$

where

$$\begin{aligned} \mathcal{A} &= \frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(t,x)\partial_{x_i x_j} + \sum_{i=1}^{d} a_i(t,x)\partial_{x_i} + a(t,x), \qquad (t,x) \in \mathcal{S}_T \\ Y &= \partial_t + \langle Bx, \nabla \rangle = \partial_t + \sum_{i,j=1}^{N} b_{ij} x_j \partial_{x_i}, \qquad x \in \mathbb{R}^N, \end{aligned} \tag{1.2}$$

where B is a constant matrix of dimension $N \times N$. Here, $\mathcal{A}$ is an elliptic operator on $\mathbb{R}^d$ and Y is a first order differential operator on $\mathbb{R} \times \mathbb{R}^N$, also called *transport* or *drift term*. We mainly focus on the case $d < N$, that is when $\mathcal{L}$ is fully degenerate, namely no coercivity condition on $\mathbb{R}^N$ is satisfied.

Motivations for the study of $\mathcal{L}$ come from physics and finance. In its most basic form, with $N = 2$ and $d = 1$,

$$\frac{1}{2}\partial_{x_1x_1} + x_1\partial_{x_2} + \partial_t \tag{1.3}$$

is the backward Kolmogorov operator of the system of stochastic equations

$$\begin{cases} dV_t = dW_t \\ dX_t = V_t dt \end{cases} \tag{1.4}$$

where W is a real Brownian motion. In the classical Langevin model, (V, X) describes the velocity and position of a particle in the phase space and is the prototype of more general kinetic models (cf. [59], [52], [53]). In mathematical finance, (V, X) represents the log-price and average processes used in modeling path-dependent financial derivatives, such as Asian options (cf. [7], [73]). The study of the fundamental solution and its regularity properties is a crucial step in tackling the martingale problem for the corresponding stochastic equations, particularly for well-posedness and pathwise uniqueness problems.

Throughout Chapters 1, 2, and 3, $\mathcal{L}$ verifies the following two structural

Assumption 1.1.1 (Coercivity on $\mathbb{R}^d$). The diffusion matrix $A := (a_{ij})_{i,j=1,\dots,d}$ is symmetric and there exists a positive constant μ such that

$$\mu^{-1}|\xi|^2 \le \sum_{i,j=1}^{d} a_{ij}(t,x)\xi_i\xi_j \le \mu|\xi|^2, \qquad x \in \mathbb{R}^N,\ \xi \in \mathbb{R}^d,$$

for almost every $t \in [0, T]$.

Assumption 1.1.2 (Hörmander condition). The vector fields $\partial_{x_1}, \dots, \partial_{x_d}$ and Y satisfy

$$\text{rank Lie}(\partial_{x_1}, \dots, \partial_{x_d}, Y) = N + 1. \tag{1.5}$$

We refer to (1.5) as a *parabolic* Hörmander condition since the drift term Y plays a key role in the generation of the Lie algebra.

Under Assumption 1.1.2, the prototype Kolmogorov operator

$$\frac{\delta}{2}\sum_{i=1}^{d} \partial_{x_i x_i} + Y \tag{1.6}$$

is hypoelliptic for any $\delta > 0$. Kolmogorov [56] (see also [46]) constructed the explicit Gaussian fundamental solution for (1.6), which is the transition density of the linear stochastic differential equation in $\mathbb{R}^N$

$$dX_t = BX_t dt + \Sigma dW_t,$$

where Σ is a $N \times d$ matrix with the block form

$$\Sigma = \begin{pmatrix} I_d \\ 0_{(N-d)\times d} \end{pmatrix}$$

and $(W_t)_{t\geq 0}$ is a d-dimensional Brownian motion.

Condition (1.5) is equivalent to the well-known Kalman rank condition for controllability in linear systems theory (cf., for instance, [75]). Also, it was shown in [60] that (1.5) is equivalent to B having the block-form

$$B = \begin{pmatrix} * & * & \cdots & * & * \\ B_1 & * & \cdots & * & * \\ 0 & B_2 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_q & * \end{pmatrix} \tag{1.7}$$

where the $*$-blocks are arbitrary and B_j is a $(d_{j-1} \times d_j)$-matrix of rank d_j with

$$d \equiv d_0 \geq d_1 \geq \cdots \geq d_q \geq 1, \qquad \sum_{i=0}^{q} d_i = N.$$

This allows to introduce natural definitions of anisotropic (quasi-)norm on $\mathbb{R}^N$.

Definition 1.1.3 (**Anisotropic norm).** For any $x \in \mathbb{R}^N$ let

$$|x|_B := \sum_{j=0}^{q} \sum_{i=\bar{d}_{j-1}+1}^{\bar{d}_j} |x_i|^{\frac{1}{2j+1}}, \qquad \bar{d}_j := \sum_{k=0}^{j} d_k. \tag{1.8}$$

We also define Q the so-called *homogeneous dimension* of $\mathbb{R}^N$ with respect to the quasi-norm $|\cdot|_B$ as

$$Q = \sum_{i=0}^{q} (2i+1) d_i. \tag{1.9}$$

The quasi-norm (1.8) can be directly related to the scaling properties of the underlying diffusion process (cf. [21], [57]). For example, the anisotropic norm for the Langevin operator (1.3) reads as $|(v,x)|_B = |v| + |x|^{\frac{1}{3}}$ for $(v,x) \in \mathbb{R}^2$ and reflects the time-scaling properties of the stochastic system (1.4), i.e. $(\Delta V)^2 \approx \Delta t$ and $(\Delta X)^2 \approx (\Delta t)^3$.

For convenience, we also fix here the following

Notation 1.1.4. Let $\kappa = (\kappa_1, \dots, \kappa_N) \in \mathbb{N}_0^N$ be a multi-index, we define the *B-length* of κ as

$$[\kappa]_B := \sum_{j=0}^{q} (2j+1) \sum_{i=\bar{d}_{j-1}+1}^{\bar{d}_j} \kappa_i.$$

1.2 Hölder spaces

In this section, we introduce the *anisotropic* and *intrinsic* Hölder spaces that appear in the main results of Chapter 2 and 3. Loosely speaking, the general idea behind the definition of these spaces is the following:

- anisotropic Hölder spaces are defined for functions of $x \in \mathbb{R}^N$, *assuming regularity in all N directions* w.r.t. an anisotropic distance that reflects the different time-scaling properties of the underlying diffusion process. This distance is defined in term of an anisotropic norm that assigns to each component of $x \in \mathbb{R}^N$ a different weight corresponding to the number of commutators of ∇_d and Y that are required to reach that direction. The definition then extends to functions defined on $\mathbb{R}^{N+1}$ by only requiring measurability and local boundedness with respect to the time variable.

- intrinsic Hölder spaces are defined for functions of $(t,x) \in \mathbb{R}^{N+1}$ that are assumed to be anisotropically Hölder continuous, in the sense above, uniformly in time. Additional Hölder regularity in the direction of the drift Y is also assumed. By means of the Hörmander condition, it is then possible to infer Hölder regularity jointly with respect to all variables. As it is standard in the framework of functional analysis on homogeneous groups (cf. [32]), the idea is to weight the Hölder exponent in terms of the formal degree of the vector fields, which is equal to 1 for $\partial_{x_1}, \dots, \partial_{x_d}$ and equal to 2 for Y.

We begin by fixing some general notation that will be utilized throughout Chapters 1, 2 and 3.

Notation 1.2.1. Let $g : \mathbb{R}^N \to \mathbb{R}$. For any $i = 1, \dots, N$, we denote by $\partial_i g(x)$ the partial derivatives of g with respect to x_i. We also denote by ∇_d the gradient operator $(\partial_1, \dots, \partial_d)$ with respect to the first d components and by ∇_d^2 the Hessian operator $(\partial_{ij})_{i,j=1,\dots,d}$ with respect to the first d components. We also consider the natural extensions of the operators ∇_d and ∇_d^2 to functions $f = f(t,x)$ defined on $\mathbb{R} \times \mathbb{R}^N$. We set the following basic functional spaces:

- C_b, the set of bounded continuous functions $g : \mathbb{R}^N \to \mathbb{R}$, equipped with the norm

$$\|g\|_{L^\infty} := \sup_{x\in\mathbb{R}^N} |g(x)|;$$

- L_t^∞, for $t > 0$, the set of measurable functions f, defined on the strip $\mathcal{S}_t$, such that the norm

$$\|f\|_{L_t^\infty} := \sup_{(s,x)\in\mathcal{S}_t} |f(s,x)|$$

  is finite.

Moreover, all the normed spaces in this section are defined for scalar valued functions and naturally extend to vector valued functions by considering the sum of the norms of their single components.

Definition 1.2.2 (**Anisotropic Hölder spaces).** Let $\alpha \in]0,3]$.

- The anisotropic Hölder norms on $\mathbb{R}^N$ are defined recursively as

$$\|g\|_{\mathbf{C}^\alpha} := \begin{cases} \|g\|_{L^\infty} + \sup\limits_{x,y\in\mathbb{R}^N} \frac{|g(x)-g(y)|}{|x-y|_B^\alpha}, & \alpha \in]0,1], \\ \|g\|_{L^\infty} + \|\nabla_d g\|_{\mathbf{C}^{\alpha-1}} + \sup\limits_{(x,\eta)\in\mathbb{R}^N\times\mathbb{R}^{N-d}} \frac{|g(x+(0,\eta))-g(x)|}{|(0,\eta)|_B^\alpha}, & \alpha \in]1,3]. \end{cases}$$

  We denote by $\mathbf{C}^\alpha$ the set of functions $g : \mathbb{R}^N \to \mathbb{R}$ such that the norm $\|g\|_{\mathbf{C}^\alpha}$ is finite. Set also $\mathbf{C}^0 := C_b$.

- For $t > 0$, the anisotropic Hölder norms on $\mathcal{S}_t$ are

$$\text{(weighted)} \quad \|f\|_{L^\infty_{t,\gamma}(\mathbf{C}^\alpha)} := \sup_{s\in\,]0,t[} \Big((t-s)^\gamma \|f(s,\cdot)\|_{\mathbf{C}^\alpha} \Big), \quad \gamma \in [0,1[,$$

$$\text{(non-weighted)} \quad \|f\|_{L^\infty_{t}(\mathbf{C}^\alpha)} := \|f\|_{L^\infty_{t,0}(\mathbf{C}^\alpha)}.$$

  We denote by $L^\infty_{t,\gamma}(\mathbf{C}^\alpha)$ and $L^\infty_{t}(\mathbf{C}^\alpha)$ the set of measurable functions $f : \mathcal{S}_t \to \mathbb{R}$ such that the norms $\|f\|_{L^\infty_{t,\gamma}(\mathbf{C}^\alpha)}$ and $\|f\|_{L^\infty_{t}(\mathbf{C}^\alpha)}$, respectively, are finite.

Before introducing the intrinsic Hölder spaces, we recall the following

Definition 1.2.3 (Lie Hölder spaces and Lie derivative). For $t > 0$ and $\alpha \in]0,2]$ we set

$$\|f\|_{C^\alpha_{Y,t}} := \sup_{(\tau,x)\in\mathcal{S}_t} \sup_{s\in]0,t[} \frac{\left|f(s,e^{(s-\tau)B}x) - f(\tau,x)\right|}{|s-\tau|^{\frac{\alpha}{2}}}.$$

We denote by $C^\alpha_{Y,t}$ the set of measurable functions $f : \mathcal{S}_t \to \mathbb{R}$ such that the norms $\|f\|_{C^\alpha_{Y,t}}$ is finite.

Moreover, we say that f *is a.e. Lie-differentiable along Y on $\mathcal{S}_t$* if there exists $F \in L^1_{\text{loc}}(]0,t[; C_b(\mathbb{R}^N))$ such that

$$f(s,e^{(s-\tau)B}x) = f(\tau,x) + \int_\tau^s F(r,e^{(r-\tau)B}x)dr, \qquad (\tau,x)\in\mathcal{S}_t,\ s\in]0,t[.$$

In that case, we set $Yf = F$ and call it an *a.e. Lie derivative of f on $\mathcal{S}_t$.*

Definition 1.2.4 (Intrinsic Hölder spaces). Let $t > 0$. The *intrinsic Hölder norms on $\mathcal{S}_t$* are defined recursively as:

$$\|f\|_{\mathbf{C}^\alpha_t} := \begin{cases} \|f\|_{L^\infty_t(\mathbf{C}^\alpha)} + \|f\|_{C^\alpha_{Y,t}}, & \alpha\in]0,1], \\ \|f\|_{L^\infty_t(\mathbf{C}^\alpha)} + \|\nabla_d f\|_{\mathbf{C}^{\alpha-1}_t} + \|f\|_{C^\alpha_{Y,t}}, & \alpha\in]1,2], \\ \|f\|_{L^\infty_t(\mathbf{C}^\alpha)} + \|\nabla_d f\|_{\mathbf{C}^{\alpha-1}_t} + \|Yf\|_{L^\infty_t(\mathbf{C}^{\alpha-2})}, & \alpha\in]2,3], \end{cases}$$

For $\alpha \in]0,3]$, we denote by $\mathbf{C}^\alpha_t$ the set of functions $f : \mathcal{S}_t \to \mathbb{R}$ such that the norm $\|f\|_{\mathbf{C}^\alpha_t}$ is finite.

We now give a few remarks concerning the latter definition.

Remark 1.2.5. Notice that the recursive Definition 1.2.4 mixes scalar and vector valued functions: we recall that also here in the recursion the norm of a vector valued function is the sum of the norms of their single components.

Remark 1.2.6 (Intrinsic vs anisotropic spaces). Obviously, the intrinsic space $\mathbf{C}^\alpha_t$ is strictly included in the anisotropic space $L^\infty_t(\mathbf{C}^\alpha)$. Note that, for $\alpha \in]0,1]$, the addition in the $\mathbf{C}^\alpha_t$-norm of the term $\|f\|_{C^\alpha_{Y,t}}$ yields Hölder regularity jointly in the time and space variables: in particular, it is standard to show that if $f \in \mathbf{C}^\alpha_t$ then

$$|f(\tau,x) - f(s,y)| \le C\|f\|_{\mathbf{C}^\alpha_t}\left(|\tau-s|^{\frac{\alpha}{2}} + |x - e^{(\tau-s)}y|^\alpha_B\right), \qquad (\tau,x),(s,y)\in\mathcal{S}_t.$$

Remark 1.2.7 (Intrinsic Taylor formula). For $\alpha \in]0,1]$, the intrinsic spaces $\mathbf{C}_t^\alpha$ and $\mathbf{C}_t^{1+\alpha}$ are equivalent to those in [72] (see also [73]). However, $\mathbf{C}_t^{2+\alpha}$ is slightly weaker than the one in [72] in that the Lie derivative Yf is not required to be in $\mathbf{C}_t^\alpha$ but only in $L_t^\infty(\mathbf{C}^\alpha)$. This difference is dictated by our assumptions on the coefficients that are merely measurable in the temporal variable: so, for a solution u of (3.1), one may expect that Yu exists in the strong sense but, in general, is not more than bounded in the Y-direction. Despite this, $\mathbf{C}_t^{2+\alpha}$ in Definition 1.2.4 is still strong enough to prove the following *intrinsic Taylor formula* as in [72, Theorem 2.10]: if $f \in \mathbf{C}_t^{2+\alpha}$ then

$$|f(\tau,x)-T_2 f\big(s,y;x-e^{(\tau-s)B}y\big)| \lesssim \|f\|_{\mathbf{C}_t^{2+\alpha}}\big(|\tau-s|+|x-e^{(\tau-s)B}y|_B^{2+\alpha}\big), \qquad (\tau,x),(s,y)\in\mathcal{S}_t, \tag{1.10}$$

where $T_2 f$ is the second order intrinsic Taylor polynomial

$$T_2 f(s,y;z) = f(s,y) + \sum_{i=1}^d z_i\,\partial_i f(s,y) + \frac{1}{2}\sum_{i,j=1}^d z_i z_j\,\partial_{ij} f(s,y).$$

Furthermore, by adding the term $(\tau-s)Yu(s,y)$ to T_2, the estimate can be improved by obtaining a term of order $o(|\tau-s|)$, as $\tau-s\to 0$, in place of $|\tau-s|$ in (1.10).

It is worth noting that, for f in the anisotropic space $L_t^\infty(\mathbf{C}^\alpha)$, *estimate* (1.10) *generally holds only for* $s=\tau$: this is the best result that one can deduce from Schauder estimates of anisotropic-to-anisotropic type.

Remark 1.2.8. To have a quick comparison with the literature on the regularity for Kolmogorov operators in (1.2), we recall that, for $\alpha \in]0,1]$:

i) the Hölder space $C^{2+\alpha}$ introduced in [66] (and adopted in [64], [81] to prove Schauder estimates), consists of functions f that, together with their second order derivatives $\partial_{x_i x_j} f$ in the non-degenerate directions $i,j=1,\dots,d$, are Hölder continuous w.r.t. the anisotropic norm (1.8). This notion is weaker than Definition 1.2.4 both in terms of the regularity of $\partial_{x_i} f$ and, more importantly, in terms of the Lipschitz continuity of f along Y (cf. (3.5)) which reveals the regularizing effect of the associated evolution semigroup;

ii) Definition 1.2.4 is similar in spirit to that proposed in [67], [24] and [72] for the study of Kolmogorov operators with Hölder coefficients: according to their definition if $f \in C^{2,\alpha}$ then Yf exists and belongs $\mathbf{C}_t^\alpha$. This is the regularity that the fundamental solution enjoys in case the coefficients of $\mathcal{A}$ are Hölder continuous in both space and time. By contrast, if $f \in \mathbf{C}_t^{2+\alpha}$ in the sense of Definition 1.2.4 then f is generally at most Lipschitz continuous along Y: this is the optimal result one can prove without assuming further regularity of the coefficients in the time variable other than measurability.

Remark 1.2.9. In [71], intrinsic Hölder spaces $C_B^{k,\alpha}(\mathcal{S}_T)$, for $\alpha \in]0,1]$ and $h = 0,1,2$, are recursively defined through the norms

$$\begin{aligned}
\|f\|_{C_B^{0,\alpha}(\mathcal{S}_T)} &:= \|f\|_{C^\alpha_{Y,T}} + \|f\|_{C^\alpha_{\nabla_d,T}}, \\
\|f\|_{C_B^{1,\alpha}(\mathcal{S}_T)} &:= \|f\|_{C^{1+\alpha}_{Y,T}} + \|\nabla_d f\|_{C_B^{0,\alpha}(\mathcal{S}_T)}, \\
\|f\|_{C_B^{2,\alpha}(\mathcal{S}_T)} &= \|\nabla_d f\|_{C_B^{1,\alpha}(\mathcal{S}_T)} + \|Yf\|_{L^\infty_T(\mathbf{C}^\alpha)},
\end{aligned}$$

where

$$\|f\|_{C^\alpha_{\nabla_d,T}} := \sup_{(s,x)\in\mathcal{S}_T} \sup_{h\in\mathbb{R}^d} \frac{|f(s,x+(h,0)) - f(s,x)|}{|h|^\alpha}.$$

The main difference between this definition and Definition 1.2.4 is that $\|f\|_{C^\alpha_{\nabla_d,T}}$ has an Hölder condition in the non-degenerate directions $x_1,\dots,x_d$ only, while $|f\|_{L^\infty_t(\mathbf{C}^\alpha)}$ has an Hölder condition in all the variables with respect to the anisotropic norm. Nevertheless these two different definition for the intrinsic Hölder norms are equivalent. The argument follows [72] and we omit the proof here. The main idea is that increments in the degenerate directions can be obtained moving in the non-degenerate directions and along the integral curve of the vector field Y: this is possible for the crucial Assumption 1.1.2.

Chapter 2

Optimal regularity for degenerate Kolmogorov equations with rough coefficients

We consider a class of degenerate equations satisfying a parabolic Hörmander condition, with coefficients that are measurable in time and Hölder continuous in the space variables. By utilizing a generalized notion of strong solution, we establish the existence of a fundamental solution and its optimal Hölder regularity, as well as Gaussian estimates. These results are key to study the backward Kolmogorov equations associated to a class of Langevin-type diffusions.

Based on a joint work ([71]) with Profs. Stefano Pagliarani and Andrea Pascucci.

2.1 Main results

Throughout this chapter, we consider an operator $\mathcal{L}$ on $\mathcal{S}_{T_0}$ of the form (1.1), for a fixed $T_0 > 0$, under Assumptions 1.1.1 and 1.1.2; moreover the coefficients of $\mathcal{L}$ satisfy the following

Assumption 2.1.1. The coefficients a_{ij}, a_i, a of $\mathcal{L}$ are in in $L^\infty_{T_0}(\mathbf{C}^\alpha)$ for some $\alpha \in]0, 1]$.

According to Assumption 2.1.1, the coefficients of $\mathcal{A}$ are intrinsically Hölder continuous in the space variables and merely measurable in the time component. For Kolmogorov operators with coefficients that are Hölder continuous in both space and time, the study of the existence of a fundamental solution goes back to the early papers [87], [51], [84] and [68]. A modern and more natural approach based on the Lie group theory was developed by [79], [23], [6] and [74]. Applications to the martingale problem for some degenerate diffusion processes are given in [69] and [70].

Major questions in the study of Kolmogorov equations are the very definition of solution and its optimal regularity properties. It is well-known that, in general, the fundamental solution is not regular enough to support the derivatives ∂_{x_i}, for $d < i \leq N$, appearing in the transport term Y. Indeed, under the Hörmander condition (1.5), these derivatives are of order three and higher in the intrinsic sense. For this reason, even for equations with Hölder coefficients, weak notions of solution have been introduced. In this regard we may identify two main streams of research. In the semigroup approach initiated by [66], solutions are defined in the *distributional* sense: in this framework, solutions do not benefit from the time-smoothing effect that is typical of parabolic equations (see, for instance, Theorem 4.3 in [81]). On the other hand, in the stream of research started by [79], solutions in the *Lie* sense are defined by regarding Y as a directional derivative. In this approach regularity properties in space and time are strictly intertwined: this allows to fully exploit the smoothing effect of the equation but makes the analysis less suitable for applications to stochastic equations.

Recently, a third notion of solution, which is a cross between the two previous ones, has been proposed in [77] with the aim of studying Langevin *stochastic PDEs* with rough coefficients. Since we are specifically interested in operators whose coefficients are only measurable in time, it seems natural to adopt this latter approach for our analysis. We introduce the following definition that is a particular case of (1.3) in [77] when $N = 2$.

Definition 2.1.2 (Strong Lie solution). Let $0 < T \leq T_0$ and $f \in L^1_{\text{loc}}(]0,T[; C_b(\mathbb{R}^N))$. A solution to equation

$$\mathcal{A}u + Yu = f \quad \text{on } \mathcal{S}_T \tag{2.1}$$

is a continuous function u such that there exist $\partial_{x_i}u, \partial_{x_i x_j}u \in L^1_{\text{loc}}([0,T[; C_b(\mathbb{R}^N))$, for $i,j = 1,\dots,d$, and $Yu = f - \mathcal{A}u$ in the sense of Definition 1.2.3, i.e.

$$u(s, e^{(s-t)B}x) = u(t,x) - \int_t^s \left(\mathcal{A}u(\tau, e^{(\tau-t)B}x) - f(\tau, e^{(\tau-t)B}x)\right) d\tau, \qquad (t,x) \in \mathcal{S}_T,\ s < T. \tag{2.2}$$

Remark 2.1.3. Notice that $s \mapsto (s, e^{(s-t)B}x)$ is the integral curve of Y starting from (t,x):

for any suitably regular function u the limit

$$Yu(t,x) := \lim_{s\to t} \frac{u(s, e^{(s-t)B}x) - u(t,x)}{s-t} \tag{2.3}$$

is the directional (or Lie) derivative along Y of u at (t,x). Thus, if the integrand in (2.2) is continuous then u is a classical (pointwise) solution of (2.1). However, as noticed in Remark 2.2.9, in general a solution u in the sense of Definition 2.1.2 is only a.e. differentiable along Y and equation (2.1) is satisfied for almost every $(t,x) \in \mathcal{S}_T$.

In order to state our first main result, we give the following

Definition 2.1.4 (Fundamental solution). A fundamental solution of $\mathcal{A}+Y$ is a function $p = p(t,x;T,y)$ defined for $t < T$ and $x, y \in \mathbb{R}^N$ such that, for any fixed $(T,y) \in \mathcal{S}_{T_0}$, we have:

i) $p(\cdot,\cdot;T,y)$ is a solution of equation $\mathcal{A}u + Yu = 0$ on $\mathcal{S}_T$ in the sense of Definition 2.1.2;

ii) for any $g \in C_b(\mathbb{R}^N)$ we have

$$\lim_{\substack{(t,x)\to(T,y)\\ t<T}} \int_{\mathbb{R}^N} p(t,x;T,\eta) g(\eta) d\eta = g(y).$$

We draw attention to the fact that in point ii) of the previous definition the limit is pointwise.

The following result states the existence of the fundamental solution p of $\mathcal{A}+Y$, as well as uniform Gaussian bounds for p and its derivatives with respect to the non-degenerate variables $x_1, \ldots, x_d$.

Theorem 2.1.5 (Existence and Gaussian bounds). *Under Assumptions 1.1.1, 1.1.2 and 2.1.1, $\mathcal{L}$ has a fundamental solution $p = p(t,x;T,y)$ in the sense of Definition 2.1.4. For every $\varepsilon > 0$ there exists a positive constant C, only dependent on $T_0, \mu, B, \varepsilon, \alpha$ and the α-Hölder norms of the coefficients, such that*

$$p(t,x;T,y) \le C\Gamma^{\mu+\varepsilon}(t,x;T,y), \tag{2.4}$$

$$|\partial_{x_i} p(t,x;T,y)| \le \frac{C}{\sqrt{T-t}} \Gamma^{\mu+\varepsilon}(t,x;T,y), \tag{2.5}$$

$$\left|\partial_{x_i x_j} p(t,x;T,y)\right| \le \frac{C}{T-t} \Gamma^{\mu+\varepsilon}(t,x;T,y), \tag{2.6}$$

for any $(T,y) \in \mathcal{S}_{T_0}$*,* $(t,x) \in \mathcal{S}_T$ *and* $i,j = 1,\dots,d$*, where* Γ^{δ} *is the Gaussian fundamental solution of* (1.6)*, whose explicitly expression is given in* (2.11)*. Moreover, there exist two positive constants* $\bar{\mu}, \bar{c}$ *such that*

$$\bar{c}\,\Gamma^{\bar{\mu}}(t,x;T,y) \le p(t,x;T,y), \tag{2.7}$$

for any $(T,y) \in \mathcal{S}_{T_0}$ *and* $(t,x) \in \mathcal{S}_T$*.*

The proof of Theorem 2.1.5 is based on a modification of Levi's parametrix technique, which allows to deal with the lack of regularity of the coefficients along the drift term Y. The main tool is the fundamental solution of a Kolmogorov operator with time-dependent measurable coefficients, also recently studied in [12]. This approach allows for a careful analysis of the optimal regularity properties of the fundamental solution p: Theorem 2.1.6 below states that p belongs to the intrinsic Hölder space $\mathbf{C}_T^{2+\alpha}$ as given by Definition 1.2.4. As the notation could be misleading, we explicitly remark that for $u \in \mathbf{C}_T^{2+\alpha}$ not even the first order derivatives $\partial_{x_i} u$, for $i > d$, necessarily exist. However, in general we cannot expect higher regularity properties for solutions to (2.1) and $\mathbf{C}_T^{2+\alpha}$-regularity is indeed optimal.

Theorem 2.1.6 (Regularity of the fundamental solution). *Under the assumptions of Theorem 2.1.5,* $p(\cdot,\cdot;T,y) \in \mathbf{C}_{\tau}^{2+\beta}$ *for every* $(T,y) \in \mathcal{S}_{T_0}$*,* $0 < \tau < T$ *and* $\beta < \alpha$*. Precisely, there exists a positive constant* C *only dependent on* $T_0, \mu, B, \beta, \alpha$ *and the* α*-Hölder norms of the coefficients, such that*

$$\|p(\cdot,\cdot;T,y)\|_{\mathbf{C}_{\tau}^{2+\beta}} \le \frac{C}{(T-\tau)^{\frac{Q+2+\beta}{2}}}.$$

Theorem 2.1.6 refines the known results about the smoothness of the fundamental solution (cf. [66], [67], [24]) and exhibits its maximal regularity properties.

The rest of the chapter is structured as follows. Section 2.2 contains the construction of the fundamental solution by means of the parametrix method: in particular, Section 2.2.2 includes the proof of Theorem 2.1.5. In Section 2.3 we prove the regularity estimates of the fundamental solution, in particular Theorem 2.1.6. In Section 2.4 we state some more properties for the fundamental solution. The appendices contain the Gaussian and potential estimates that are employed in the proofs.

For reader's convenience, we recall that we shall always denote by $\mathcal{S}_T$ the strip $]0,T[\times\mathbb{R}^N$; also, in the following table we collect the notations used for the main functional spaces:

Notation	Functional space	Reference
$\mathbf{C}^\alpha$	Anisotropic Hölder spaces on $\mathbb{R}^N$	Def. 1.2.2
$C^\alpha_{Y,T}$	Lie Hölder spaces on $\mathcal{S}_T$	Def. 1.2.3
$\mathbf{C}^\alpha_T,\ k=0,1,2$	Intrinsic Hölder spaces on $\mathcal{S}_T$	Def. 1.2.4

2.2 Parametrix construction

Let Assumptions 1.1.1, 1.1.2 and 2.1.1 be satisfied. The first step of the parametrix method is to set a kernel $\mathbf{P} = \mathbf{P}(t,x;T,y)$ that serves as proxy for the fundamental solution, called *parametrix*. We denote by $\mathcal{A}^{(s,v)}$ the operator obtained by freezing the second-order coefficients of $\mathcal{A}$ along the integral curve of the vector field Y passing through $(s,v) \in \mathcal{S}_{T_0}$ and neglecting the lower order terms. Namely we consider the operator

$$\mathcal{A}^{(s,v)} := \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t, e^{(t-s)B}v)\partial_{x_i x_j}, \qquad (t,x) \in \mathcal{S}_{T_0}. \tag{2.8}$$

One can directly prove that the fundamental solution of

$$\mathcal{A}^{(s,v)} + Y,$$

in the sense of Definition 2.1.4, is given by

$$\Gamma^{(s,v)}(t,x;T,y) = \mathbf{G}(\mathcal{C}^{(s,v)}(t,T), y - e^{(T-t)B}x), \qquad (T,y) \in \mathcal{S}_{T_0},\ (t,x) \in \mathcal{S}_T,$$

where

$$\mathbf{G}(\mathcal{C}, z) := \frac{1}{\sqrt{(2\pi)^N \det \mathcal{C}}} e^{-\frac{1}{2}\langle \mathcal{C}^{-1} z, z\rangle}$$

is the Gaussian kernel on $\mathbb{R}^N$ and

$$\mathcal{C}^{(s,v)}(t,T) := \int_t^T e^{(T-\tau)B} A^{(s,v)}(\tau) e^{(T-\tau)B^*} d\tau, \tag{2.9}$$

$$A^{(s,v)}(\tau) := \begin{pmatrix} A_0(\tau, e^{(\tau-s)B}v) & 0 \\ 0 & 0 \end{pmatrix}, \qquad A_0 = (a_{ij})_{i,j=1,\dots,d}\,. \tag{2.10}$$

Remark 2.2.1. Clearly $\Gamma^{(s,v)}(t,x;T,y)$ is of class C^∞ as a function of x and only absolutely continuous along the integral curves of Y as a function of (t,x).

Remark 2.2.2. In the particular case of $A_0 \equiv \delta I_d$ for some $\delta > 0$, where I_d is the $(d \times d)$-identity matrix, the Kolmogorov operator $\mathcal{A}^{(s,v)} + Y$ reads as in (1.6) and its fundamental solution reduces to

$$\Gamma^{\delta}(t, x; T, y) := \mathbf{G}(\delta \mathcal{C}(T - t), y - e^{(T-t)B} x), \tag{2.11}$$

with

$$\mathcal{C}(t) = \int_0^t e^{(t-\tau)B} \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} e^{(t-\tau)B^*} d\tau. \tag{2.12}$$

Proceeding as in [69] and [77], we define the parametrix function $\mathbf{P}(t, x; T, y)$ as

$$\mathbf{P}(t, x; T, y) := \Gamma^{(T,y)}(t, x; T, y), \qquad (T, y) \in \mathcal{S}_{T_0},\ (t, x) \in \mathcal{S}_T, \tag{2.13}$$

and we refer to it as to the *time-dependent parametrix* in order to emphasize the fact that it is obtained by freezing only the space variable of the coefficients of $\mathcal{A}$.

Remark 2.2.3. Since $\Gamma^{(s,v)}$ is the fundamental solution of $\mathcal{A}^{(s,v)} + Y$, we have

$$(\mathcal{A}^{(T,y)} + Y)\mathbf{P}(\cdot, \cdot; T, y) = 0 \quad \text{on } \mathcal{S}_T, \tag{2.14}$$

in the sense of Definition 2.1.2, for any $(T, y) \in \mathcal{S}_{T_0}$.

Remark 2.2.4. In [23], where the variable coefficients of $\mathcal{A}$ are assumed intrinsically Hölder continuous in space and time, the parametrix is defined as the fundamental solution of the operator obtained by freezing the second order coefficients of $\mathcal{A}$ in both time and space variables, i.e.

$$\frac{1}{2} \sum_{i,j=1}^{d} a_{ij}(s, v) \partial_{x_i x_j} + Y.$$

As we shall see below, the choice of freezing the coefficients only in the space variable, along the integral curve of Y as in (2.8), is necessary in order to deal with the lack of regularity in the time variable.

Once the parametrix function is defined, the parametrix construction prescribes that a fundamental solution of $\mathcal{A} + Y$ is sought in the form

$$p(t, x; T, y) = \mathbf{P}(t, x; T, y) + \int_t^T \int_{\mathbb{R}^N} \mathbf{P}(t, x; \tau, \eta) \varphi(\tau, \eta; T, y) d\eta d\tau, \tag{2.15}$$

where φ is an unknown function. We now perform some heuristic computations that will lead to a fixed-point equation for φ. Assuming that $p(t,x;T,y)$ in (2.15) is a fundamental solution of $\mathcal{A}+Y$, we obtain

$$0 = (\mathcal{A}+Y)p(t,x;T,y) = (\mathcal{A}+Y)\mathbf{P}(t,x;T,y)+(\mathcal{A}+Y)\int_t^T\int_{\mathbb{R}^N}\mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau.$$

Furthermore, by formally differentiating and employing $p(t,x;t,\cdot)=\delta_x$ we also have

$$\begin{aligned}&(\mathcal{A}+Y)\int_t^T\int_{\mathbb{R}^N}\mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau\\ &=\int_t^T\int_{\mathbb{R}^N}(\mathcal{A}+Y)\mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau-\varphi(t,x;T,y).\end{aligned}$$

Therefore, $\varphi(t,x;T,y)$ must solve the Volterra integral equation

$$\varphi(t,x;T,y) = (\mathcal{A}+Y)\mathbf{P}(t,x;T,y)+\int_t^T\int_{\mathbb{R}^N}(\mathcal{A}+Y)\mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau. \quad (2.16)$$

Now, owing to Remark 2.2.3, equation (2.16) can be written as

$$\varphi(t,x;T,y) = (\mathcal{A}-\mathcal{A}^{(T,y)})\mathbf{P}(t,x;T,y)+\int_t^T\int_{\mathbb{R}^N}(\mathcal{A}-\mathcal{A}^{(\tau,\eta)})\mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau, \quad (2.17)$$

whose solution can be determined by an iterative procedure, which leads to the series representation

$$\varphi(t,x;T,y) = \sum_{k\geq 1}\varphi_k(t,x;T,y) \quad (2.18)$$

where

$$\begin{cases}\varphi_1(t,x;T,y) := \left(\mathcal{A}-\mathcal{A}^{(T,y)}\right)\mathbf{P}(t,x;T,y),\\ \varphi_{k+1}(t,x;T,y) := \displaystyle\int_t^T\int_{\mathbb{R}^N}(\mathcal{A}-\mathcal{A}^{(\tau,\eta)})\mathbf{P}(t,x;\tau,\eta)\varphi_k(\tau,\eta;T,y)d\eta d\tau, \qquad k\in\mathbb{N}.\end{cases} \quad (2.19)$$

In order to make the previous arguments rigorous one has to prove that:

- the series defined by (2.18)-(2.19) is uniformly convergent on $\mathcal{S}_T$. At this stage one also obtains a uniform upper bound and a Hölder estimate for φ;
- p defined by (2.15) is actually a fundamental solution of $\mathcal{A}+Y$. In this step one also establishes the Gaussian estimates on p and its derivatives that appear in Theorem 2.1.5.

2.2.1 Convergence of the series and estimates on φ

Proposition 2.2.5. *For every $(T,y) \in \mathcal{S}_{T_0}$ the series in (2.18) converges uniformly in $(t,x) \in \mathcal{S}_T$ and the function $\varphi = \varphi(t,x;T,y)$ solves the integral equation (2.17) on $\mathcal{S}_T$. Furthermore, for every $\varepsilon > 0$ and $0 < \delta < \alpha$, there exists a positive constant C, only dependent on $T_0, \mu, B, \delta, \alpha, \varepsilon$ and the α-Hölder norms of the coefficients, such that*

$$|\varphi(t,x;T,y)| \le \frac{C}{(T-t)^{1-\frac{\alpha}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y), \tag{2.20}$$

$$|\varphi(t,x;T,y) - \varphi(t,v;T,y)| \le \frac{C|x-v|_B^{\alpha-\delta}}{(T-t)^{1-\frac{\delta}{2}}}\left(\Gamma^{\mu+\varepsilon}(t,x;T,y) + \Gamma^{\mu+\varepsilon}(t,v;T,y)\right), \tag{2.21}$$

for any $(T,y) \in \mathcal{S}_{T_0}$ and $(t,x),(t,v) \in \mathcal{S}_T$.

To avoid repeating the arguments already used in [23], we limit ourself to highlighting the parts of the proof that differ significantly from the classical case.

Proof. We first prove that there exists a positive κ such that

$$|(\mathcal{A} - \mathcal{A}^{(T,y)})\mathbf{P}(t,x;T,y)| \le \frac{\kappa}{(T-t)^{1-\alpha/2}}\Gamma^{\mu+\varepsilon}(t,x;T,y), \qquad (T,y) \in \mathcal{S}_{T_0},\ (t,x) \in \mathcal{S}_T. \tag{2.22}$$

Assume for simplicity that $a_i, a \equiv 0$, the general case being a straightforward extension. By definition (2.8) we have

$$|(\mathcal{A} - \mathcal{A}^{(T,y)})\mathbf{P}(t,x;T,y)| \le \frac{1}{2}\sum_{i,j=1}^{d} |a_{ij}(t,x) - a_{ij}(t,e^{-(T-t)B}y)| \times |\partial_{x_i x_j}\mathbf{P}(t,x;T,y)| \tag{2.23}$$

(by the Hölder regularity of a_{ij} and the Gaussian estimate (2.B.3))

$$\le \kappa \frac{|x - e^{-(T-t)B}y|_B^{\alpha}}{T-t}\,\Gamma^{\mu+\varepsilon/2}(t,x;T,y). \tag{2.24}$$

The estimate (2.B.1) then yields (2.22).

For any $(T,y) \in \mathcal{S}_{T_0}$ and $(t,x) \in \mathcal{S}_T$, (2.19) and (2.22) imply

$$|\varphi_1(t,x;T,y)| \le \frac{\kappa}{(T-t)^{1-\alpha/2}}\,\Gamma^{\mu+\varepsilon}(t,x;T,y)$$

and

$$|\varphi_2(t,x;T,y)| \le \int_t^T \int_{\mathbb{R}^N} \left|\left(\mathcal{A} - \mathcal{A}^{(\tau,\eta)}\right)\mathbf{P}(t,x;\tau,\eta)\right| \times |\varphi_1(\tau,\eta;T,y)|d\eta d\tau$$

$$\leq \kappa^2 \int_t^T \frac{1}{(\tau-t)^{1-\alpha/2}} \frac{1}{(T-\tau)^{1-\alpha/2}} \int_{\mathbb{R}^N} \Gamma^{\mu+\varepsilon}(t,x;\tau,\eta)\Gamma^{\mu+\varepsilon}(\tau,\eta;T,y) d\eta d\tau =$$

(by the Chapman-Kolmogorov identity and solving the integral in $d\tau$)

$$= \kappa^2 \frac{\Gamma^2_{\text{Euler}}\left(\frac{\alpha}{2}\right)}{(T-t)^{1-\alpha}\Gamma_{\text{Euler}}(\alpha)} \Gamma^{\mu+\varepsilon}(t,x;T,y).$$

Proceeding by induction, it is straightforward to verify that

$$|\varphi_n(t,x;T,y)| \leq \kappa^n \frac{\Gamma^n_{\text{Euler}}\left(\frac{\alpha}{2}\right)}{(T-t)^{1-\frac{\alpha n}{2}}\Gamma_{\text{Euler}}\left(\frac{\alpha n}{2}\right)} \Gamma^{\mu+\varepsilon}(t,x;T,y), \qquad n \in \mathbb{N}.$$

This proves the uniform convergence of the series on $\mathcal{S}_T$, which in turn implies that φ satisfies (2.17), as well as the estimate (2.20).

The proof of (2.21) is a technical modification of the arguments in [23, Lemma 6.1], which is necessary to account for the different parametrix function. For sake of brevity, we leave the details to the reader. □

Remark 2.2.6. The proof above is particularly informative to understand the choice of the parametrix function in relation to the lack of regularity of the coefficients with respect to the time variable. In particular, in passing from (2.24) to (2.22), we take advantage of the increment $|x - e^{-(T-t)B}y|_B^\alpha$ in order to recover the integrability of the singularity in time. In the classical case, namely when the coefficient a_{ij} is also Hölder continuous in time, the parametrix function is obtained by freezing the variable coefficients in both space and time (see Remark 2.2.4). In (2.23), this choice leads to increments of type

$$\left|a_{ij}(t,x) - a_{ij}(T,y)\right|,$$

which is clearly not helpful if a_{ij} does not exhibit any regularity in time.

Furthermore, note that the coefficients have to be frozen in the space variable along the integral curve of Y: freezing the coefficients at a fixed point y would yield an increment of type $|x-y|_B^\alpha$ in (2.24), which does not allow to employ the Gaussian estimates in (2.B.1) to control the singularity.

2.2.2 Proof that p is a fundamental solution and Gaussian bounds

We now prove the first part of Theorem 2.1.5, concerning the existence of the fundamental solution of $\mathcal{L}$. This is achieved by proving that the candidate solution $p = p(t,x;T,y)$ defined through (2.15) satisfies points i) and ii) of Definition 2.1.4. The innovative part of the proof consists in showing point i), which is $p(\cdot,\cdot;T,y)$ solves the equation

$$\mathcal{A}u + Yu = 0 \quad \text{on } \mathcal{S}_T \tag{2.25}$$

in the integral sense of Definition 2.1.2. Once more, we provide the details of the parts that significantly depart from the classical case.

For any $(T,y) \in \mathcal{S}_{T_0}$, let us rewrite $p(t,x;T,y)$ as

$$p(t,x;T,y) = \mathbf{P}(t,x;T,y) + \Phi(t,x;T,y), \qquad (t,x) \in \mathcal{S}_T,$$

where we set

$$\Phi(t,x;T,y) := \int_t^T \int_{\mathbb{R}^N} \mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau. \tag{2.26}$$

The strategy of the proof is to first show that p possesses the regularity required in order to qualify as a fundamental solution, and then to check that it actually solves equation (2.25). As pointed out in Remark 2.14, the parametrix $\mathbf{P} = \mathbf{P}(t,x;T,y)$ is an integral solution to (2.14). In particular, it is smooth in the variable x and absolutely continuous along Y. As for $\Phi = \Phi(t,x;T,y)$, the next result shows that it is twice differentiable w.r.t. $x_1,\dots,x_d$ and states some Gaussian bounds on the derivatives.

Proposition 2.2.7. *For any $(T,y) \in \mathcal{S}_{T_0}$, $(t,x) \in \mathcal{S}_T$ and $i,j = 1,\dots,d$, there exist*

$$\partial_{x_i}\Phi(t,x;T,y) = \int_t^T \int_{\mathbb{R}^N} \partial_{x_i}\mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau,$$
$$\partial_{x_i x_j}\Phi(t,x;T,y) = \int_t^T \int_{\mathbb{R}^N} \partial_{x_i x_j}\mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau,$$

and, for any $\varepsilon > 0$ we have

$$|\Phi(t,x;T,y)| \leq C(T-t)^{\frac{\alpha}{2}}\Gamma^{\mu+\varepsilon}(t,x;T,y),$$
$$|\partial_{x_i}\Phi(t,x;T,y)| \leq \frac{C}{(T-t)^{\frac{1-\alpha}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y),$$
$$\left|\partial_{x_i x_j}\Phi(t,x;T,y)\right| \leq \frac{C}{(T-t)^{\frac{2-\alpha}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y),$$

where C denotes a positive constant, only dependent on $T_0, \mu, B, \alpha, \varepsilon$ and the α-Hölder norms of the coefficients.

Proof. By the definition of Φ in (2.26) we have

$$\Phi(t,x;T,y) = \int_t^T J(t,x;\tau;T,y)d\tau,$$

with J defined as in (2.C.1). The potential estimates of Proposition 2.C.1 upon integrating in τ, yield the result. $\square$

The following result shows that $\Phi(\cdot,\cdot;T,y)$ is also Lipschitz continuous along the integral curves of Y.

Lemma 2.2.8. *For every $(T,y) \in \mathcal{S}_{T_0}$ and $(t,x) \in \mathcal{S}_T$, we have*

$$\Phi\big(s, e^{(s-t)B}x;T,y\big) - \Phi(t,x;T,y) = -\int_t^s F(\tau,x;T,y)d\tau, \qquad s \in [t,T[,$$

where

$$F(\tau,x;T,y) := \int_\tau^T \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau, e^{(\tau-t)B}x;r,\eta\big)\varphi(r,\eta;T,y)d\eta dr + \varphi\big(\tau, e^{(\tau-t)B}x;T,y\big). \tag{2.27}$$

Proof. For any $s \in [t,T[$ one can write

$$\begin{aligned}&\Phi\big(s, e^{(s-t)B}x;T,y\big) - \Phi(t,x;T,y)\\ &= \underbrace{\int_s^T \int_{\mathbb{R}^N} \big(\mathbf{P}\big(s, e^{(s-t)B}x;r,\eta\big) - \mathbf{P}(t,x;r,\eta)\big)\,\varphi(r,\eta;T,y)d\eta dr}_{=:G(t,x)}\\ &\quad - \underbrace{\int_t^s \int_{\mathbb{R}^N} \mathbf{P}(t,x;r,\eta)\varphi(r,\eta;T,y)d\eta dr}_{=:H(t,x)}.\end{aligned}$$

First, we study the term $G(t,x)$. Remark 2.2.3 yields

$$G(t,x) = \int_s^T \int_{\mathbb{R}^N} \left(\int_t^s -\mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau, e^{(\tau-t)B}x;r,\eta\big)d\tau\right)\varphi(r,\eta;T,y)d\eta dr.$$

By (2.B.3) and Assumption 2.1.1, for every $\varepsilon > 0$ we have

$$\big|\mathcal{A}^{(r,\eta)}\mathbf{P}(\tau, e^{(\tau-t)B}x;r,\eta)\big| \leq \frac{C}{r-\tau}\Gamma^{\mu+\varepsilon}(\tau, e^{(\tau-t)B}x;r,\eta), \qquad t<\tau<s<r<T.$$

Therefore, considering also (2.20), for any $r \in]s,T]$, the function

$$(\tau,\eta) \mapsto |\mathcal{A}^{(r,\eta)}\mathbf{P}(\tau, e^{(\tau-t)B}x;r,\eta)\varphi(r,\eta;T,y)|$$

is integrable on $[t,s]\times\mathbb{R}^N$. Thus Fubini's theorem yields

$$\int_{\mathbb{R}^N}\left(\int_t^s \mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau,e^{(\tau-t)B}x;r,\eta\big)d\tau\right)\varphi(r,\eta;T,y)d\eta$$
$$=\int_t^s\int_{\mathbb{R}^N}\mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau,e^{(\tau-t)B}x;r,\eta\big)\varphi(r,\eta;T,y)d\eta d\tau.$$

Moreover, by the potential estimate (2.C.3) with $\delta=\frac{\alpha}{2}$, for any $\varepsilon>0$ we have

$$\left|\int_{\mathbb{R}^N}\mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau,e^{(\tau-t)B}x;r,\eta\big)\varphi(r,\eta;T,y)d\eta\right|\le\frac{C}{(T-r)^{1-\frac{\alpha}{4}}(r-\tau)^{1-\frac{\alpha}{4}}}\Gamma^{\mu+\varepsilon}\big(\tau,e^{(\tau-t)B}x;T,y\big).\tag{2.28}$$

As the right-hand side term is integrable over $[t,s]\times[s,T]$ as a function of (τ,r), we can apply once more Fubini's theorem to conclude that

$$G(t,x)=-\int_t^s\int_s^T\int_{\mathbb{R}^N}\mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau,e^{(\tau-t)B}x;r,\eta\big)\varphi(r,\eta;T,y)d\eta dr d\tau.\tag{2.29}$$

Let us consider $H(t,x)$. For every $n\in\mathbb{N}$, we define $\varepsilon_n(r):=\frac{1}{n}(r-t)$. Note that, for every $r\in]t,s[$ we have $r-\varepsilon_n(r)\ge t$. Hence

$$\begin{aligned}H(t,x)=&\underbrace{\int_t^s\int_{\mathbb{R}^N}\mathbf{P}\big(r-\varepsilon_n(r),e^{(r-\varepsilon_n(r)-t)B}x;r,\eta\big)\varphi(r,\eta;T,y)d\eta dr}_{=:\widetilde{H}_n(t,x)}\\&-\underbrace{\int_t^s\int_{\mathbb{R}^N}\left(\mathbf{P}\big(r-\varepsilon_n(r),e^{(r-\varepsilon_n(r)-t)B}x;r,\eta\big)-\mathbf{P}(t,x;r,\eta)\right)\varphi(r,\eta;T,y)d\eta dr}_{=:H_n(t,x)}.\end{aligned}$$

Once more, Remark 2.2.3 yields

$$H_n(t,x)=\int_t^s\int_{\mathbb{R}^N}\left(\int_t^{r-\varepsilon_n(r)}\mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau,e^{(\tau-t)B}x;r,\eta\big)d\tau\right)\varphi(r,\eta;T,y)d\eta dr$$

(applying Fubini's theorem as above)

$$=\int_t^s\int_t^{r-\varepsilon_n(r)}\int_{\mathbb{R}^N}\mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau,e^{(\tau-t)B}x;r,\eta\big)\varphi(r,\eta;T,y)d\eta d\tau dr$$

(setting $\delta_n(\tau)=\frac{\tau-t}{n-1}$ and applying Fubini's theorem again)

$$=\int_t^{s-\varepsilon_n(s)}\int_{\tau+\delta_n(\tau)}^s\int_{\mathbb{R}^N}\mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau,e^{(\tau-t)B}x;r,\eta\big)\varphi(r,\eta;T,y)d\eta dr d\tau$$

(by (2.28) and applying Lebesgue's dominated convergence theorem)

$$\xrightarrow[n\to\infty]{} \int_t^s \int_\tau^s \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)}\mathbf{P}\big(\tau, e^{(\tau-t)B}x; r,\eta\big)\varphi(r,\eta;T,y)d\eta dr d\tau.$$

On the other hand, by the potential estimate (2.C.2), for any $n \in \mathbb{N}$ we have

$$\left| \int_{\mathbb{R}^N} \mathbf{P}\big(r-\varepsilon_n(r), e^{(r-\varepsilon_n(r)-t)B}x; r,\eta\big)\varphi(r,\eta;T,y)d\eta \right| \leq C\frac{\Gamma^{\mu+\varepsilon}(\tau, e^{(\tau-t)B}x;T,y)}{(T-r)^{1-\frac{\alpha}{2}}(T-r)^{\frac{Q}{2}}}, \qquad r\in[t,s].$$

Thus Lebesgue's dominated convergence theorem yields

$$\lim_{n\to\infty} \widetilde{H}_n(t,x) = \int_t^s \lim_{n\to\infty} \int_{\mathbb{R}^N} \mathbf{P}\big(r-\varepsilon_n(r), e^{(r-\varepsilon_n(r)-t)B}x; r,\eta\big)\varphi(r,\eta;T,y)d\eta\, dr$$

(by (2.31), since $\eta \mapsto \varphi(r,\eta;T,y)$ is a bounded and continuous function for every $r \in [t,s]$)

$$= \int_t^s \varphi\big(r, e^{(r-t)B}x;T,y\big)dr.$$

We have proved that

$$H(t,x) = \int_t^s \int_\tau^s \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)}\mathbf{P}(\tau, e^{(\tau-t)B}x;r,\eta)\varphi(r,\eta;T,y)d\eta dr d\tau + \int_t^s \varphi(\tau, e^{(\tau-t)B}x;T,y)d\tau.$$

This and (2.29) prove the statement. □

We are now in the position to prove Theorem 2.1.5, namely that $p = p(t,x;T,y)$ defined by (2.15) is a fundamental solution of $\mathcal{A}+Y$ in the sense of definition Definition 2.1.4, and that the Gaussian bounds from (2.4) to (2.7) are satisfied.

Proof of Theorem 2.1.5. Let $p = p(t,x;T,y)$ be defined by (2.15).

Step 1. We show that $p = p(t,x;T,y)$ satisfies point i) of Definition 2.1.4, namely that $p(\cdot,\cdot;T,y)$ is an integral solution to (2.25) on $\mathcal{S}_T$ in the sense of Definition 2.1.2. By Lemma 2.2.8, we have

$$\begin{aligned} p\big(s, e^{(s-t)B}x;T,y\big) - p(t,x;T,y) &= \mathbf{P}(s, e^{(s-t)B}x;T,y) - \mathbf{P}(t,x;T,y) \\ &\quad + \Phi(s, e^{(s-t)B}x;T,y) - \Phi(t,x;T,y) \\ &= -\int_t^s \Big(\mathcal{A}^{(r,\eta)}\mathbf{P}(\tau, e^{(\tau-t)B}x;r,\eta) + F(\tau,x;T,y)\Big)d\tau. \end{aligned} \tag{2.30}$$

Furthermore, by (2.27) and since $\varphi(t,x;T,y)$ solves the integral equation (2.16), we obtain

$$\begin{aligned}\mathcal{A}^{(r,\eta)}\mathbf{P}(\tau,e^{(\tau-t)B}x;r,\eta)+F(\tau,x;T,y) &= \mathcal{A}\mathbf{P}(\tau,e^{(\tau-t)B}x;T,y)\\ &\quad+\int_\tau^T\int_{\mathbb{R}^N}\mathcal{A}\mathbf{P}(\tau,e^{(\tau-t)B}x;r,\eta)\varphi(r,\eta;T,y)d\eta dr\end{aligned}$$

(by Proposition 2.2.7)

$$\begin{aligned}&= \mathcal{A}\mathbf{P}(\tau,e^{(\tau-t)B}x;T,y)+\mathcal{A}\Phi(\tau,e^{(\tau-t)B}x;T,y)\\ &= \mathcal{A}p(\tau,e^{(\tau-t)B}x;T,y),\end{aligned}$$

which, together with (2.30), concludes the proof.

Step 2. We show that $p=p(t,x;T,y)$ satisfies point ii) of Definition 2.1.4. In light of the estimate (2.20), it is straightforward to see that

$$|\Phi(t,x;T,y)|\le C(T-t)^{\frac{\alpha}{2}}\Gamma^{\mu+\varepsilon}(t,x;T,y),\qquad (T,y)\in\mathcal{S}_{T_0},\ (t,x)\in\mathcal{S}_T.$$

Therefore, it is enough to prove that, for any fixed $(T,y)\in]0,T_0[\times\mathbb{R}^N$, we have

$$\lim_{\substack{(t,x)\to(T,y)\\ t<T}}\int_{\mathbb{R}^N}\mathbf{P}(t,x;T,\eta)f(\eta)d\eta=f(y),\qquad f\in C_b(\mathbb{R}^N). \tag{2.31}$$

Recalling the definition of the parametrix $\mathbf{P}$, we add and subtract to obtain

$$\begin{aligned}\int_{\mathbb{R}^N}\mathbf{P}(t,x;T,\eta)f(\eta)d\eta &= \int_{\mathbb{R}^N}\Gamma^{(T,\eta)}(t,x;T,\eta)f(\eta)d\eta\\ &= \int_{\mathbb{R}^N}\Gamma^{(T,y)}(t,x;T,\eta)f(\eta)d\eta\\ &\quad+\underbrace{\int_{\mathbb{R}^N}\left(\Gamma^{(T,\eta)}(t,x;T,\eta)-\Gamma^{(T,y)}(t,x;T,\eta)\right)f(\eta)d\eta}_{=:J(t,x)}.\end{aligned}$$

Furthermore, by estimate (2.B.4), for every $\varepsilon>0$ one has

$$|J(t,x)|\le C\int_{\mathbb{R}^N}|y-\eta|_B^\alpha\Gamma^{\mu+\varepsilon}(t,x;T,\eta)d\eta.$$

Eventually, (2.31) follows from classical arguments.

Step 3. We show the upper Gaussian bounds (2.4)-(2.5)-(2.6) for p and its derivatives. The proof of the lower Gaussian bound (2.7) is similar to that of Theor. 4.7 in [76] and Section 5.1.4. in [77], thus we omit it for sake of brevity.

The Gaussian bounds of Proposition 2.B.2 and the definition of parametrix (2.13) yield the estimates (2.4)-(2.5)-(2.6) for $\mathbf{P} = \mathbf{P}(t,x;T,y)$. The estimates of Proposition 2.2.7 and the fact that $p = \mathbf{P} + \Phi$ conclude the proof. □

Remark 2.2.9. Any integral solution u to equation (2.2) on $\mathcal{S}_T$ in the sense of Definition 2.1.2, is Lie-differentiable along Y almost everywhere on $\mathcal{S}_T$. Indeed, the set H_T of $(t,x) \in \mathcal{S}_T$ such that $Yu(t,x)$ in (2.3) exists finite, is measurable as the limit

$$\limsup_{\tau \to t^+} \frac{u\big(\tau, e^{(\tau-t)B}x\big) - u(t,x)}{\tau - t}$$

is a measurable function of (t,x) and the same holds for lim inf. This is a straightforward consequence of the continuity of u along the integral curves of Y. The fact that H_T has null Lebesgue measure stems from Fubini's theorem, as u is absolutely continuous along the integral curves of Y and the map

$$(\tau, y) \mapsto (\tau, e^{\tau B} y)$$

is a diffeomorphism on $\mathcal{S}_T$.

2.3 Regularity of the fundamental solution

In this section we prove Theorem 2.1.6. Since $p(\cdot,\cdot;T,y)$ can be represented as in (2.15), we need to study the regularity of $\mathbf{P}(\cdot,\cdot;T,y)$ and $\Phi(\cdot,\cdot;T,y)$. While the former term can be easily dealt with by means of the Gaussian estimates of Appendix 2.A, the latter has to be treated more carefully. We start with the proof of Theorem 2.1.6, which is based on the regularity estimates for $\Phi(\cdot,\cdot;T,y)$ and $\mathbf{P}(\cdot,\cdot;T,y)$ proved in Section 2.3.1 and Section 2.3.2, respectively.

Proof of Theorem 2.1.6. Let $\beta < \alpha$. For fixed $(T,y) \in \mathcal{S}_{T_0}$, we set

$$f(t,x) := p(t,x;T,y), \qquad (t,x) \in \mathcal{S}_T.$$

We first note that, by definition of fundamental solution, (3.5) is satisfied with $f_Y = -\mathcal{A}f$. Furthermore, for any $t \in]0,T[$, by Definition 1.2.4 and the representation (2.15) we have

$$\|f\|_{\mathbf{C}_t^{2+\beta}} = N_{\mathbf{P},1} + N_{\mathbf{P},2} + N_{\Phi,1} + N_{\Phi,2},$$

where

$$N_{\mathbf{P},1} := \sum_{i=1}^{d} \|\partial_{x_i}\mathbf{P}(\cdot,\cdot;T,y)\|_{C^{1+\beta}_{Y,t}} + \sum_{i,j=1}^{d} \left(\|\partial_{x_i x_j}\mathbf{P}(\cdot,\cdot;T,y)\|_{C^{\beta}_{Y,t}} + \|\partial_{x_i x_j}\mathbf{P}(\cdot,\cdot;T,y)\|_{L^\infty_t(\mathbf{C}^\beta)}\right),$$

$$N_{\mathbf{P},2} := \|\mathcal{A}\mathbf{P}(\cdot,\cdot;T,y)\|_{L^\infty_t(\mathbf{C}^\beta)},$$

$$N_{\Phi,1} := \sum_{i=1}^{d} \|\partial_{x_i}\Phi(\cdot,\cdot;T,y)\|_{C^{1+\beta}_{Y,t}} + \sum_{i,j=1}^{d} \left(\|\partial_{x_i x_j}\Phi(\cdot,\cdot;T,y)\|_{C^{\beta}_{Y,t}} + \|\partial_{x_i x_j}\Phi(\cdot,\cdot;T,y)\|_{L^\infty_t(\mathbf{C}^\beta)}\right),$$

$$N_{\Phi,2} := \|\mathcal{A}\Phi(\cdot,\cdot;T,y)\|_{L^\infty_t(\mathbf{C}^\beta)}.$$

Now, the estimates of Lemma 2.3.4 yield

$$N_{\mathbf{P},1} \le \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}. \tag{2.32}$$

To bound $N_{\mathbf{P},2}$, first fix $i,j = 1,\dots,d$ and note that, by estimate (2.B.3), we obtain

$$\sup_{x\in\mathbb{R}^N} \left|\partial_{x_i x_j}\mathbf{P}(s,x;T,y)\right| \le \frac{C}{(T-t)^{\frac{Q+2}{2}}}, \qquad s<t. \tag{2.33}$$

Furthermore, (2.32) combined with Remark 1.2.7 yield

$$\sup_{x,v\in\mathbb{R}^N} \frac{|\partial_{x_i x_j}\mathbf{P}(s,x;T,y) - \partial_{x_i x_j}\mathbf{P}(s,v;T,y)|}{|x-v|_B^\beta} \le \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}, \qquad s<t. \tag{2.34}$$

Thus, by (2.33)-(2.34) we obtain

$$\|\partial_{x_i x_j}\mathbf{P}(s,\cdot;T,y)\|_{\mathbf{C}^\beta_t} \le \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}, \qquad s<t,$$

which in turn implies

$$\|\partial_{x_i x_j}\mathbf{P}(\cdot,\cdot;T,y)\|_{L^\infty_t(\mathbf{C}^\beta)} \le \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}.$$

This, together with Assumption 2.1.1, prove

$$N_{\mathbf{P},2} \le \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}.$$

The bound for $N_{\Phi,1}$ stems from the estimates of Proposition 2.3.2, which yield

$$N_{\Phi,1} \le \frac{C}{(T-t)^{\frac{Q+2-(\alpha-\beta)}{2}}} \le \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}.$$

Eventually, the bound for $N_{\Phi,2}$ follows from the same arguments used to bound $N_{\mathbf{P},2}$. □

The rest of this section is devoted to the results utilized in the proof of Theorem 2.1.6. It is useful to introduce the following

Notation 2.3.1. Let $f = f(t, x; T, y)$ be a function defined for $(T, y) \in \mathcal{S}_{T_0}$ and $(t, x) \in \mathcal{S}_T$, suitably differentiable w.r.t. x. For any $i = 1, \dots, N$, we set

$$\partial_i f(t, x; T, y) := \partial_{x_i} f(t, x; T, y),$$

and we adopt analogous notations for the higher-order derivatives.

This notation is useful in order to compose partial derivatives with other functions. For instance, if $g = g(t, x)$ is a given function, then

$$\partial_i f\big(t, g(t, x); T, y\big) = \partial_{z_i} f(t, z; T, y)|_{z=g(t,x)}.$$

2.3.1 Regularity estimates of Φ

Now prove the Hölder estimates for $\Phi(\cdot, \cdot; T, y)$. We recall that Q denotes the homogeneous dimension of $\mathbb{R}^N$ as in (1.9).

Proposition 2.3.2. *For every $\varepsilon > 0$ and $0 < \beta < \alpha$ there exists a positive constant C, only dependent on $T_0, \mu, B, \varepsilon, \alpha, \beta$ and the α-Hölder norms of the coefficients, such that, for any $i, j, k = 1, \dots, d$, we have*

$$\left|\partial_i \Phi(s, e^{(s-t)B}x; T, y) - \partial_i \Phi(t, x; T, y)\right| \le C(s-t)^{\frac{1+\beta}{2}} \frac{(T-t)^{Q/2}}{(T-s)^{\frac{Q+2-(\alpha-\beta)}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{2.35}$$

$$\left|\partial_{ij} \Phi(s, e^{(s-t)B}x; T, y) - \partial_{ij} \Phi(t, x; T, y)\right| \le C(s-t)^{\frac{\beta}{2}} \frac{(T-t)^{Q/2}}{(T-s)^{\frac{Q+2-(\alpha-\beta)}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{2.36}$$

$$|\partial_{ij} \Phi(t, x + h\mathbf{e}_k; T, y) - \partial_{ij} \Phi(t, x; T, y)| \le C|h|^{\beta} \frac{\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)}{(T-t)^{\frac{2-(\alpha-\beta)}{2}}}, \tag{2.37}$$

for every $(T, y) \in \mathcal{S}_{T_0}$, $(t, x) \in \mathcal{S}_T$, $t < s < T$ and $h \in \mathbb{R}$.

The proof of estimates (2.35)-(2.36) relies on the following

Lemma 2.3.3. *Let $(T, y) \in \mathcal{S}_{T_0}$. Then, for any $i = 1, \dots, d$, the function $u := \partial_i \mathbf{P}(\cdot, \cdot; T, y)$ is a strong Lie solution to the equation*

$$\mathcal{A}u + Yu = -\sum_{j=1}^{d+d_1} b_{ji} \partial_j \mathbf{P}(\cdot, \cdot; T, y) \quad \text{on } \mathcal{S}_T,$$

in the sense of Definition 2.1.2.

Proof. We note that

$$[\partial_i, Y]\mathbf{P}(t,x;T,y) = [\partial_i, \langle Bx, \nabla\rangle + \partial_t]\mathbf{P}(t,x;T,y) = \sum_{j=1}^{d+d_1} b_{ji}\partial_j \mathbf{P}(t,x;T,y),$$

for every $x \in \mathbb{R}^N$ and for almost every $t \in [0,T[$, where, in the last equality, we used that $b_{ji} = 0$ if $j > d+d_1$. While it is obvious that the previous identity holds for smooth functions of (t,x), one can directly check that $\partial_i\partial_t\mathbf{P}(t,x;T,y) = \partial_t\partial_i\mathbf{P}(t,x;T,y)$ and thus the identity holds for the parametrix too. Therefore, we obtain

$$\begin{aligned}
&\partial_i\mathbf{P}(s, e^{(s-t)B}x;\tau,\eta) - \partial_i\mathbf{P}(t,x;\tau,\eta)\\
&= \int_t^s (Y\partial_i\mathbf{P})(r, e^{(r-t)B}x;\tau,\eta)dr\\
&= \int_t^s \Big((\partial_i Y\mathbf{P})(r, e^{(r-t)B}x;\tau,\eta) - [\partial_i, Y]\mathbf{P}(r, e^{(r-t)B}x;\tau,\eta)\Big)dr
\end{aligned}$$

(by Remark 2.2.3)

$$= -\int_t^s \Big((\partial_i\mathcal{A}^{(\tau,\eta)}\mathbf{P})(r, e^{(r-t)B}x;\tau,\eta) + \sum_{j=1}^{d+d_1} b_{ji}\partial_j\mathbf{P}(r, e^{(r-t)B}x;\tau,\eta)\Big)dr$$

(since $\partial_i\mathcal{A}^{(\tau,\eta)} = \mathcal{A}^{(\tau,\eta)}\partial_i$)

$$= -\int_t^s \Big(\big(\mathcal{A}^{(\tau,\eta)}\partial_i\mathbf{P}\big)(r, e^{(r-t)B}x;\tau,\eta) + \sum_{j=1}^{d+d_1} b_{ji}\partial_j\mathbf{P}(r, e^{(r-t)B}x;\tau,\eta)\Big)dr.$$

□

We are now in the position to prove Proposition 2.3.2.

Proof of Proposition 2.3.2. Let $(T,y) \in \mathcal{S}_{T_0}$, $(t,x) \in \mathcal{S}_T$, $t < s < T$ and $h \in \mathbb{R}$ be fixed. Also fix $i,j,k \in \{1,\dots,d\}$. First we prove (2.35). By adding and subtracting, we have

$$\begin{aligned}
&\partial_i\Phi(s, e^{(s-t)B}x;T,y) - \partial_i\Phi(t,x;T,y)\\
&= \int_s^T \int_{\mathbb{R}^N} \underbrace{\Big(\partial_i\mathbf{P}(s, e^{(s-t)B}x;\tau,\eta) - \partial_i\mathbf{P}(t,x;\tau,\eta)\Big)}_{=:I(\tau,\eta)} \varphi(\tau,\eta;T,y)d\eta d\tau\\
&\quad \underbrace{-\int_t^s \int_{\mathbb{R}^N} \partial_i\mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau}_{=:L}.
\end{aligned}$$

We consider the first term. By Lemma 2.3.3 and swapping the integrals as in the proof of Proposition 2.2.8, we have

$$\int_s^T \int_{\mathbb{R}^N} I(\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau$$
$$= -\int_s^T\int_t^s\int_{\mathbb{R}^N}\Big(\big(\mathcal{A}^{(\tau,\eta)}\partial_i\mathbf{P}\big)(r,e^{(r-t)B}x;\tau,\eta) + \sum_{j=1}^{d+d_1} b_{ji}\partial_j\mathbf{P}(r,e^{(r-t)B}x;\tau,\eta)\Big)\varphi(\tau,\eta;T,y)d\eta dr d\tau.$$

Therefore, the estimates of Proposition 2.C.1 with $\delta = (\alpha-\beta)/2$ yield

$$\Big|\int_s^T \int_{\mathbb{R}^N} I(\tau,\eta)\varphi(\tau,\eta;T,y)d\eta d\tau\Big|$$
$$\leq \int_s^T\int_t^s \frac{C}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-r)^{\frac{3}{2}-\frac{\alpha+\beta}{4}}}\Gamma^{\mu+\varepsilon}(r,e^{(r-t)B}x;T,y)dr d\tau$$

(by a standard estimate on $\Gamma^{\mu+\varepsilon}(r,e^{(r-t)B}x;T,y)$)

$$\leq C\underbrace{\int_s^T\int_t^s \frac{1}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-r)^{\frac{3}{2}-\frac{\alpha+\beta}{4}}}dr d\tau}_{=:K}\left(\frac{T-t}{T-s}\right)^{Q/2}\Gamma^{\mu+\varepsilon}(t,x;T,y). \tag{2.38}$$

We now bound K:

$$\begin{aligned}
K =& \int_t^s\int_s^T \frac{1}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-r)^{\frac{3}{2}-\frac{\alpha+\beta}{4}}}d\tau dr \\
\leq& \int_t^s\int_s^T \frac{1}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-r)^{1-\frac{\alpha-\beta}{4}}}d\tau \frac{1}{(s-r)^{\frac{1}{2}-\frac{\beta}{2}}}dr
\end{aligned}$$

(solving the integral in $d\tau$)

$$\begin{aligned}
\leq& C\int_t^s \frac{1}{(T-r)^{1-\frac{\alpha-\beta}{2}}}\frac{1}{(s-r)^{\frac{1}{2}-\frac{\beta}{2}}}dr \\
\leq& \frac{C}{(T-s)^{1-\frac{\alpha-\beta}{2}}}\int_t^s \frac{1}{(s-r)^{\frac{1-\beta}{2}}}dr \\
\leq& \frac{C}{(T-s)^{1-\frac{\alpha-\beta}{2}}}(s-t)^{\frac{1+\beta}{2}}.
\end{aligned} \tag{2.39}$$

On the other hand, estimate (2.C.3) with $\delta = \alpha-\beta$ yields

$$|L| \leq \int_t^s \frac{C}{(T-\tau)^{1-\frac{\alpha-\beta}{2}}(\tau-t)^{\frac{1}{2}-\frac{\beta}{2}}}d\tau\, \Gamma^{\mu+\varepsilon}(t,x;T,y)$$

$$\leq \frac{C}{(T-s)^{1-\frac{\alpha-\beta}{2}}} \int_t^s \frac{1}{(\tau-t)^{\frac{1-\beta}{2}}} d\tau\, \Gamma^{\mu+\varepsilon}(t,x;T,y)$$
$$\leq \frac{C}{(T-s)^{1-\frac{\alpha-\beta}{2}}} (s-t)^{\frac{1+\beta}{2}} \Gamma^{\mu+\varepsilon}(t,x;T,y).$$

This, together with (2.38)-(2.39), proves (2.35). Estimate (2.36) can be obtained following the same arguments.

We finally prove (2.37). By Proposition 2.2.7 we have

$$\partial_{ij}\Phi(t,x+h\mathbf{e}_k;T,y) - \partial_{ij}\Phi(t,x;T,y)$$
$$= \int_t^T \underbrace{\int_{\mathbb{R}^N} \big(\partial_{ij}\mathbf{P}(t,x+h\mathbf{e}_k;\tau,\eta) - \partial_{ij}\mathbf{P}(t,x;\tau,\eta)\big)\varphi(\tau,\eta;T,y) d\eta}_{=:I(\tau)}\, d\tau.$$

We first prove that

$$|I(\tau)| \leq C \frac{|h|^\beta}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-t)^{1-\frac{\alpha-\beta}{4}}} \left(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y) + \Gamma^{\mu+\varepsilon}(t,x;T,y)\right), \qquad \tau \in]t,T[.$$

We consider the case $\tau - t \geq h^2$. By the mean-value theorem, there exists a real $\bar{h}$ with $|\bar{h}| \leq |h|$ such that

$$|\partial_{ij}\mathbf{P}(t,x+h\mathbf{e}_k;\tau,\eta) - \partial_{ij}\mathbf{P}(t,x;\tau,\eta)| = |h| \left|\partial_{ijk}\mathbf{P}(t,x+\bar{h}\mathbf{e}_k;\tau,\eta)\right|.$$

Therefore, by the estimate (2.C.3) with $\delta = (\alpha-\beta)/2$, we have

$$|I(\tau)| \leq C \frac{|h|}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-t)^{\frac{3}{2}-\frac{\alpha+\beta}{4}}} \Gamma^{\mu+\varepsilon}(t,x+\bar{h}\mathbf{e}_k;T,y)$$

(since $\tau - t \geq h^2$)

$$\leq C \frac{|h|^\beta}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-t)^{1-\frac{\alpha-\beta}{4}}} \Gamma^{\mu+\varepsilon}(t,x+\bar{h}\mathbf{e}_k;T,y)$$

(by standard estimates on $\Gamma^{\mu+\varepsilon}(t,x+\bar{h}\mathbf{e}_k;T,y)$ with $\tau - t \geq h^2$)

$$\leq C \frac{|h|^\beta}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-t)^{1-\frac{\alpha-\beta}{4}}} \big(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y) + \Gamma^{\mu+\varepsilon}(t,x;T,y)\big).$$

We now consider the case $\tau - t < h^2$. Employing triangular inequality and estimate (2.C.3) with $\delta = (\alpha-\beta)/2$, we get

$$|I(\tau)| \leq \frac{C}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-t)^{1-\frac{\alpha+\beta}{4}}} \big(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y) + \Gamma^{\mu+\varepsilon}(t,x;T,y)\big)$$

(since $\tau - t < h^2$)

$$\leq C\frac{|h|^{\beta}}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-t)^{1-\frac{\alpha-\beta}{4}}}\big(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y)+\Gamma^{\mu+\varepsilon}(t,x;T,y)\big).$$

Therefore, combining the previous estimates, we obtain

$$\begin{aligned}\left|\int_t^T I(\tau)d\tau\right| &\leq C|h|^{\beta}\int_t^T\frac{1}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-t)^{1-\frac{\alpha-\beta}{4}}}d\tau\big(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y)+\Gamma^{\mu+\varepsilon}(t,x;T,y)\big)\\ &\leq C|h|^{\beta}\frac{1}{(T-t)^{1-\frac{\alpha-\beta}{2}}}\big(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y)+\Gamma^{\mu+\varepsilon}(t,x;T,y)\big),\end{aligned}$$

which proves (2.37). □

2.3.2 Regularity estimates for the parametrix

We have the following Hölder estimates for $\mathbf{P}$.

Lemma 2.3.4. *Let $0 \leq \beta \leq \alpha$. Then for every $\varepsilon > 0$ there exists a positive constant C, only dependent on $T_0, \mu, B, \varepsilon, \alpha, \beta$ and the α-Hölder norms of the coefficients, such that for any $i, j, k = 1, \dots, d$ we have*

$$\left|\partial_i\mathbf{P}(s,e^{(s-t)B}x;T,y)-\partial_i\mathbf{P}(t,x;T,y)\right| \leq C(s-t)^{\frac{1+\beta}{2}}\frac{(T-t)^{Q/2}}{(T-s)^{\frac{Q+2+\beta}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y), \tag{2.40}$$

$$\left|\partial_{ij}\mathbf{P}(s,e^{(s-t)B}x;T,y)-\partial_{ij}\mathbf{P}(t,x;T,y)\right| \leq C(s-t)^{\frac{\beta}{2}}\frac{(T-t)^{Q/2}}{(T-s)^{\frac{Q+2+\beta}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y), \tag{2.41}$$

$$\left|\partial_{ij}\mathbf{P}(t,x+h\mathbf{e}_k;T,y)-\partial_{ij}\mathbf{P}(t,x;T,y)\right| \leq C|h|^{\beta}\frac{1}{(T-t)^{\frac{2+\beta}{2}}}\big(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y)+\Gamma^{\mu+\varepsilon}(t,x;T,y)\big), \tag{2.42}$$

for any $(T,y) \in \mathcal{S}_{T_0}$, $(t,x) \in \mathcal{S}_T$, $t < s < T$ and $h \in \mathbb{R}$.

Proof. We first consider (2.40). By Lemma 2.3.3 we have

$$\partial_i\mathbf{P}(s,e^{(s-t)B}x;T,y)-\partial_i\mathbf{P}(t,x;T,y) = -\int_t^s\Big(\mathcal{A}^{(T,y)}\partial_i\mathbf{P}(r,e^{(r-t)B}x;T,y)+\sum_{j=1}^{d+d_1}b_{ji}\partial_j\mathbf{P}(r,e^{(r-t)B}x;T,y)\Big)dr$$

Therefore, by boundedness of the coefficients of $\mathcal{A}^{(T,y)}$ and the estimates of Proposition 2.B.2, we obtain

$$\left|\partial_i\mathbf{P}(s,e^{(s-t)B}x;T,y)-\partial_i\mathbf{P}(t,x;T,y)\right| \leq \int_t^s\frac{C}{(T-r)^{\frac{3}{2}}}\Gamma^{\mu+\varepsilon}(r,e^{(r-t)B}x;T,y)dr$$

$$\le \int_t^s \frac{C}{(T-r)^{\frac{3}{2}}} dr \left(\frac{T-t}{T-s}\right)^{Q/2} \Gamma^{\mu+\varepsilon}(t,x;T,y)$$

(for any $\beta \le 1$)

$$\le C \frac{(s-t)^{\frac{1+\beta}{2}}}{(T-s)^{1+\frac{\beta}{2}}} \left(\frac{T-t}{T-s}\right)^{Q/2} \Gamma^{\mu+\varepsilon}(t,x;T,y).$$

The proof of (2.41) is based on analogous arguments.

We finally prove (2.42). As for (2.37), we first consider the case $T-t \ge h^2$. By the mean-value theorem, there exists a real $\bar{h}$ with $|\bar{h}| \le |h|$ such that

$$|\partial_{ij}\mathbf{P}(t,x+h\mathbf{e}_k;T,y) - \partial_{ij}\mathbf{P}(t,x;T,y)| = |h| \left|\partial_{ijk}\mathbf{P}(t,x+\bar{h}\mathbf{e}_k;T,y)\right|$$

(by estimate (2.B.3))

$$\le C \frac{|h|}{(T-t)^{\frac{3}{2}}} \Gamma^{\mu+\varepsilon}(t,x+\bar{h}\mathbf{e}_k;T,y)$$

(since $T-t \ge h^2$ and by standard estimates on $\Gamma^{\mu+\varepsilon}(t,x+\bar{h}\mathbf{e}_k;T,y)$)

$$\le C \frac{|h|^\beta}{(T-t)^{1+\frac{\beta}{2}}} \left(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y) + \Gamma^{\mu+\varepsilon}(t,x;T,y)\right).$$

We now consider $T-t < h^2$. Employing triangular inequality and estimate (2.B.3) yields

$$|\partial_{ij}\mathbf{P}(t,x+h\mathbf{e}_k;T,y) - \partial_{ij}\mathbf{P}(t,x;T,y)| \le \frac{C}{T-t} \left(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y) + \Gamma^{\mu+\varepsilon}(t,x;T,y)\right)$$

(since $T-t < h^2$)

$$\le C \frac{|h|^\beta}{(T-t)^{1+\frac{\beta}{2}}} \left(\Gamma^{\mu+\varepsilon}(t,x+h\mathbf{e}_k;T,y) + \Gamma^{\mu+\varepsilon}(t,x;T,y)\right).$$

This concludes the proof of (2.42). □

2.4 More properties for the fundamental solution

The following proposition contains further useful properties that allow to view the fundamental solution as the transition probability density of a Markovian process. We omit the proofs that are based on rather standard arguments.

Proposition 2.4.1. *Under the assumptions of Theorem 2.1.5 we have:*

i) the Chapman-Kolmogorov identity

$$p(t,x;T,y) = \int_{\mathbb{R}^N} p(t,x;s,\eta)p(s,\eta;T,y)d\eta, \qquad t < s < T, \ x,y \in \mathbb{R}^N;$$

ii) if the zeroth order coefficient a of $\mathcal{A}$ is constant, i.e. $a(t,x) = \bar{a}$, then

$$\int_{\mathbb{R}^N} p(t,x;T,y)dy = e^{\bar{a}(T-t)}, \qquad t < T, \ x \in \mathbb{R}^N.$$

2.A Gaussian estimates

We prove Gaussian estimates that are crucial in the analysis of Sections 2.2 and 2.3. Here we follow the ideas in [23, Section 3], but with some technical difference. Namely, in the aforementioned paper the Kolmogorov operator acts on the forward variables of $\Gamma^{(s,v)}(t,x;T,y)$, whereas here we consider $\mathcal{L} = \mathcal{A} + Y$ acting on the backward variables (t,x). This has an impact on the spatial derivatives, which contain additional factors that require a careful analysis.

Throughout the appendix we suppose that Assumptions 1.1.1, 1.1.2 and 2.1.1 are satisfied and fix $(s,v) \in \mathcal{S}_{T_0}$. Denoting by B_0 the matrix B with null $*$-blocks, we define the $N \times N$ matrices

$$\mathcal{C}_0(t) := \int_0^t e^{(t-\tau)B_0} \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} e^{(t-\tau)B_0^*} d\tau,$$
$$\mathcal{C}_0^{(s,v)}(t,T) := \int_t^T e^{(T-\tau)B_0} A^{(s,v)}(\tau) e^{(T-\tau)B_0^*} d\tau,$$

with $A^{(s,v)}$ as defined in (3.10). As an immediate consequence of Assumption 1.1.1 we can compare the quadratic forms associated to $\mathcal{C}^{(s,v)}$ (as in (2.9)), $\mathcal{C}_0^{(s,v)}$ with $\mathcal{C}(T-t)$ (as in (2.12)), $\mathcal{C}_0(T-t)$, respectively:

$$\frac{1}{\mu}\mathcal{C}(T-t) \leq \mathcal{C}^{(s,v)}(t,T) \leq \mu\mathcal{C}(T-t),$$
$$\frac{1}{\mu}\mathcal{C}_0(T-t) \leq \mathcal{C}_0^{(s,v)}(t,T) \leq \mu\mathcal{C}_0(T-t), \tag{2.A.1}$$

for any $t \leq T$. Moreover, an asymptotic comparison near 0 of $\mathcal{C}^{(s,v)}$ and $\mathcal{C}_0^{(s,v)}$ holds:

Lemma 2.A.1. *There exist two positive constants C and δ, only dependent on μ and B, such that*

$$\frac{1}{2\mu}\mathcal{C}_0(T-t) \leq \mathcal{C}^{(s,v)}(t,T) \leq 2\mu\mathcal{C}_0(T-t),$$
$$\frac{1}{(2\mu)^N}\det\mathcal{C}_0(T-t) \leq \det\mathcal{C}^{(s,v)}(t,T) \leq (2\mu)^N\det\mathcal{C}_0(T-t),$$

for any $0 < T-t < \delta$. Analogous estimates hold for $\left(\mathcal{C}^{(s,v)}(t,T)\right)^{-1}$.

Proof. It follows from the same arguments of [60, Lemma 3.1]: the proof is only based on the properties of the matrices A and B, and it is not relevant whether A has constant or time-dependent entries. □

Remark 2.A.2. We note that $|\cdot|_B$ is homogeneous with respect to the family of dilations defined by the matrices

$$D(\lambda) := \operatorname{diag}(\lambda I_d, \lambda^3 I_{d_1}, \dots, \lambda^{2q+1} I_{d_q}), \qquad \lambda \geq 0.$$

In [60, Proposition 2.3] it is proved that

$$\mathcal{C}_0(t) = D(\sqrt{t})\mathcal{C}_0(1)D(\sqrt{t}), \qquad t \geq 0. \tag{2.A.2}$$

Therefore, for $0 < T-t < \delta$ with δ as in Lemma 2.A.1,

$$\frac{(T-t)^Q}{(2\mu)^N}\det\mathcal{C}_0(1) \leq \det\mathcal{C}^{(s,v)}(t,T) \leq (2\mu)^N (T-t)^Q \det\mathcal{C}_0(1).$$

To compute the spatial derivatives of $\Gamma^{(s,v)}(t,x;T,y)$ it is useful noticing that

$$\Gamma^{(s,v)}(t,x;T,y) = \mathbf{G}\big(H^{(s,v)}(t,T), e^{-(T-t)B}y - x\big), \qquad (T,y) \in \mathcal{S}_{T_0},\ (t,x) \in \mathcal{S}_T,$$

where

$$H^{(s,v)}(t,T) := e^{-(T-t)B}\mathcal{C}^{(s,v)}(t,T)e^{-(T-t)B^*}.$$

Since $\mathcal{C}^{(s,v)}(t,T)$ is symmetric positive definite and $e^{-(T-t)B}$ is non-singular, then $H^{(s,v)}(t,T)$ is symmetric and positive definite for every $0 \leq t < T$.

In order to give estimates on the matrix $H^{(s,v)}$ we need to study the elements of e^{tB}. We recall the block partition (1.7) of the matrix B: for $h,k = 0,\dots,q$, we denote the $d_h \times d_k$ block of B by

$$\mathcal{Q}_{hk} := (b_{ij})_{\substack{i=\bar{d}_{h-1}+1,\dots,\bar{d}_h \\ j=\bar{d}_{k-1}+1,\dots,\bar{d}_k}},$$

with $\bar{d}_h$ as in (1.8). Note that by (1.7) we have

$$\begin{cases} \mathcal{Q}_{hk} = 0_{d_h \times d_k} & \text{if } h > k+1, \\ \mathcal{Q}_{hk} = B_h & \text{if } h = k+1, \\ \mathcal{Q}_{hk} = * & \text{if } h < k+1. \end{cases} \tag{2.A.3}$$

Analogously, for $n \in \mathbb{N}$, we can consider the same block decomposition for B^n. We denote by $\mathcal{Q}_{hk}^{(n)}$ the $d_h \times d_k$ block of B^n.

Lemma 2.A.3. *Let $h, k = 0, \dots, q$ and $n \in \mathbb{N}$. Then*

$$\mathcal{Q}_{hk}^{(n)} = 0_{d_h \times d_k}, \qquad h > k+n, \tag{2.A.4}$$

which is $(B^n)_{ij} = 0$ if $i \in \{\bar{d}_{h-1}+1, \dots, \bar{d}_h\}$ and $j \in \{\bar{d}_{k-1}+1, \dots, \bar{d}_k\}$.

Proof. We proceed by induction on n. The case of $n = 1$ is obvious (see (2.A.3)). Now we assume that (2.A.4) holds for a certain $n \in \mathbb{N}$. For $h > k+n+1$ we have

$$\mathcal{Q}_{hk}^{(n+1)} = \sum_{m=0}^{q} \mathcal{Q}_{hm}^{(n)} \mathcal{Q}_{mk}.$$

If $m < h-n$, then $\mathcal{Q}_{hm}^{(n)} = 0_{d_h \times d_m}$ by inductive hypothesis; if $m \geq h-n$, then $m > k+1$ and $\mathcal{Q}_{mk} = 0_{d_m \times d_k}$. Therefore $\mathcal{Q}_{hk}^{(n+1)} = 0_{d_h \times d_k}$. □

Lemma 2.A.4. *Let $h, k = 1, \dots, q$ such that $h - k =: n \in \mathbb{N}$. For any $i \in \{\bar{d}_{h-1}+1, \dots, \bar{d}_h\}$ and $j \in \{\bar{d}_{k-1}+1, \dots, \bar{d}_k\}$ we have*

$$\left(e^{tB}\right)_{ij} = O(t^n), \quad \text{as } t \to 0.$$

Proof. From Lemma 2.A.3 we have that $(B^m)_{ij} = 0$ for every $m = 0, \dots, n-1$, since $\mathcal{Q}_{hk}^{(m)} = 0_{d_h \times d_k}$ for $h - k = n > m$. Therefore

$$(e^{tB})_{ij} = \frac{t^n (B^n)_{ij}}{n!} + O(t^{n+1}), \quad \text{as } t \to 0.$$

□

Lemma 2.A.5. *There exists a positive constant C that only depends on μ, B and T_0 such that, for every $i, j = 1, \dots, d$ and $k = d+1, \dots, d+d_1$,*

$$\left|\left(H^{(s,v)}(t,T)^{-1}x\right)_i\right| \leq \frac{C}{\sqrt{T-t}} \left|D(\sqrt{T-t})^{-1} e^{(T-t)B} x\right|, \tag{2.A.5}$$

$$\left|(H^{(s,v)}(t,T)^{-1})_{ij}\right| \leq \frac{C}{T-t}, \tag{2.A.6}$$

$$\left|\left(H^{(s,v)}(t,T)^{-1}x\right)_k\right| \leq \frac{C}{(T-t)^{\frac{3}{2}}}\left|D(\sqrt{T-t})^{-1}e^{(T-t)B}x\right|, \tag{2.A.7}$$

$$\left|\left(H^{(s,v)}(t,T)^{-1}\right)_{ik}\right| \leq \frac{C}{(T-t)^2}, \tag{2.A.8}$$

for any $0 < T < T_0$ *and* $(t,x) \in \mathcal{S}_T$.

Proof. We prove the first inequality. Setting $\tau = T-t$, we have

$$\begin{aligned}\left|\left(H^{(s,v)}(t,T)^{-1}x\right)_i\right| &= \frac{1}{\sqrt{\tau}}\left|\left(D(\sqrt{\tau})e^{\tau B^*}\mathcal{C}^{(s,v)}(t,T)^{-1}e^{\tau B}x\right)_i\right| \\ &\leq \frac{1}{\sqrt{\tau}}\sum_{n=1}^{N}\left|\left(D(\sqrt{\tau})e^{\tau B^*}D(\sqrt{\tau})^{-1}\right)_{in}\right|\left\|D(\sqrt{\tau})\mathcal{C}^{(s,v)}(t,T)^{-1}D(\sqrt{\tau})\right\|\left|D(\sqrt{\tau})^{-1}e^{\tau B}x\right|.\end{aligned}$$

By Lemma 2.A.1 there exists a positive constant δ such that, if $0 < \tau < \delta$, we have

$$\begin{aligned}\left\|D(\sqrt{\tau})\mathcal{C}^{(s,v)}(t,T)^{-1}D(\sqrt{\tau})\right\| &\leq \sup_{|y|=1}\langle D(\sqrt{\tau})\mathcal{C}^{(s,v)}(t,T)^{-1}D(\sqrt{\tau})y,y\rangle \\ &\leq 2\mu\sup_{|y|=1}\langle \mathcal{C}_0(\tau)^{-1}D(\sqrt{\tau})y, D(\sqrt{\tau})y\rangle = 2\mu\|\mathcal{C}_0(1)^{-1}\|,\end{aligned}$$

where the last equality follows from (2.A.2). If $\delta \leq \tau < T_0$, by equation (2.A.1) we have

$$\left\|D(\sqrt{\tau})\mathcal{C}^{(s,v)}(t,T)^{-1}D(\sqrt{\tau})\right\| \leq \mu\left\|D(\sqrt{\tau})\mathcal{C}(\tau)^{-1}D(\sqrt{\tau})\right\|,$$

which is bounded by a constant that depends only on μ, T_0 and B.

In order to conclude the proof of (2.A.5), we let h_n be the only $h \in \{0,\dots,q\}$ such that $\bar{d}_{h-1}+1 \leq n \leq \bar{d}_h$. Then, by Lemma 2.A.4, since $i \in \{1,\dots,d\}$, we obtain

$$\begin{aligned}\left(D(\sqrt{\tau})e^{\tau B^*}D(\sqrt{\tau})^{-1}\right)_{in} =& D(\sqrt{\tau})_{ii}\left(e^{\tau B^*}\right)_{in}D(\sqrt{\tau})^{-1}_{nn} \\ =& \tau^{\frac{1}{2}}\left(e^{\tau B}\right)_{ni}\tau^{-\frac{2h_n+1}{2}} = O(1) \quad \text{as } \tau \to 0.\end{aligned}$$

Estimate (2.A.6) follows from (2.A.5) choosing $x = \mathbf{e}_j$. Estimates (2.A.7) and (2.A.8) can be proved following the same arguments, noticing that for $k = d+1,\dots,d+d_1$ we have $D(\tau)_{kk} = \tau^3$. □

2.B Gaussian estimates

Finally, we provide Gaussian estimates for $\Gamma^{(s,v)}(t,x;T,y)$ and its derivatives up to the fourth order that will be used to study the Hölder regularity of the second order derivatives of the fundamental solution via the representation (2.15)-(2.19). The following result can be proved as [23, Proposition 3.5].

Lemma 2.B.1. *For every $\beta \geq 0$ and $\varepsilon > 0$ there exists a positive constant C, only dependent on T_0, μ, B, ε and β, such that*

$$|w_i|^\beta \, \Gamma^{(s,v)}(t,x;T,y) \leq C\Gamma^{\mu+\varepsilon}(t,x;T,y), \qquad (T,y) \in \mathcal{S}_{T_0},\ (t,x) \in \mathcal{S}_T,\ i = 1,\dots,N, \tag{2.B.1}$$

where

$$w = D(\sqrt{T-t})^{-1}\left(y - e^{(T-t)B}x\right).$$

Combining Lemmas 2.A.5 and 2.B.1 with [23, Proposition 3.1, 3.6 and Lemma 5.2], some lengthy but straightforward computations show the following

Proposition 2.B.2. *We have*

$$\frac{1}{\mu^N}\Gamma^{\frac{1}{\mu}}(t,x;T,y) \leq \Gamma^{(s,v)}(t,x;T,y) \leq \mu^N \Gamma^\mu(t,x;T,y). \tag{2.B.2}$$

for any $(T,y) \in \mathcal{S}_{T_0}$ and $(t,x) \in \mathcal{S}_T$. Moreover, for every $\varepsilon > 0$ and $\nu \in \mathbb{N}_0^N$ with $[\nu]_B \leq 4$, there exists a positive constant C, only dependent on T_0, μ, B and ε, such that

$$|\partial_x^\nu \Gamma^{(s,v)}(t,x;T,y)| \leq \frac{C}{(T-t)^{\frac{[\nu]_B}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y), \tag{2.B.3}$$

$$\left|\partial_x^\nu \Gamma^{(s,v)}(t,x;T,y) - \partial_x^\nu \Gamma^{(s,w)}(t,x;T,y)\right| \leq C\frac{|v-w|_B^\alpha}{(T-t)^{\frac{[\nu]_B}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y), \tag{2.B.4}$$

for any $(T,y) \in \mathcal{S}_{T_0}$, $(t,x) \in \mathcal{S}_T$ and $w \in \mathbb{R}^N$.

2.C Potential estimates

We study $\Phi = \Phi(t,x;T,y)$ in (2.26) and its derivatives w.r.t. to the variables $x_1,\dots,x_d$. To do so, we have to deal with some singular integrals. We follow the steps in [23, Section 5], but we remark that the estimates of Proposition 2.C.1 extend the ones in the aforementioned paper to higher order derivatives. This is needed to prove the optimal regularity of $\Phi(t,x;T,y)$ and thereafter of $p(t,x;T,y)$.

We set

$$J(t,x;\tau;T,y) := \int_{\mathbb{R}^N} \mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta, \qquad (T,y) \in \mathcal{S}_{T_0},\ (t,x) \in \mathcal{S}_T,\ \tau \in]t,T[. \tag{2.C.1}$$

Proposition 2.C.1. *For every $\varepsilon > 0$, $\nu \in \mathbb{N}_0^N$ with $[\nu]_B \leq 4$ and $0 < \delta < \alpha$, there exists a positive constant C, only dependent on $N, T_0, \mu, B, \delta, \alpha$ and ε, such that,*

$$|J(t,x;\tau;T,y)| \leq \frac{C}{(T-\tau)^{1-\frac{\alpha}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y) \tag{2.C.2}$$

$$|\partial_x^\nu J(t,x;\tau;T,y)| \leq \frac{C}{(T-\tau)^{1-\frac{\delta}{2}}(\tau-t)^{\frac{[\nu]_B-(\alpha-\delta)}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y), \tag{2.C.3}$$

for every $(T,y) \in \mathcal{S}_{T_0}$, $(t,x) \in \mathcal{S}_T$ and $\tau \in]t,T[$.

Proof. The proof relies on Proposition 2.2.5: (2.C.2) can be easily obtained by applying estimate (2.B.2) to $\mathbf{P}(t,x;\tau,\eta)$, estimate (2.20) to $\varphi(\tau,\eta;T,y)$ and the Chapman-Kolmogorov identity.

We provide a full proof of (2.C.3) in the case of $\partial_x^\nu = \partial_{x_i x_j}$, with $i,j \leq d$, the proof for higher order derivatives being analogous. The idea is to combine (2.21) with the techniques in [23, Proposition 5.3] and [79, Proposition 3.2]. Let $(t,x) \in \mathcal{S}_T$ and $\tau \in]t,T[$ be fixed. By estimates (2.B.3) and (2.20), we have

$$\partial_{x_i x_j} J(t,x;\tau;T,y) = \int_{\mathbb{R}^N} \partial_{x_i x_j}\mathbf{P}(t,x;\tau,\eta)\varphi(\tau,\eta;T,y)d\eta.$$

We set $\bar{t} = \frac{t+T}{2}$ and consider two separate cases:
Case $\bar{t} < \tau < T$. By (2.B.3) and (2.20), we have that for every $\varepsilon > 0$ and $0 < \delta < \alpha$ there exists a positive constant C such that

$$|\partial_{x_i x_j} J(t,x;\tau;T,y)| \leq \int_{\mathbb{R}^N} \frac{C}{(T-\tau)^{1-\frac{\alpha}{2}}(\tau-t)}\Gamma^{\mu+\varepsilon}(t,x;\tau,\eta)\Gamma^{\mu+\varepsilon}(\tau,\eta;T,y)d\eta$$

(by the Chapman-Kolmogorov equation)

$$\leq \frac{C}{(T-\tau)^{1-\frac{\alpha}{2}}(\tau-t)}\Gamma^{\mu+\varepsilon}(t,x;T,y)$$

(since $T-\tau < \tau - t$)

$$\leq \frac{C}{(T-\tau)^{1-\frac{\delta}{2}}(\tau-t)^{1-\frac{\alpha-\delta}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y).$$

Case $t < \tau \leq \bar{t}$. Here we need to handle with care the singularity of $\partial_{x_i x_j}\mathbf{P}(t,x;\tau,\eta)$ for small $\tau - t$. Note that in this case the following inequalities hold true:

$$\tau - t \leq \frac{T-t}{2} \leq T - \tau < T - t. \tag{2.C.4}$$

We have

$$\partial_{x_i x_j} J(t,x;\tau;T,y) = K_1 + K_2 + K_3,$$

where, setting $\xi = e^{(\tau-t)B}x$,

$$K_1 := \int_{\mathbb{R}^N} \partial_{x_i x_j}\Gamma^{(\tau,\eta)}(t,x;\tau,\eta)\big(\varphi(\tau,\eta;T,y) - \varphi(\tau,\xi;T,y)\big)d\eta,$$
$$K_2 := \varphi(\tau,\xi;T,y)\int_{\mathbb{R}^N}\left(\partial_{x_i x_j}\Gamma^{(\tau,\eta)}(t,x;\tau,\eta) - \partial_{x_i x_j}\Gamma^{(\tau,v)}(t,x;\tau,\eta)\big|_{v=\xi}\right)d\eta,$$
$$K_3 := \varphi(\tau,\xi;T,y)\int_{\mathbb{R}^N}\partial_{x_i x_j}\Gamma^{(\tau,v)}(t,x;\tau,\eta)\big|_{v=\xi}d\eta.$$

We first consider K_1. By (2.21) and (2.B.3), for every $\varepsilon > 0$ and $0 < \delta < \alpha$ there exists a positive constant C such that

$$|K_1| \le \frac{C}{(T-\tau)^{1-\frac{\delta}{2}}}\int_{\mathbb{R}^N}\frac{|\eta-\xi|_B^{\alpha-\delta}}{(\tau-t)}\Gamma^{\mu+\frac{\varepsilon}{2}}(t,x;\tau,\eta)\left(\Gamma^{\mu+\varepsilon}(\tau,\xi;T,y) + \Gamma^{\mu+\varepsilon}(\tau,\eta;T,y)\right)d\eta$$

(by (2.B.1))

$$\le \frac{C}{(T-\tau)^{1-\frac{\delta}{2}}}\int_{\mathbb{R}^N}\frac{1}{(\tau-t)^{1-\frac{\alpha-\delta}{2}}}\Gamma^{\mu+\varepsilon}(t,x;\tau,\eta)\left(\Gamma^{\mu+\varepsilon}(\tau,\xi;T,y) + \Gamma^{\mu+\varepsilon}(\tau,\eta;T,y)\right)d\eta$$

(integrating in η and by the Chapman-Kolmogorov identity)

$$\le \frac{C}{(T-\tau)^{1-\frac{\delta}{2}}(\tau-t)^{1-\frac{\alpha-\delta}{2}}}\left(\Gamma^{\mu+\varepsilon}(\tau,\xi;T,y) + \Gamma^{\mu+\varepsilon}(t,x;T,y)\right)$$

(by (2.C.4))

$$\le \frac{C}{(T-\tau)^{1-\frac{\delta}{2}}(\tau-t)^{1-\frac{\alpha-\delta}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y).$$

Consider now K_2. By (2.20) and (2.B.4), we obtain

$$|K_2| \le C\frac{\Gamma^{\mu+\varepsilon}(\tau,\xi;T,y)}{(T-\tau)^{1-\frac{\alpha}{2}}}\int_{\mathbb{R}^N}\frac{|\eta-\xi|_B^{\alpha}}{\tau-t}\Gamma^{\mu+\varepsilon}(t,x;\tau,\eta)d\eta$$

(by (2.B.1) and integrating in η)

$$\le \frac{C}{(T-\tau)^{1-\frac{\alpha}{2}}(\tau-t)^{1-\frac{\alpha}{2}}}\Gamma^{\mu+\varepsilon}(\tau,\xi;T,y)$$

(again by (2.C.4))

$$\le \frac{C}{(T-\tau)^{1-\frac{\delta}{2}}(\tau-t)^{1-\frac{\alpha-\delta}{2}}}\Gamma^{\mu+\varepsilon}(t,x;T,y).$$

Finally, $K_3 = 0$ since

$$\int_{\mathbb{R}^N} \partial_{x_i x_j} \Gamma^{(\tau,v)}(t, x; \tau, \eta) d\eta = \partial_{x_i x_j} \int_{\mathbb{R}^N} \Gamma^{(\tau,v)}(t, x; \tau, \eta) d\eta = 0$$

for any $v \in \mathbb{R}^N$. $\square$

Chapter 3

Optimal Schauder estimates for kinetic Kolmogorov equations with time measurable coefficients

We prove global Schauder estimates for kinetic Kolmogorov equations with coefficients that are Hölder continuous in the spatial variables but only measurable in time. Compared to other available results in the literature, our estimates are optimal in the sense that the inherent Hölder spaces are the strongest possible under the given assumptions: in particular, under a parabolic Hörmander condition, we introduce Hölder norms defined in terms of the intrinsic geometry that the operator induces on the space-time variables. The technique is based on the existence and the regularity estimates of the fundamental solution of the equation, recently proved in [71]. These results are essential for studying backward Kolmogorov equations associated with kinetic-type diffusions, e.g. stochastic Langevin equation.

Based on a joint work ([65]) with Profs. Stefano Pagliarani and Andrea Pascucci.

3.1 Introduction

In recent years, several sharp Schauder estimates have been proved for Kolmogorov equations with coefficients that are Hölder-continuous in the space-variables but only measurable in time. In this chapter we prove global Schauder estimates which we claim to be *optimal*, meaning that the inherent Hölder spaces are the strongest possible under the given assumptions on the coefficients. In particular, our results include and improve some known estimates in the framework of non-divergence form operators satisfying a parabolic Hörmander condition.

A prototype example of the class under consideration is

$$\frac{\sigma^2(t,v,x)}{2}\partial_{vv} + v\partial_x + \partial_t, \qquad (t,v,x) \in \mathbb{R}^3,$$

which is the backward Kolmogorov operator of the system of stochastic differential equations

$$\begin{cases} dV_t = \sigma(t, V_t, X_t)dW_t, \\ dX_t = V_t dt, \end{cases}$$

where W is a real Brownian motion. The study of these models is motivated by several applications, including kinetic theory and finance. In the classical Langevin model, (V, X) describes velocity and position of a particle in the phase space and is a pilot example of more complex kinetic models (cf. [59], [52], [53]). In mathematical finance, (V, X) represents the log-price and average processes utilized in modeling path-dependent financial derivatives, such as Asian options (cf. [7], [75]).

We consider the Cauchy problem

$$\begin{cases} \mathcal{L}u = f & \text{on } \mathcal{S}_T, \\ u(T, \cdot) = g & \text{on } \mathbb{R}^N, \end{cases} \tag{3.1}$$

posed on the strip

$$\mathcal{S}_T :=]0, T[\times \mathbb{R}^N.$$

Generally speaking, Schauder estimates give a bound of some Hölder norm of the solution u in terms of some (possibly different) Hölder norms of the data, namely the coefficients of $\mathcal{L}$, the non-homogeneous term f and the datum g. The "strength" of a Schauder estimate depends on the norms involved and this is a sensitive issue in the theory of degenerate PDEs: clearly, for a given Hölder norm on the solution u, the weaker the norms on the data, the stronger the Schauder estimate; conversely, if the norms on the data are given, then the stronger the norm on u the stronger the Schauder estimate. Now, in the literature on degenerate Kolmogorov equations, we may recognize at least two notions of Hölder norm as well as variants of them: the so-called *anisotropic* and *intrinsic* norms, whose precise definitions are given in Section 1.2. Intuitively, the former norm takes into account the anisotropic behavior in space induced by the underlying diffusion, but does not require any time-regularity. The intrinsic norm, by opposite, is induced by the geometric properties of the differential operator $\mathcal{L}$ and takes into account the regularity along the vector field Y (and thus along the time variable). Therefore, the intrinsic norm is stronger in the sense that it allows to see the full regularizing effect (in both space and time) of the fundamental solution of $\mathcal{L}$.

Roughly speaking, we may catalogue the known Schauder estimates for solutions to (3.1) as follows:

- *anisotropic-to-anisotropic*: the anisotropic norms of the data bound the anisotropic norm of the solution, as in [66], [64], [81], [17] and [44];
- *intrinsic-to-intrinsic*: the intrinsic norms of the data bound the intrinsic norm of the solution, as in [67], [24], [43], [52] and [80];
- *anisotropic-to-intrinsic*: the anisotropic norms of the data bound the intrinsic norm of the solution. This class is stronger than the two above: only recently, partial results were proved in [26] and [8].

Our main result, Theorem 3.2.2, provides an optimal *anisotropic-to-intrinsic* global Schauder estimate which improves the results in [26] and [8] in a subtle but crucial way, as explained in Section 3.2.2. For instance, the Hölder norm we adopt for the solution u is strong enough to derive intrinsic Taylor formulas and therefore also an Itô formula for the underlying diffusion processes, which is a fundamental tool in stochastic calculus.

Our estimate is global in that it holds all the way up to the boundary, with an explosion factor that depends on the regularity of g, namely $(T-t)^{-\frac{2+\alpha-\beta}{2}}$: here α and β represent the Hölder exponents of the solution u and of the terminal datum g, respectively. In particular, it is possible to recognize two limiting cases: $\beta = 2+\alpha$, no explosion close to the boundary; $\beta = 0$ (g only continuous), maximum explosion rate. As a corollary, we obtain a sharp regularity estimate (Corollary 3.2.3) in the Y direction, at the boundary $t = T$. Although this is an expected phenomenon, the quantitative characterization of the penalty term is novel, to the best of our knowledge, for degenerate Kolmogorov operators in the context of variable coefficients or of intrinsic Hölder spaces. We refer the reader to [88] for the case of constant diffusion coefficients and anisotropic Besov-Hölder spaces. We also mention that intrinsic embedding theorems of Sobolev type were recently proved in [35] and [78], and L^p global Schauder estimates recently appeared in [25].

Theorem 3.2.2 comprises a well-posedness result for (3.1). The proof of Theorem 3.2.2 goes as follows. First we define a candidate solution u via Duhamel principle. After proving that it is actually a solution to (3.1), we prove the regularity estimates for the two convolution terms that constitute u, from which the Schauder estimate follows. The proof critically relies on the recent results in [71], where the existence of the fundamental solution of $\mathcal{L}$ was established, together with optimal Hölder estimates, by means of a suitable modification of the parametrix technique, already employed in [79] and [23] in the case of intrinsic Hölder-continuous coefficients.

The rest of the chapter is organized as follows. Section 3.2 contains the main results and a detailed comparison with the related literature. Precisely, in Section 3.2.1 we state our main result, Theorem 3.2.2, and comment on it; Section 3.2.2 contains the comparison with the literature. Section 3.3 is entirely devoted to the proof Theorem 3.2.2.

3.2 Schauder estimates

In this section we state the main results of this chapter. We start by introducing the assumption on the regularity for the coefficients of $\mathcal{L}$, that is exactly Assumption 2.1.1 in slightly different setting:

Assumption 3.2.1. The coefficients a_{ij}, a_i, a of $\mathcal{L}$ are in $L^\infty_T(\mathbf{C}^{\bar{\alpha}})$ for some $\bar{\alpha} \in]0,1]$.

3.2.1 Main result

We remark that if $u \in L^\infty_T(\mathbf{C}^{2+\alpha})$ or $u \in \mathbf{C}^{2+\alpha}_T$ then the derivatives $\partial_{x_i} u, \partial_{x_i x_j} u$ exist in the classical sense only for $1 \leq i, j \leq d$. Indeed, u is not regular enough to support even the first-order derivatives ∂_{x_i} for $d < i \leq N$ and the full gradient ∇u appearing in the drift of $\mathcal{L}$ must be interpreted in a suitable way. Thus, in accordance with the intrinsic Hölder spaces defined above, we consider solutions for the kinetic equation in the sense of Definition 2.1.2. Moreover, for $g \in C(\mathbb{R}^N)$, a solution to the Cauchy problem (3.1) is a solution u to (2.1), which can be extended continuously to $]0,T] \times \mathbb{R}^N$ with $u(T, \cdot) = g$.

Theorem 3.2.2. *Let Assumptions 1.1.1, 1.1.2 and 3.2.1 be in force. Then, for any* $\alpha \in]0, \bar{\alpha}[$, $g \in \mathbf{C}^{\beta}$ *with* $\beta \in [0, 2+\alpha]$ *and* $f \in L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})$ *with* $\gamma \in [0, 1[$, *there exists a unique bounded strong Lie solution* u *to the Cauchy problem* (3.1). *Furthermore, we have*

$$\|u\|_{\mathbf{C}_t^{2+\alpha}} \leq C\left((T-t)^{-\frac{2+\alpha-\beta}{2}}\|g\|_{\mathbf{C}^{\beta}} + (T-t)^{-\gamma}\|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})}\right), \qquad 0 < t < T, \tag{3.2}$$

where C *is a positive constant which depends only on* $T, B, \bar{\alpha}, \alpha, \beta, \gamma$ *and on the norms* $L^{\infty}_{T}(\mathbf{C}^{\bar{\alpha}})$ *of the coefficients of* $\mathcal{A}$. *In particular,* C *does not depend on* t.

We illustrate the Schauder estimate through particular instances; by linearity, we can treat the cases $g = 0$ and $f = 0$ separately:

- [**Case** $g = 0$] Estimate (3.2) reads as

$$\|u\|_{\mathbf{C}_t^{2+\alpha}} \leq C(T-t)^{-\gamma}\|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})}, \qquad 0 < t < T. \tag{3.3}$$

  In particular, if $\gamma = 0$ then f is bounded and (3.3) holds true up to the boundary $t = T$:

$$\|u\|_{\mathbf{C}_T^{2+\alpha}} \leq C\,\|f\|_{L^{\infty}_{T}(\mathbf{C}^{\alpha})}.$$

- [**Case** $f = 0$] We have two extreme cases:

  ◇ If $\beta = 0$, that is g is only bounded and continuous, then the solution has the same explosion behavior of the fundamental solution as $t \to T^-$ (cf. [71]), which is

$$\|u\|_{\mathbf{C}_t^{2+\alpha}} \leq C(T-t)^{-\frac{2+\alpha}{2}}\|g\|_{L^{\infty}}.$$

  ◇ If $\beta = 2+\alpha$ then the solution is $(2+\alpha)$-Hölder continuous up to the boundary

$$\|u\|_{\mathbf{C}_t^{2+\alpha}} \leq C\|g\|_{\mathbf{C}^{2+\alpha}}. \tag{3.4}$$

Estimate (3.4) entails a regularity result along Y at the boundary $t = T$, which is reported in the following

Corollary 3.2.3. *Let the assumptions of Theorem 3.2.2 be in force with* $\gamma = 0$. *The solution* u *to the Cauchy problem* (3.1) *satisfies*

$$\|u(t, e^{B(T-t)}\cdot) - u(T, \cdot)\|_{L^{\infty}} \leq C(T-t)^{\frac{\beta}{2}}, \qquad 0 < t < T,$$

if $\beta \in]0, 2]$ *and*

$$\|u(t, e^{B(T-t)}\cdot) - u(T, \cdot) + Yu(T, \cdot)(T-t)\|_{L^{\infty}} \leq C(T-t)^{\frac{\beta-2}{2}}, \qquad 0 < t < T,$$

if $\beta \in]2, 2+\bar{\alpha}[$.

Remark 3.2.4. Recall that (see Remark 1.2.7) Theorem 3.2.2 does not imply, under the assumptions therein, that the Lie derivative Yu is jointly continuous in time and space variables. However, under the additional assumption that f and a_{ij}, a_i, a are continuous on $\mathcal{S}_T$, it can be seen by Definitions 1.2.4 and 2.1.2, together with Remark 1.2.6, that Yu turns out to be continuous on S_T. In particular, u is Lie differentiable along Y everywhere on $\mathcal{S}_T$, and thus equation (2.1) is satisfied pointwise everywhere on $\mathcal{S}_T$.

3.2.2 Comparison with the literature

Global *anisotropic-to-anisotropic* Schauder estimates were proved by Lunardi [66], Lorenzi [64], Priola [81], Menozzi [17] under different assumptions on the coefficients, stronger or equivalent to ours. These estimates read as follows: if u is a solution of the Cauchy problem (3.1) then

$$\|u\|_{L^\infty_T(\mathbf{C}^{2+\alpha})} \leq C\left(\|g\|_{\mathbf{C}^{2+\alpha}} + \|f\|_{L^\infty_T(\mathbf{C}^{\alpha})}\right). \tag{3.5}$$

Estimate (3.5) is similar to (3.2) for the particular choice $\beta = 2$ and $\gamma = 0$; however, by Remark 1.2.6, estimate (3.5) is weaker than (3.2) due to the strict inclusion of intrinsic into anisotropic spaces. Moreover, the above mentioned papers assume a smooth datum, $g \in \mathbf{C}^{2+\alpha}$, missing the smoothing effect of the equation, which is well-known in the uniformly parabolic case (cf. [61]). On the other hand, the class of equations considered by Menozzi [17] allows for more general drift terms. Zhang [88] proved general estimates in the context of Besov-Holder spaces, which, at order two read as

$$\|u\|_{L^\infty_T(\mathbf{C}^{2+\alpha})} \leq C(T-t)^{-\frac{2+\alpha-\beta}{2}}\|g\|_{\mathbf{C}^{\beta}}. \tag{3.6}$$

Once more, this estimate is proved for the anisotropic Hölder norm and thus only catches the smoothing effect of the kernel with respect to the spatial variables. Also, (3.6) is proved in the case of $\mathcal{A}$ with constant coefficients.

Global intrinsic-to-intrinsic Schauder estimates were obtained by Imbert and Mouhot [52] for B as in (3.8), assuming coefficients in $\mathbf{C}^{\alpha}_{T}$. As noticed in [26], the results in [52] do not cover the case of some elementary smooth functions, for instance, $d = 1$ and $f(t, x_1, x_2) = \sin x_2$. We also mention that *interior estimates* were obtained by Manfredini [67], Di Francesco and Polidoro[24], Henderson and Snelson [43] and very recently by Polidoro, Rebucci and Stroffolini [80] for operators with Dini continuous coefficients.

Global anisotropic-to-intrinsic Schauder estimates were obtained by Biagi and Bramanti [8] under the additional assumption that the $*$-blocks in (1.7) are null and therefore $\mathcal{L}$ is homogeneous w.r.t. a suitable family of dilations: if u is a solution to $\mathcal{L}u = f$ then

$$\|\nabla_d^2 u\|_{L_T^\infty(\mathbf{C}^\alpha)} + \|Yu\|_{L_T^\infty(\mathbf{C}^\alpha)} + \|\nabla_d u\|_{\mathbf{C}_T^\alpha} + \|u\|_{\mathbf{C}_T^\alpha} \le C\left(\|u\|_{L_T^\infty} + \|f\|_{L_T^\infty(\mathbf{C}^\alpha)}\right). \tag{3.7}$$

Estimate (3.7) is weaker than (3.2) because the norm on the l.h.s. of (3.7) is smaller than the norm in $\mathbf{C}_T^{2+\alpha}$: the Hölder semi-norms $\|\nabla_d u\|_{C_{Y,T}^{1+\alpha}}$ and $\|\nabla_d^2 u\|_{C_{Y,T}^{\alpha}}$ along the direction Y are missing. Moreover, the optimal regularity for the second derivatives $\nabla_d^2 u$ is obtained only *locally*.

The closest results to ours were recently obtained by Dong and Yastrzhembskiy [26] for the Langevin operator in $\mathbb{R}^{2d+1}$, that is for B in (1.7) in the particular form

$$B = \begin{pmatrix} 0 & 0 \\ I_d & 0 \end{pmatrix}, \tag{3.8}$$

where I_d is the $d \times d$ identity matrix and for the Cauchy problem with null-terminal datum, $g = 0$. The techniques in [26] are based on a kernel free approach inspired by Campanato's ideas. Despite the proof is quite different, the estimates are very similar to ours except they miss the optimal regularity along Y of the first order derivatives $\nabla_d u$: we remark that this piece of information is crucial to guarantee the validity of Taylor formulas like (1.10). As already mentioned, the latter is a basic tool to prove probabilistic results such as the Itô formula (cf. [57]), as well as analytical (cf. [52]) and asymptotic results (cf. [73]).

3.3 Proof of Theorem 3.2.2

The proof of Theorem 3.2.2 relies on the results of Chapter 2 (see also [71]); for the readers' convenience we recall them here. A fundamental solution to the operator $\mathcal{L}$ was constructed in the form

$$p(t,x;T,y) = \mathbf{P}(t,x;T,y) + \Phi(t,x;T,y), \qquad 0 < t < T, \quad x,y \in \mathbb{R}^N, \tag{3.9}$$

where:

- the function $\mathbf{P}$ is a so-called *parametrix*, which is defined as

$$\mathbf{P}(t,x;s,y) := \Gamma^{(s,y)}(t,x;s,y), \qquad 0 < t < s < T, \quad x,y \in \mathbb{R}^N,$$

where, for $(\tau, v) \in \mathcal{S}_T$, we set

$$\Gamma^{(\tau,v)}(t,x;s,y) := \mathbf{G}\big(\mathcal{C}^{(\tau,v)}(t,s), y - e^{(s-t)B}x\big), \qquad 0 < t < s < T, \quad x, y \in \mathbb{R}^N,$$

with

$$\mathbf{G}(\mathcal{C}, z) := \frac{1}{\sqrt{(2\pi)^N \det \mathcal{C}}} e^{-\frac{1}{2}\langle \mathcal{C}^{-1} z, z\rangle},$$

and

$$\mathcal{C}^{(\tau,v)}(t,s) := \int_t^s e^{(s-r)B} A^{(\tau,v)}(r) e^{(s-r)B^*} dr,$$

$$A^{(\tau,v)}(r) := \begin{pmatrix} A_0(r, e^{(r-\tau)B}v) & 0 \\ 0 & 0 \end{pmatrix}, \qquad A_0 = (a_{ij})_{i,j=1,\dots,d}; \tag{3.10}$$

- the function $\Phi(t,x;T,y)$ is a remainder enjoying suitable regularity estimates, which are recalled in Proposition 3.3.6 below.

The strategy of our proof is to define a candidate solution to the Cauchy problem (3.1) in the form

$$u(t,x) := \int_{\mathbb{R}^N} p(t,x;T,\eta) g(\eta) d\eta - \int_t^T \int_{\mathbb{R}^N} p(t,x;\tau,\eta) f(\tau,\eta) d\eta d\tau, \qquad (t,x) \in \mathcal{S}_T, \tag{3.11}$$

and prove that u: (i) is the unique bounded solution to (3.1) and (ii) satisfies the estimate (3.2).

Note that u in (3.11) can be written as

$$u(t,x) = V_g(t,x) - V_{\mathbf{P},f}(t,x) - V_{\Phi,f}(t,x), \qquad (t,x) \in \mathcal{S}_T \tag{3.12}$$

with

$$V_g(t,x) = \int_{\mathbb{R}^N} p(t,x;T,\eta) g(\eta) d\eta,$$

and

$$V_{\mathbf{P},f}(t,x) := \int_t^T \int_{\mathbb{R}^N} \mathbf{P}(t,x;\tau,\eta) f(\tau,\eta) d\eta d\tau, \qquad V_{\Phi,f}(t,x) := \int_t^T \int_{\mathbb{R}^N} \Phi(t,x;\tau,\eta) f(\tau,\eta) d\eta d\tau.$$

We now prove our main result, Theorem 3.2.2. The proof is based on the sharp regularity estimates for V_g, $V_{\mathbf{P},f}$ and $V_{\Phi,f}$ contained in Propositions 3.3.11 and 3.3.10.

Proof of Theorem 3.2.2 (well-posedness of (3.1)*).* The uniqueness of the solution follows from standard arguments: we refer to [33] for a detailed proof.

We prove that u as defined in (3.11) is a solution to the Cauchy problem (3.1) in the sense of Definition 2.1.2. The facts that $\nabla_d u, \nabla_d^2 u$ have the required regularity on $\mathcal{S}_T$, and that u can be extended continuously to the closure of $\mathcal{S}_T$ in a way that $u(T,\cdot) \equiv g$, are straightforward consequences of the estimates of Propositions 3.3.10-3.3.11 and of the Dirac delta property of p, that is

$$\lim_{\substack{(t,x)\to(T,y)\\ t<T}} \int_{\mathbb{R}^N} p(t,x;T,\eta)\varphi(\eta)d\eta = \varphi(y), \qquad \varphi \in C_b(\mathbb{R}^N).$$

To prove that $Yu = f - \mathcal{A}u$ in the sense of Definition 1.2.3, we can consider separately, by linearity, two cases. We state here, once and for all, that all the applications of Fubini's theorem throughout this proof are justified by the estimates of Propositions 3.3.10 and 3.3.11.

Case $f \equiv 0$. We have

$$\begin{aligned} u(s,e^{(s-t)B}x) - u(t,x) &= V_g(s,e^{(s-t)B}x) - V_g(t,x) \\ &= \int_{\mathbb{R}^N} \big(p(s,e^{(s-t)B}x;T,\eta) - p(t,x;T,\eta)\big) g(\eta)d\eta \end{aligned}$$

($(Y+\mathcal{A})p(\cdot,\cdot;T,\eta) = 0$ on $\mathcal{S}_T$ in the sense of Definition 2.1.2)

$$= -\int_{\mathbb{R}^N} \int_t^s \mathcal{A}p(r,e^{(r-t)B}x;T,\eta)dr\, g(\eta)d\eta = -\int_t^s \mathcal{A}u(r,e^{(r-t)B}x)dr,$$

where, in the last equality, we employed Fubini's theorem and (3.25)-(3.26) to move the operator $\mathcal{A}$ out of the integral on $\mathbb{R}^N$.

Case $g \equiv 0$. We have

$$\begin{aligned} u(s,e^{(s-t)B}x) - u(t,x) &= \underbrace{-\int_s^T \int_{\mathbb{R}^N} \big(p(s,e^{(s-t)B}x;\tau,\eta) - p(t,x;\tau,\eta)\big) f(\tau,\eta)d\eta d\tau}_{=:I} \\ &\quad + \int_t^s \int_{\mathbb{R}^N} p(t,x;\tau,\eta) f(\tau,\eta)d\eta d\tau. \end{aligned}$$

As $(Y+\mathcal{A})p(\cdot,\cdot;\tau,\eta) = 0$ on $\mathcal{S}_\tau$ in the sense of Definition 2.1.2, for any $\tau \in]s,T[$, we have

$$p(s,e^{(s-t)B}x;\tau,\eta) - p(t,x;\tau,\eta) = -\int_t^s \mathcal{A}p(r,e^{(r-t)B}x;\tau,\eta)dr.$$

Therefore, by Fubini's theorem, we have

$$\begin{aligned} I =& \int_t^s \int_s^T \int_{\mathbb{R}^N} \mathcal{A}p(r, e^{(r-t)B}x; \tau, \eta) f(\tau, \eta) d\eta d\tau dr \\ =& \int_t^s \int_r^T \int_{\mathbb{R}^N} \mathcal{A}p(r, e^{(r-t)B}x; \tau, \eta) f(\tau, \eta) d\eta d\tau dr \\ & - \int_t^s \int_r^s \int_{\mathbb{R}^N} \mathcal{A}p(r, e^{(r-t)B}x; \tau, \eta) f(\tau, \eta) d\eta d\tau dr \\ =& I_1 + I_2. \end{aligned}$$

By moving the operator $\mathcal{A}$ out of the integral in $d\eta d\tau$, we obtain

$$I_1 = -\int_t^s \mathcal{A}u(r, e^{(r-t)B}x) dr,$$

and thus, to show $(Y + \mathcal{A})u = f$ on $\mathcal{S}_T$, we need to prove that

$$I_2 = -\int_t^s f(r, e^{(r-t)B}x) dr + \int_t^s \int_{\mathbb{R}^N} p(t, x; \tau, \eta) f(\tau, \eta) d\eta d\tau. \tag{3.13}$$

By applying Fubini's theorem we obtain

$$I_2 = \int_t^s \underbrace{\int_t^\tau \int_{\mathbb{R}^N} \mathcal{A}p(r, e^{(r-t)B}x; \tau, \eta) f(\tau, \eta) d\eta \, dr}_{=:J(\tau)} \, d\tau.$$

Now, for any $\varepsilon \in]0, \tau - t[$, we can write

$$J(\tau) = \underbrace{\int_{\mathbb{R}^N} \int_t^{\tau-\varepsilon} \mathcal{A}p(r, e^{(r-t)B}x; \tau, \eta) f(\tau, \eta) dr d\eta}_{=:J_1^\varepsilon(\tau)} + \underbrace{\int_{\tau-\varepsilon}^\tau \int_{\mathbb{R}^N} \mathcal{A}p(r, e^{(r-t)B}x; \tau, \eta) f(\tau, \eta) d\eta \, dr}_{=:J_2^\varepsilon(\tau)}.$$

As $(Y + \mathcal{A})p(\cdot, \cdot; \tau, \eta) = 0$ on $\mathcal{S}_\tau$, we obtain

$$J_1^\varepsilon(\tau) = -\int_{\mathbb{R}^N} \big(p(\tau - \varepsilon, e^{(\tau-\varepsilon-t)B}x; \tau, \eta) - p(t, x; \tau, \eta)\big) f(\tau, \eta) d\eta,$$

and since $f(\tau, \cdot)$ is bounded and continuous, the Dirac delta property of p yields

$$J_1^\varepsilon(\tau) \to -f(\tau, e^{(\tau-t)B}x) + \int_{\mathbb{R}^N} p(t, x; \tau, \eta) f(\tau, \eta) d\eta, \qquad \text{as } \varepsilon \to 0^+. \tag{3.14}$$

We finally consider J_2^ε. By the estimates of Lemma 3.3.4 and of Proposition 3.3.6, together with (3.9), we have

$$\left| \int_{\mathbb{R}^N} \mathcal{A}p(r, e^{(r-t)B}x; \tau, \eta) f(\tau, \eta) d\eta \right| \le \frac{C}{(T-\tau)^\gamma (\tau - r)^{\frac{2-\alpha}{2}}} \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)},$$

that is integrable in dr on $[t,\tau]$. Then, Lebesgue dominated convergence theorem yields

$$J_2^\varepsilon(\tau) \to 0, \qquad \text{as } \varepsilon \to 0^+.$$

This and (3.14) yield (3.13). □

Proof of Theorem 3.2.2 (estimate (3.2)*).* As u is a solution to (3.1), $Yu = f - \mathcal{A}u$ is a Lie derivative of u on $\mathcal{S}_T$. Therefore, we have

$$\|u\|_{\mathbf{C}_t^{2+\alpha}} \leq \|u\|_{L_t^\infty(\mathbf{C}^{2+\alpha})} + \|\nabla_d u\|_{\mathbf{C}_t^{1+\alpha}} + \|\mathcal{A}u\|_{L_t^\infty(\mathbf{C}^\alpha)} + \|f\|_{L_t^\infty(\mathbf{C}^\alpha)}.$$

Now we recall that, for $\alpha \in]0,1]$, the intrinsic spaces $\mathbf{C}_t^\alpha, \mathbf{C}_t^{1+\alpha}$ are exactly equivalent to those in [72]. This is a consequence of the intrinsic Taylor formula of Theorem 2.10 in the latter reference, with $n = 0,1$. In particular, we have

$$\|\nabla_d u\|_{\mathbf{C}_t^{1+\alpha}} \leq C(\|\nabla_d u\|_{L_t^\infty} + \|\nabla_d^2 u\|_{L_t^\infty} + \|\nabla_d u\|_{C_{Y,t}^{1+\alpha}} + \|\nabla_d^2 u\|_{C_{Y,t}^\alpha} + \|\nabla_d^2 u\|_{C_{\nabla_d,t}^\alpha}).$$

Therefore, the estimates of Propositions 3.3.10 and 3.3.11 yield

$$\|\nabla_d u\|_{\mathbf{C}_t^{1+\alpha}} \leq C\Big((T-t)^{-\frac{2+\alpha-\beta}{2}}\|g\|_{\mathbf{C}^\beta} + (T-t)^{-\gamma}\|f\|_{L_{T,\gamma}^\infty(\mathbf{C}^\alpha)}\Big).$$

To obtain the same estimate for $\|u\|_{L_t^\infty(\mathbf{C}^{2+\alpha})}$, it is enough to prove it for

$$\sup_{s\in]0,t[} \sup_{(x,\eta)\in\mathbb{R}^N\times\mathbb{R}^{N-d}} \frac{|u(s,x+(0,\eta)) - u(s,x)|}{|(0,\eta)|_B^{2+\alpha}},$$

which can be done by proceeding as in the proofs of (3.30) and (3.41): we omit the details for brevity. The same estimate for $\|u\|_{L_t^\infty(\mathbf{C}^{2+\alpha})}$ holds. Furthermore, by the regularity assumptions on the coefficients of $\mathcal{A}$, we have

$$\|\mathcal{A}u\|_{L_t^\infty(\mathbf{C}^\alpha)} \leq C(\|u\|_{L_t^\infty(\mathbf{C}^\alpha)} + \|\nabla_d u\|_{\mathbf{C}_t^{1+\alpha}}).$$

Finally, we have $\|f\|_{L_t^\infty(\mathbf{C}^\alpha)} \leq (T-t)^{-\gamma}\|f\|_{L_{T,\gamma}^\infty(\mathbf{C}^\alpha)}$ and thus (3.2).

□

The rest of the section is devoted to proving the regularity estimates employed in the proof of Theorem 3.2.2. Hereafter, we denote, indistinctly, by C any positive constant depending at most on $T, B, \mu, \bar{\alpha}, \alpha, \beta, \gamma$ and on the $L_T^\infty(\mathbf{C}^{\bar{\alpha}})$ norms of the coefficients of $\mathcal{A}$. We also introduce the following

Notation 3.3.1. For any $f = f(t, x; T, y)$ and $i = 1, \dots, N$, we set

$$\partial_i f(t, x; T, y) := \partial_{x_i} f(t, x; T, y),$$

and we adopt analogous notations for the higher-order derivatives. Thus, ∂_i always denotes a derivative with respect to the first set of space variables. Some caution is necessary when considering the composition of f with a given function $F = F(x)$: $\partial_i f\big(t, F(x); T, y\big)$ denotes the derivative $\partial_{z_i} f(t, z; T, y)|_{z=F(x)}$, and similarly for higher order derivatives. We also denote by e_k the k-th element of the canonical basis of $\mathbb{R}^N$.

3.3.1 Preliminaries results

We first recall the useful result [71, Lemma 3.3]:

Lemma 3.3.2. *Let $(t, y) \in \mathcal{S}_T$. Then, for any $i = 1, \dots, d$, the function $u := \partial_i \mathbf{P}(\cdot, \cdot; t, y)$ is a strong Lie solution to the equation*

$$\mathcal{A}^{(t,y)} u + Y u = -\sum_{j=1}^{d+d_1} b_{ji} \partial_j \mathbf{P}(\cdot, \cdot; t, y) \text{ on } \mathcal{S}_t,$$

in the sense of Definition 2.1.2.

In order to state the next preliminary lemma, we fix the following

Notation 3.3.3. Let $\kappa = (\kappa_1, \dots, \kappa_N) \in \mathbb{N}_0^N$ be a multi-index. Recalling (1.8), we define the *B-length* of κ as

$$[\kappa]_B := \sum_{j=0}^{q} (2j+1) \sum_{i=\bar{d}_{j-1}+1}^{\bar{d}_j} \kappa_i.$$

Moreover, we recall the definition of the *homogeneous dimension* Q in (1.9).

We have the following potential estimates, whose proof is identical to the one of [71, Proposition B.2].

Lemma 3.3.4. *For any $\kappa \in \mathbb{N}_0^N$ with $[\kappa]_B \leq 4$, we have*

$$\left| \int_{\mathbb{R}^N} \mathbf{P}(t, x; \tau, \eta) f(\tau, \eta) d\eta \right| \leq \frac{C}{(T-\tau)^{\gamma}} \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)} \tag{3.15}$$

$$\left| \int_{\mathbb{R}^N} \partial_x^\kappa \mathbf{P}(t, x; \tau, \eta) f(\tau, \eta) d\eta \right| \leq \frac{C}{(T-\tau)^{\gamma} (\tau - t)^{\frac{[\kappa]_B - \alpha}{2}}} \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)} \tag{3.16}$$

for every $0 < t < \tau < T$ and $x \in \mathbb{R}^N$.

We now recall the estimates for $\mathbf{P}$ and Φ proved in [71, Propositions 2.7-3.2 and Lemma 3.4], which are stated in terms of the Gaussian density Γ^δ defined, for any $\delta > 0$, by

$$\Gamma^\delta(t,x;\tau,y) := \mathbf{G}\big(\delta\mathcal{C}(\tau-t), y - e^{(\tau-t)B}x\big), \qquad 0 \le t < \tau \le T,\ x,y \in \mathbb{R}^N,$$

with

$$\mathcal{C}(t) = \int_0^t e^{(t-\tau)B}\begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} e^{(t-\tau)B^*}d\tau.$$

We also recall that μ is fixed in Assumption 1.1.1 as the ellipticity constant for the matrix A.

Proposition 3.3.5. *For any $i,j,k = 1,\dots,d$, we have*

$$\begin{aligned}
|\mathbf{P}(t,x;\tau,y)| &\le C\,\Gamma^{2\mu}(t,x;\tau,y),\\
|\partial_i\mathbf{P}(t,x;\tau,y)| &\le \frac{C}{(\tau-t)^{\frac{1}{2}}}\Gamma^{2\mu}(t,x;\tau,y),\\
|\partial_{ij}\mathbf{P}(t,x;\tau,y)| &\le \frac{C}{\tau-t}\Gamma^{2\mu}(t,x;\tau,y),
\end{aligned}$$

and

$$\left|\partial_i\mathbf{P}(s,e^{(s-t)B}x;\tau,y) - \partial_i\mathbf{P}(t,x;\tau,y)\right| \le C\frac{(s-t)^{\frac{1+\alpha}{2}}(\tau-t)^{\frac{Q}{2}}}{(\tau-s)^{\frac{Q+1+\alpha}{2}}}\Gamma^{2\mu}(t,x;\tau,y),$$

$$\left|\partial_{ij}\mathbf{P}(s,e^{(s-t)B}x;\tau,y) - \partial_{ij}\mathbf{P}(t,x;\tau,y)\right| \le C\frac{(s-t)^{\frac{\alpha}{2}}(\tau-t)^{\frac{Q}{2}}}{(\tau-s)^{\frac{Q+2+\alpha}{2}}}\Gamma^{2\mu}(t,x;\tau,y), \tag{3.17}$$

$$|\partial_{ij}\mathbf{P}(t,x+h\mathbf{e}_k;\tau,y) - \partial_{ij}\mathbf{P}(t,x;\tau,y)| \le C|h|^\alpha\frac{\Gamma^{2\mu}(t,x+h\mathbf{e}_k;\tau,y)+\Gamma^{2\mu}(t,x;\tau,y)}{(\tau-t)^{\frac{2+\alpha}{2}}}, \tag{3.18}$$

for any $0<t<s<\tau<T$, $x \in \mathbb{R}^N$ and $h \in \mathbb{R}$.

Proposition 3.3.6. *For any $i,j,k = 1,\dots,d$, we have*

$$|\Phi(t,x;\tau,y)| \le C(\tau-t)^{\frac{\bar\alpha}{2}}\Gamma^{2\mu}(t,x;\tau,y), \tag{3.19}$$

$$|\partial_i\Phi(t,x;\tau,y)| \le \frac{C}{(\tau-t)^{\frac{1-\bar\alpha}{2}}}\Gamma^{2\mu}(t,x;\tau,y), \tag{3.20}$$

$$|\partial_{ij}\Phi(t,x;\tau,y)| \le \frac{C}{(\tau-t)^{\frac{2-\bar\alpha}{2}}}\Gamma^{2\mu}(t,x;\tau,y), \tag{3.21}$$

and

$$\left|\partial_i\Phi(s,e^{(s-t)B}x;\tau,y) - \partial_i\Phi(t,x;\tau,y)\right| \le C\frac{(s-t)^{\frac{1+\alpha}{2}}(\tau-t)^{\frac{Q}{2}}}{(\tau-s)^{\frac{Q+2+\alpha-\bar\alpha}{2}}}\Gamma^{2\mu}(t,x;\tau,y), \tag{3.22}$$

$$\left|\partial_{ij}\Phi(s,e^{(s-t)B}x;\tau,y)-\partial_{ij}\Phi(t,x;\tau,y)\right|\leq C\frac{(s-t)^{\frac{\alpha}{2}}(\tau-t)^{\frac{Q}{2}}}{(\tau-s)^{\frac{Q+2+\alpha-\bar{\alpha}}{2}}}\Gamma^{2\mu}(t,x;\tau,y), \tag{3.23}$$

$$|\partial_{ij}\Phi(t,x+h\mathbf{e}_k;\tau,y)-\partial_{ij}\Phi(t,x;\tau,y)|\leq C|h|^{\alpha}\frac{\Gamma^{2\mu}(t,x+h\mathbf{e}_k;\tau,y)+\Gamma^{2\mu}(t,x;\tau,y)}{(\tau-t)^{\frac{2-(\alpha-\bar{\alpha})}{2}}}, \tag{3.24}$$

for any $0<t<s<\tau<T$, $x\in\mathbb{R}^N$ *and* $h\in\mathbb{R}$.

By the Lemma 3.3.4 and Proposition 3.3.6, we have the following, direct,

Proposition 3.3.7. *For* $\mathbf{F}=\mathbf{P},\Phi$, *we have*

$$\nabla_d V_{\mathbf{F},f}(t,x)=\int_t^T\int_{\mathbb{R}^N}\nabla_d\mathbf{F}(t,x;\tau,\eta)f(\tau,\eta)d\eta d\tau,$$
$$\nabla_d^2 V_{\mathbf{F},f}(t,x)=\int_t^T\int_{\mathbb{R}^N}\nabla_d^2\mathbf{F}(t,x;\tau,\eta)f(\tau,\eta)d\eta d\tau,$$

for any $0<t<T$ *and* $x\in\mathbb{R}^N$.

The following identities directly stem from the boundedness assumption on g and from Propositions 3.3.5 and 3.3.6.

Proposition 3.3.8. *We have*

$$\nabla_d V_g(t,x)=\int_{\mathbb{R}^N}\nabla_d\, p(t,x;\tau,y)g(y)dy, \tag{3.25}$$

$$\nabla_d^2 V_g(t,x)=\int_{\mathbb{R}^N}\nabla_d^2\, p(t,x;\tau,y)g(y)dy, \tag{3.26}$$

for any $0<t<\tau<T$, $x\in\mathbb{R}^N$.

In the sequel we will make use of the special functions ${}_2F_1$ and $\mathbf{B}$, which denote the Gaussian hypergeometric function and the incomplete Beta function, respectively.

Remark 3.3.9. We recall the following known properties. For any $\gamma\in[0,1)$ and $\alpha\in(0,1)$ we have:

(a) ${}_2F_1\left(\frac{\alpha-1}{2},\gamma;\frac{1+\alpha}{2};\cdot\right)$ is bounded on $[0,1/2]$;

(b) there exists $\kappa=\kappa(\alpha,\gamma)>0$ such that

$$\mathbf{B}(x,\alpha/2,1-\gamma)\leq\kappa\, x^{\alpha/2},\qquad x\in[0,1].$$

3.3.2 Hölder estimates for $V_{\mathbf{P},f}$ and $V_{\Phi,f}$

In this section we prove the following Hölder estimates for $V_{\mathbf{P},f}$ and $V_{\Phi,f}$, on which the proof of Theorem 3.2.2 relies.

Proposition 3.3.10. *For $\mathbf{F} = \mathbf{P}, \Phi$, and for any $i, j, k = 1, \dots, d$, we have*

$$|V_{\mathbf{F},f}(t,x)| \le C(T-t)^{-\gamma+1}\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}, \tag{3.27}$$

$$|\partial_i V_{\mathbf{F},f}(t,x)| \le C(T-t)^{-\gamma+\frac{1+\alpha}{2}}\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}, \tag{3.28}$$

$$|\partial_{ij} V_{\mathbf{F},f}(t,x)| \le C(T-t)^{-\gamma+\frac{\alpha}{2}}\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}, \tag{3.29}$$

and

$$|\partial_{ij} V_{\mathbf{F},f}(t,x+h\mathbf{e}_k) - \partial_{ij} V_{\mathbf{F},f}(t,x)| \le C|h|^\alpha (T-t)^{-\gamma}\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}, \tag{3.30}$$

$$|\partial_i V_{\mathbf{F},f}\big(s, e^{(s-t)B}x\big) - \partial_i V_{\mathbf{F},f}(t,x)| \le C(s-t)^{\frac{1+\alpha}{2}}(T-s)^{-\gamma}\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}, \tag{3.31}$$

$$|\partial_{ij} V_{\mathbf{F},f}\big(s, e^{(s-t)B}x\big) - \partial_{ij} V_{\mathbf{F},f}(t,x)| \le C(s-t)^{\frac{\alpha}{2}}(T-s)^{-\gamma}\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}, \tag{3.32}$$

for any $0 < t < s < T$, $x \in \mathbb{R}^N$ and $h \in \mathbb{R}$.

Proof of Proposition 3.3.10 for $\mathbf{F} = \mathbf{P}$. Estimates (3.27)-(3.28)-(3.29) are a straightforward consequence of estimates (3.15), (3.16) with $\partial^\kappa = \partial_i$, and (3.16) with $\partial^\kappa = \partial_{ij}$, respectively.

We now fix $0 < t < T$, $x \in \mathbb{R}^N$ and prove (3.30) in two separate cases.
Case $2h^2 \le T - t$. We define

$$I(\tau) := \int_{\mathbb{R}^N} \big(\partial_{ij}\mathbf{P}(t, x+h\mathbf{e}_k; s, y) - \partial_{ij}\mathbf{P}(t, x; s, y)\big) f(s,y)dy$$

so that

$$|\partial_{ij} V_{\mathbf{P},f}(t,x+h\mathbf{e}_k) - \partial_{ij} V_{\mathbf{P},f}(t,x)| = \underbrace{\int_{t+h^2}^T I(\tau)d\tau}_{=:I_1} + \underbrace{\int_t^{t+h^2} I(\tau)d\tau}_{=:I_2}.$$

We consider I_1. By the mean-value theorem, there exists a real $\bar h$ with $|\bar h| \le |h|$ such that

$$|\partial_{ij}\mathbf{P}(t,x+h\mathbf{e}_k;\tau,\eta) - \partial_{ij}\mathbf{P}(t,x;\tau,\eta)| = |h|\left|\partial_{ijk}\mathbf{P}(t,x+\bar h\mathbf{e}_k;\tau,\eta)\right|.$$

Therefore, by the estimate (3.16) with $\partial_x^\kappa = \partial_{ijk}$ we have

$$|I_1| \le C\int_{t+h^2}^T \frac{|h|}{(T-\tau)^\gamma(\tau-t)^{\frac{3-\alpha}{2}}}d\tau\ \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}. \tag{3.33}$$

Now, a direct computation yields

$$\int_{t+h^2}^{T} \frac{1}{(T-\tau)^{\gamma}(\tau-t)^{\frac{3-\alpha}{2}}} d\tau$$
$$= \frac{\Gamma_E\left(\frac{\alpha-1}{2}\right)}{(T-t)^{\gamma}} \left((T-t)^{\frac{\alpha-1}{2}} \frac{\Gamma_E(1-\gamma)}{\Gamma_E\left(\frac{1+\alpha}{2}-\gamma\right)} - |h|^{\alpha-1} {}_2F_1\left(\frac{\alpha-1}{2}, \gamma; \frac{1+\alpha}{2}; \frac{h^2}{T-t}\right) \right),$$

where Γ_E and ${}_2F_1$ denote, respectively, the Euler Gamma and the Gaussian hypergeometric functions. This, together with (3.33), $2h^2 \leq T-t$ and Remark 3.3.9-(a), proves

$$|I_1| \leq C|h|^{\alpha} \|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})} (T-t)^{-\gamma}.$$

We now consider I_2. By employing triangular inequality and estimate (3.16) with $\partial_x^{\kappa} = \partial_{ij}$ we obtain

$$|I_2| \leq \int_t^{t+h^2} \frac{C}{(T-\tau)^{\gamma}(\tau-t)^{\frac{2-\alpha}{2}}} d\tau \; \|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})}. \tag{3.34}$$

A direct computation yields

$$\int_t^{t+h^2} \frac{1}{(T-\tau)^{\gamma}(\tau-t)^{\frac{2-\alpha}{2}}} d\tau = (T-t)^{\frac{\alpha}{2}-\gamma} \mathbf{B}\left(\frac{h^2}{T-t}, \frac{\alpha}{2}, 1-\gamma\right),$$

where $\mathbf{B}$ denotes the incomplete Beta function. This, together with (3.34) and Remark 3.3.9-(b), proves

$$|I_2| \leq C|h|^{\alpha} \|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})} (T-t)^{-\gamma},$$

and thus (3.30) when $2h^2 \leq T-t$.
Case $2h^2 > T-t$. By employing the triangular inequality and estimate (3.16) with $\partial_x^{\kappa} = \partial_{ij}$ we obtain

$$\begin{aligned}
&|\partial_{ij} V_{\mathbf{P},f}(t, x+h\mathbf{e}_k) - \partial_{ij} V_{\mathbf{P},f}(t,x)| \\
&\leq C \int_t^T \frac{\|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})}}{(T-\tau)^{\gamma}(\tau-t)^{\frac{2-\alpha}{2}}} d\tau \\
&\leq C \, \|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})} (T-t)^{\alpha/2-\gamma} \\
&\leq C \, |h|^{\alpha} \|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})} (T-t)^{-\gamma},
\end{aligned}$$

as $T-t \leq 2h^2$.

We now prove (3.31). By adding and subtracting, we have

$$
\begin{aligned}
&\partial_i V_{\mathbf{P},f}\big(s, e^{(s-t)B}x\big) - \partial_i V_{\mathbf{P},f}(t,x) \\
&= \int_s^T \int_{\mathbb{R}^N} \Big(\partial_i \mathbf{P}(s, e^{(s-t)B}x; \tau, \eta) - \partial_i \mathbf{P}(t, x; \tau, \eta)\Big) f(\tau,\eta) d\eta d\tau \\
&\quad - \int_t^s \int_{\mathbb{R}^N} \partial_i \mathbf{P}(t, x; \tau, \eta) f(\tau,\eta) d\eta d\tau \\
&= I + L.
\end{aligned}
$$

Estimate (3.16) with $\partial_x^\kappa = \partial_i$ yields

$$
|L| \le C \int_t^s \frac{1}{(T-\tau)^\gamma (\tau - t)^{\frac{1-\alpha}{2}}} d\tau \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)} \le C(s-t)^{\frac{1+\alpha}{2}} (T-s)^{-\gamma} \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}.
$$

We now prove

$$
|I| \le C(T-s)^{-\gamma} h^{\frac{1+\alpha}{2}} \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}. \tag{3.35}
$$

Set $h := s - t$ and consider, once more, two separate cases.

Case $2h < T - s$. We split the integral

$$
\begin{aligned}
I = &\underbrace{\int_s^{s+h} \int_{\mathbb{R}^N} \Big(\partial_i \mathbf{P}(s, e^{(s-t)B}x; \tau, \eta) - \partial_i \mathbf{P}(t, x; \tau, \eta)\Big) f(\tau,\eta) d\eta}_{=: H_1} \\
&+ \underbrace{\int_{s+h}^T \int_{\mathbb{R}^N} \Big(\partial_i \mathbf{P}(s, e^{(s-t)B}x; \tau, \eta) - \partial_i \mathbf{P}(t, x; \tau, \eta)\Big) f(\tau,\eta) d\eta}_{=: H_2}.
\end{aligned}
$$

We estimate H_1. By the triangular inequality and (3.16) with $\partial_x^\kappa = \partial_i$ we have

$$
|H_1| \le C \int_s^{s+h} \frac{1}{(T-\tau)^\gamma} \left(\frac{1}{(\tau - s)^{\frac{1-\alpha}{2}}} + \frac{1}{(\tau - t)^{\frac{1-\alpha}{2}}} \right) d\tau \, \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}
$$

(since $\tau - s < \tau - t$)

$$
\le C \int_s^{s+h} \frac{1}{(T-\tau)^\gamma (\tau - s)^{\frac{1-\alpha}{2}}} d\tau \, \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}
$$

(by a direct computation)

$$
\le C(T-s)^{-\gamma} (T-s)^{\frac{1+\alpha}{2}} B\Big(\frac{h}{T-s}, \frac{1+\alpha}{2}, 1-\gamma\Big) \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}.
$$

Therefore by Remark 3.3.9-(b) we have

$$|H_1| \le C(T-s)^{-\gamma} h^{\frac{1+\alpha}{2}} \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}. \tag{3.36}$$

We now consider H_2. By Lemma 3.3.2 and by Fubini's theorem we have

$$H_2 = -\int_{s+h}^T \int_t^s \int_{\mathbb{R}^N} \Big(\big(\mathcal{A}^{(\tau,\eta)} \partial_i \mathbf{P}\big)(r, e^{(r-t)B}x; \tau, \eta) + \sum_{j=1}^{d+d_1} b_{ji} \partial_j \mathbf{P}(r, e^{(r-t)B}x; \tau, \eta) \Big) f(\tau, \eta) d\eta dr d\tau.$$

Therefore, the estimates (3.16) with $[\kappa]_B = 3$ and the regularity assumptions on the coefficients yield

$$\begin{aligned} |H_2| &\le C\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)} \int_{s+h}^T \int_t^s \frac{1}{(T-\tau)^\gamma (\tau - r)^{\frac{3-\alpha}{2}}} dr d\tau \\ &\le Ch\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)} \int_{s+h}^T \frac{1}{(T-\tau)^\gamma (\tau-s)^{\frac{3-\alpha}{2}}} d\tau \end{aligned}$$

(by direct computation)

$$\begin{aligned} &= Ch\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)} \frac{\Gamma_E\left(\frac{\alpha-1}{2}\right)}{(T-s)^\gamma} \\ &\quad \times \left((T-s)^{\frac{\alpha-1}{2}} \frac{\Gamma_E(1-\gamma)}{\Gamma_E\left(\frac{1+\alpha}{2} - \gamma\right)} - h^{\frac{\alpha-1}{2}} {}_2F_1\left(\frac{\alpha-1}{2}, \gamma; \frac{1+\alpha}{2}; \frac{h}{T-s}\right) \right). \end{aligned}$$

(as $2h \le T-s$ and by Remark 3.3.9-(a))

$$\le C(T-s)^{-\gamma} h^{\frac{1+\alpha}{2}} \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)},$$

which, together with (3.36), yields (3.35).

Case $2h \ge T-s$. By triangular inequality, and by (3.16) with $\partial_x^\kappa = \partial_i$, we obtain

$$|I| \le C \int_s^T \frac{1}{(T-\tau)^\gamma} \left(\frac{1}{(\tau-s)^{\frac{1-\alpha}{2}}} + \frac{1}{(\tau-t)^{\frac{1-\alpha}{2}}} \right) \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)} d\tau$$

(since $\tau - s < \tau - t$)

$$\begin{aligned} &\le C \int_s^T \frac{1}{(T-\tau)^\gamma (\tau-s)^{\frac{1-\alpha}{2}}} d\tau \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)} \\ &\le C(T-s)^{\frac{1+\alpha}{2}-\gamma} \|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)} \end{aligned}$$

(as $T-s\le 2h$)

$$\le Ch^{\frac{1+\alpha}{2}}(T-s)^{-\gamma}\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)},$$

which is (3.35). The proof of (3.32) is completely analogous, and thus is omitted for brevity. □

Proof of Proposition 3.3.10 for $\mathbf{F}=\Phi$. Estimates (3.27)-(3.28)-(3.29) can be easily obtained from estimates (3.19)-(3.20)-(3.21), respectively. The details are omitted for sake of brevity.

By (3.24) we obtain

$$\begin{aligned}&|\partial_{ij}V_{\Phi,f}(t,x+h\mathbf{e}_k)-\partial_{ij}V_{\Phi,f}(t,x)|\\ &\le C\int_t^T\int_{\mathbb{R}^N}|h|^\alpha\frac{\Gamma^{2\mu}(t,x+h\mathbf{e}_k;\tau,\eta)+\Gamma^{2\mu}(t,x;\tau,\eta)}{(T-\tau)^\gamma(\tau-t)^{\frac{2-(\bar\alpha-\alpha)}{2}}}d\eta\,d\tau\,\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}\end{aligned}$$

(integrating in η)

$$\begin{aligned}&\le C|h|^\alpha\int_t^T\frac{1}{(T-\tau)^\gamma(\tau-t)^{\frac{2-(\bar\alpha-\alpha)}{2}}}d\tau\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}\\ &\le C|h|^\alpha(T-t)^{-\gamma+\frac{\bar\alpha-\alpha}{2}}\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)},\end{aligned}$$

which proves (3.30).

We now prove (3.31). By adding and subtracting, we have

$$\begin{aligned}\partial_iV_{\mathbf{\Phi},f}\big(s,e^{(s-t)B}x\big)-\partial_iV_{\mathbf{\Phi},f}(t,x)=&\int_s^T\underbrace{\int_{\mathbb{R}^N}\Big(\partial_i\Phi(s,e^{(s-t)B}x;\tau,\eta)-\partial_i\Phi(t,x;\tau,\eta)\Big)f(\tau,\eta)d\eta}_{=:I(\tau)}\,d\tau\\ &-\underbrace{\int_t^s\int_{\mathbb{R}^N}\partial_i\Phi(t,x;\tau,\eta)f(\tau,\eta)d\eta d\tau}_{=:L}.\end{aligned}$$

We first bound the first integral. By applying (3.22) in the case $s-t<\tau-s$, and (3.20) in the case $s-t\ge\tau-s$, we obtain

$$I(\tau)\le C\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}(s-t)^{\frac{1+\alpha}{2}}(T-\tau)^{-\gamma}(\tau-s)^{-1+\frac{\bar\alpha-\alpha}{2}},$$

which yields

$$\Big|\int_s^T I(\tau)d\tau\Big|\le C\|f\|_{L^\infty_{T,\gamma}(\mathbf{C}^\alpha)}(T-s)^{-\gamma+\frac{\bar\alpha-\alpha}{2}}(s-t)^{\frac{1+\alpha}{2}}. \tag{3.37}$$

As for L, estimate (3.20) yields

$$|L| \le C \int_t^s \frac{1}{(T-\tau)^{\gamma}(\tau-t)^{\frac{1-\bar{\alpha}}{2}}} \int_{\mathbb{R}^N} \Gamma^{2\mu}(t,x;\tau,\eta) d\eta d\tau \|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})}$$

(by integrating in η, and since $T-s \le T-\tau$)

$$\le C(T-s)^{-\gamma} \int_t^s \frac{1}{(\tau-t)^{\frac{1-\bar{\alpha}}{2}}} d\tau \|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})} \le C(T-s)^{-\gamma+\frac{\bar{\alpha}-\alpha}{2}}(s-t)^{\frac{1+\alpha}{2}} \|f\|_{L^{\infty}_{T,\gamma}(\mathbf{C}^{\alpha})},$$

which, together with (3.37), proves (3.31).

The proof of (3.32) is completely analogous, by employing (3.21)-(3.23) in place of (3.20)-(3.22). □

3.3.3 Hölder estimates for V_g

In this section we prove the following Hölder estimates for V_g, on which the proof of Theorem 3.2.2 relies.

Proposition 3.3.11. *For any $i,j,k = 1,\dots,d$, we have*

$$|V_g(t,x)| \le C\|g\|_{\mathbf{C}^{\beta}}, \tag{3.38}$$

$$|\partial_i V_g(t,x)| \le C(T-t)^{-\frac{(1-\beta)\vee 0}{2}} \|g\|_{\mathbf{C}^{\beta}}, \tag{3.39}$$

$$|\partial_{ij} V_g(t,x)| \le C(T-t)^{-\frac{(2-\beta)\vee 0}{2}} \|g\|_{\mathbf{C}^{\beta}}, \tag{3.40}$$

and

$$|\partial_{ij} V_g(t,x+h\mathbf{e}_k) - \partial_{ij} V_g(t,x)| \le C|h|^{\alpha}(T-t)^{-\frac{2+\alpha-\beta}{2}} \|g\|_{\mathbf{C}^{\beta}}, \tag{3.41}$$

$$|\partial_i V_g\big(s, e^{(s-t)B}x\big) - \partial_i V_g(t,x)| \le C(s-t)^{\frac{1+\alpha}{2}}(T-s)^{-\frac{2+\alpha-\beta}{2}} \|g\|_{\mathbf{C}^{\beta}}, \tag{3.42}$$

$$|\partial_{ij} V_g\big(s, e^{(s-t)B}x\big) - \partial_{ij} V_g(t,x)| \le C(s-t)^{\frac{\alpha}{2}}(T-s)^{-\frac{2+\alpha-\beta}{2}} \|g\|_{\mathbf{C}^{\beta}}, \tag{3.43}$$

for any $0<t<s<T$, $x \in \mathbb{R}^N$ and $h \in \mathbb{R}$.

Recall that, by assumption, $g \in \mathbf{C}^{\beta}$ with $\beta \in [0, 2+\alpha]$. Therefore, for any fixed $\bar{x} \in \mathbb{R}^N$, the following truncated Taylor polynomials are well defined

$$\widetilde{T}_{\beta,\bar{x}} g(y) := \begin{cases} \psi(y-\bar{x}) g(\bar{x}), & \text{if } \beta \in]0,1], \\ \psi(y-\bar{x}) \Big(g(\bar{x}) + \sum_{i=1}^{d} \partial_i g(\bar{x})(y_i - \bar{x}_i) \Big), & \text{if } \beta \in]1,2], \\ \psi(y-\bar{x}) \Big(g(\bar{x}) + \sum_{i=1}^{d} \partial_i g(\bar{x})(y_i - \bar{x}_i) + \frac{1}{2} \sum_{i,j=1}^{d} \partial_{ij} g(\bar{x})(y_i - \bar{x}_i)(y_j - \bar{x}_j) \Big), & \text{if } \beta \in]2,3], \end{cases}$$

for any $y \in \mathbb{R}^N$, where ψ is a cut-off function such that $\psi(x) = 1$ if $|x|_B \le 1$. We also set the remainder

$$R^g_{\beta,\bar{x}}(y) := g(y) - \widetilde{T}_{\beta,\bar{x}}g(y), \qquad \bar{x}, y \in \mathbb{R}^N.$$

The next lemma is a straightforward consequence of the definition of anisotropic norm $\|\cdot\|_{\mathbf{C}^\beta}$.

Lemma 3.3.12 (Taylor formula). *We have*

$$|R^g_{\beta,\bar{x}}(y)| \le C\|g\|_{\mathbf{C}^\beta}|y - \bar{x}|^\beta_B, \qquad y, \bar{x} \in \mathbb{R}^N.$$

Remark 3.3.13. For any $\bar{x} \in \mathbb{R}^N$, we have

$$\|\widetilde{T}_{\beta,\bar{x}}g\|_{\mathbf{C}^{2+\alpha}_T} + \|(Y + \mathcal{A})\widetilde{T}_{\beta,\bar{x}}g\|_{L^\infty_T(\mathbf{C}^\alpha)} \le C\|g\|_{\mathbf{C}^\beta}. \tag{3.44}$$

Now set $u_{\bar{x}}(t,x) := V_g(t,x) - \widetilde{T}_{\beta,\bar{x}}g(x)$ so that

$$V_g(t,x) = \widetilde{T}_{\beta,\bar{x}}g(x) + u_{\bar{x}}(t,x).$$

By the first part of Theorem 3.2.2, V_g is the solution to the Cauchy problem 3.1 with $f = 0$. Therefore, it is easy to check that $u_{\bar{x}}$ is the solution to the Cauchy problem 3.1 with

$$f = -(Y + \mathcal{A})\widetilde{T}_{\beta,\bar{x}}g$$

and terminal datum given by $R^g_{\bar{x},\beta}$. In particular (see (3.12)), $u_{\bar{x}}$ is of the form

$$u_{\bar{x}} = V_{R^g_{\bar{x},\beta}} - V_{\mathbf{P},f} - V_{\Phi,f}.$$

Therefore, owing to Proposition 3.3.10 and to (3.44), in order to prove the inequalities in Proposition 3.3.11, it is sufficient to prove them for $V_{R^g_{\bar{x},\beta}}$, with an arbitrary $\bar{x} \in \mathbb{R}^N$.

Proof of Proposition 3.3.11. Let $0 < t < s < T$, $x \in \mathbb{R}^N$ and $h \in \mathbb{R}$ be fixed. For brevity, we only prove (3.40), (3.41) and (3.43), the proofs of (3.38), (3.39) and (3.42) begin simpler.

We first prove (3.40). By Remark 3.3.13, it is enough to prove the estimate for $V_{R^g_{\xi,\beta}}$ with $\xi := e^{(T-t)B}x$. By (3.26), which remains true for $V_{R^g_{\xi,\beta}}$, and Lemma 3.3.12, we obtain

$$|\partial_{ij}V_{R^g_{\xi,\beta}}(t,x)| \le C\|g\|_{\mathbf{C}^\beta}\int_{\mathbb{R}^N}\frac{\Gamma^{2\mu}(t,x;T,\eta)}{T-t}|\eta - \xi|^\beta_B d\eta$$

(by the Gaussian estimates in [71, Lemma A.6], and integrating in η)

$$\leq C\|g\|_{\mathbf{C}^\beta}(T-t)^{-\frac{2-\beta}{2}}.$$

We now prove (3.41) by considering two separate cases.
Case $h^2 \geq T-t$. By employing triangular inequality, we have

$$|\partial_{ij}V_g(t,x+h\mathbf{e}_k) - \partial_{ij}V_g(t,x)| \leq |\partial_{ij}V_g(t,x+h\mathbf{e}_k)| + |\partial_{ij}V_g(t,x)|$$

(by (3.40) that we have just proved)

$$\leq \frac{C}{(T-t)^{\frac{2-\beta}{2}}}\|g\|_{\mathbf{C}^\beta} \leq C\frac{|h|^\alpha}{(T-t)^{\frac{2-\beta+\alpha}{2}}}\|g\|_{\mathbf{C}^\beta},$$

where we used $T-t \leq h^2$ in the last inequality.
Case $h^2 < T-t$. Once more, by Remark 3.3.13, it is enough to prove the estimate for $V_{R^g_{\xi,\beta}}$ with $\xi := e^{(T-t)B}x$. First we note that, setting $\xi' := e^{(T-t)B}(x+h\mathbf{e}_k)$, [71, Lemma A.4] yields

$$|\xi'-\xi|_B = |he^{(T-t)B}\mathbf{e}_k|_B \leq C\sum_{j=0}^{q}|h(T-t)^j|^{\frac{1}{2j+1}} \leq C(T-t)^{\frac{1}{2}},$$

and thus

$$\Gamma^{2\mu}(t,x+h\mathbf{e}_k;T,\eta)|\eta-\xi|_B^\beta \leq \Gamma^{2\mu}(t,x+h\mathbf{e}_k;T,\eta)(|\eta-\xi'|_B^\beta + |\xi'-\xi|_B^\beta)$$

(by [71, Lemma A.6])

$$\begin{aligned}&\leq C(T-t)^{\frac{\beta}{2}}\Gamma^{3\mu}(t,x+h\mathbf{e}_k;T,\eta) + \Gamma^{2\mu}(t,x+h\mathbf{e}_k;T,\eta)|\xi'-\xi|_B^\beta \\ &\leq C(T-t)^{\frac{\beta}{2}}\Gamma^{3\mu}(t,x+h\mathbf{e}_k;T,\eta). \end{aligned} \tag{3.45}$$

By Proposition 3.3.8 and Lemma 3.3.12, we obtain

$$\begin{aligned}&|\partial_{ij}V_{R^g_{\xi,\beta}}(t,x+h\mathbf{e}_k) - \partial_{ij}V_{R^g_{\xi,\beta}}(t,x)| \\ \leq{}& C\|g\|_{\mathbf{C}^\beta}\int_{\mathbb{R}^N}|\partial_{ij}p(t,x+h\mathbf{e}_k;T,\eta) - \partial_{ij}p(t,x;T,\eta)||\eta-\xi|_B^\beta d\eta\end{aligned}$$

(by (3.18)-(3.24))

$$\leq C\|g\|_{\mathbf{C}^\beta}|h|^\alpha\int_{\mathbb{R}^N}\frac{\Gamma^{2\mu}(t,x+h\mathbf{e}_k;T,\eta)+\Gamma^{2\mu}(t,x;T,\eta)}{(T-t)^{\frac{2+\alpha}{2}}}|\eta-\xi|_B^\beta d\eta$$

(by the estimate (3.45) and, once more, the Gaussian estimates in [71, Lemma A.6])

$$\leq C\|g\|_{\mathbf{C}^\beta}|h|^\alpha \int_{\mathbb{R}^N} \frac{\Gamma^{3\mu}(t, x + h\mathbf{e}_k; T, \eta) + \Gamma^{2\mu}(t, x; T, \eta)}{(T-t)^{\frac{2+\alpha-\beta}{2}}} d\eta$$

(integrating in η)

$$\leq C\|g\|_{\mathbf{C}^\beta}|h|^\alpha (T-t)^{\frac{-2-\alpha+\beta}{2}}.$$

We finally prove (3.42) by considering two separate cases.
Case $s - t \geq T - s$. By employing triangular inequality, we have

$$\begin{aligned}|\partial_{ij} V_g\big(s, e^{(s-t)B}x\big) - \partial_{ij} V_g(t,x)|& \\ &\leq |\partial_{ij} V_g\big(s, e^{(s-t)B}x\big)| + |\partial_{ij} V_g(t,x)|\end{aligned}$$

(by (3.40))

$$\begin{aligned}&\leq C\|g\|_{\mathbf{C}^\beta}\big((T-t)^{\frac{-2+\beta}{2}} + (T-s)^{\frac{-2+\beta}{2}}\big) \\ &\leq C\|g\|_{\mathbf{C}^\beta}(s-t)^{\frac{\alpha}{2}}(T-t)^{\frac{-2-\alpha+\beta}{2}},\end{aligned}$$

where we used $T - s \leq s - t \leq T - t$ in the last inequality.
Case $s - t < T - s$. Once more, by Remark 3.3.13, it is enough to prove the estimate for $V_{R^g_{\xi,\beta}}$ with $\xi := e^{(T-t)B}x$. By Proposition 3.3.8 and Lemma 3.3.12 we obtain

$$\begin{aligned}&|\partial_{ij} V_{R^g_{\xi,\beta}}\big(s, e^{(s-t)B}x\big) - \partial_{ij} V_{R^g_{\xi,\beta}}(t,x)| \\ &\leq C\|g\|_{\mathbf{C}^\beta} \int_{\mathbb{R}^N} |\partial_{ij} p(s, e^{(s-t)B}x; T, \eta) - \partial_{ij} p(t, x; T, \eta)||\eta - \xi|_B^\beta d\eta\end{aligned}$$

(by (3.17)-(3.23), and since $s - t < T - s$)

$$\leq C\|g\|_{\mathbf{C}^\beta} \frac{(s-t)^{\frac{\alpha}{2}}}{(T-s)^{1+\frac{\alpha}{2}}} \int_{\mathbb{R}^N} \Gamma^{2\mu}(t, x; T, \eta)|\eta - \xi|_B^\beta d\eta$$

(by the Gaussian estimates in [71, Lemma A.6], and integrating in $d\eta$)

$$\leq C\|g\|_{\mathbf{C}^\beta}(T-s)^{\frac{-1-\alpha+\beta}{2}}(s-t)^{\frac{\alpha}{2}}.$$

□

Chapter 4

A kinetic Nash inequality and precise boundary behavior of the kinetic Fokker-Planck equation

In this chapter, we prove a kinetic Nash type inequality and adapt it to a new functional inequality for functions in a kinetic Sobolev space with absorbing boundary conditions on the half-space. As an application, we address the boundary behavior of the kinetic Fokker-Planck equations in the half-space. Our main result is the sharp regularity of the solution at the absorbing boundary and grazing set.

Based on a joint work ([42]) with Profs. Christopher Henderson and Weinan Wang.

4.1 Introduction

4.1.1 The equation

We study the homogeneous kinetic Fokker-Planck equation in the half-space with absorbing boundary conditions:

$$\begin{cases} (\partial_t + v \cdot \nabla_x) f = \Delta_v f & \text{in } \mathbb{R}_+ \times \mathbb{H}^d \times \mathbb{R}^d, \\ f(t,x,v) = 0 & \text{on } \mathbb{R}_+ \times \gamma_-, \\ f(0,\cdot,\cdot) = f_{\text{in}} & \text{in } \mathbb{H}^d \times \mathbb{R}^d, \end{cases} \tag{4.1}$$

where we let $\mathbb{R}_+ = (0, \infty)$, $\mathbb{R}_- = (-\infty, 0)$,

$$\mathbb{H}^d = \left\{ (x_1, \ldots, x_d) \in \mathbb{R}^d : x_1 > 0 \right\}, \quad \text{and} \quad \gamma_\pm = \{(x, v) : x_1 = 0, \mp v_1 > 0\}.$$

We assume that f_{in} is a nonnegative, measurable function that is an element of a certain weighted L^1-space. We refer to γ_- as the *incoming* portion of the boundary and γ_+ as the *outgoing* portion of the boundary. The sign convention may appear strange above, but we follow the standard notation in the general case: the minus sign corresponds to the negativity of $v \cdot \eta_x$, where η_x is the outward pointing unit normal on the physical space boundary. In our case $\eta_x = (-1, 0, \cdots, 0)$. The set where $v \cdot \eta_x = 0$ is called the "grazing set." In our case this is when $x_1 = 0 = v_1$.

4.1.2 Informal discussion of the main results

Our goal is to understand the precise boundary behavior of (4.1). In particular, we are interested in the sharp regularity on γ_-. We note that the interior regularity is quite well-understood; see [56] for the homogeneous equation and [1, 3, 5, 6, 8, 11, 24, 26, 27, 37, 39, 43, 44, 52, 60, 63, 67, 71] for more recent results with varying degrees of inhomogeneity. More generally, we refer to the review [4]. Let us note that the literature is quite large, so the above is unfortunately only a small sample of related works. Briefly, though, the major source of difficulty for (4.1) is the lack of diffusion in x. Instead, one must use "hypoellipticity" to import the v-regularity (generated by the Δ_v term on the right hand side) to (t, x)-regularity via the transport term $\partial_t + v \cdot \nabla_x$.

To illustrate the boundary regularity, let us briefly introduce a (nontrivial) steady solution to (4.1). As it is convenient to introduce a steady solution to the adjoint problem at the same time, we do so here. These solutions are

$$\begin{cases} v \cdot \nabla_x \varphi = \Delta_v \varphi & \text{in } \mathbb{H}^d \times \mathbb{R}^d, \\ \varphi = 0 & \text{on } \gamma_-, \end{cases} \quad \text{and} \quad \begin{cases} -v \cdot \nabla_x \widetilde{\varphi} = \Delta_v \widetilde{\varphi} & \text{in } \mathbb{H}^d \times \mathbb{R}^d, \\ \widetilde{\varphi} = 0 & \text{on } \gamma_+. \end{cases}$$

It is easy to see that, with an abuse of notation,

$$\varphi(x, v) = \varphi(x_1, v_1) = \widetilde{\varphi}(x_1, -v_1).$$

Following [38, Lemma 2.1], we have the asymptotics of φ, and, thus, also $\widetilde{\varphi}$, given by

$$\varphi(x,v) \approx \begin{cases} \frac{x_1}{v_1^{5/2}} \exp\left\{-\frac{v_1^3}{9x_1}\right\} & \text{if } 0 \le x_1 \le v_1^3, \\ x_1^{1/6} & \text{if } x_1 \ge |v_1|^3, \\ \sqrt{|v_1|} & \text{if } 0 \le x_1 \le -v_1^3. \end{cases} \tag{4.2}$$

Given this, it is natural to expect that the behavior f is, roughly, exponentially small as $x_1 \to 0$ with $v_1 > 0$ and $C_x^{1/6} C_v^{1/2}$ as $(x_1, v_1) \to (0,0)$. This aligns with what is well-understood about kinetic equations: the bottleneck to regularity occurs at the "grazing set."

Our goal is to make this precise by both identifying *exactly* the behavior conjectured in the previous paragraph and understanding the norms that control f near the boundary. Our approach is to develop a kinetic boundary Nash inequality that allows for an $L^1_w \to L^2$ estimate, where "w" stands for "weighted." By using adjointness, we get then an $L^2 \to L^\infty_w$ estimate. In analogy with the heat equation, one expects

$$f(t,x,v) \lesssim \frac{\|\widetilde{\varphi} f_{\text{in}}\|_{L^1}}{t^{2d+1/2}} \varphi(x,v), \tag{4.3}$$

where the power of t follows by scaling arguments and the $\widetilde{\varphi}$ appears because $\|\widetilde{\varphi} f(t)\|_{L^1}$ is a conserved quantity. To dwell on the last point a moment longer, observe that

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{H}^d \times \mathbb{R}^d} f(t,x,v) \widetilde{\varphi}(x,v)\, dx\, dv &= \int_{\mathbb{H}^d \times \mathbb{R}^d} \left[(\Delta_v - v \cdot \nabla_x) f \right] \widetilde{\varphi}\, dx\, dv \\ &= \int_{\mathbb{H}^d \times \mathbb{R}^d} f \left(\Delta_v + v \cdot \nabla_x \right) \widetilde{\varphi}\, dx\, dv = 0. \end{aligned} \tag{4.4}$$

This approach to (4.3) is outlined in greater detail in Section 4.2. Using standard interior estimates along with (4.3), one can easily show that f is $C^{1/2}_{\text{kin}} \approx C_t^{1/3} C_x^{1/6} C_v^{1/2}$ up to the boundary and smooth in the interior.

Actually, (4.3) does *not* hold! Roughly, if f_{in} is supported where $\widetilde{\varphi}$ is exponentially small, that is, $v_1 \ll -1$ and $-v_1 \ll x_1 \ll -v_1^3$, the right hand side of (4.3) will be exponentially small. On the other hand, $f(1,x,v)$ will remain constant order in $v_1 + \text{supp}(f_{\text{in}})$; that is, the set obtained by applying transport to the support of $\text{supp}(f_{\text{in}})$. This behavior is clearly not consistent with (4.3). Roughly, this is related to the fact that $\widetilde{\varphi}$ "feels" infinite time scales while f only "feels" the time interval $[0,t]$. This is important here (and not for the heat equation) because transport does not (locally in v) have infinite speed of propagation while diffusion does. Regardless, (4.3) is a good indication of our main result Theorem 4.1.1 and how the proof proceeds.

In the process of proving our boundary Nash inequality, we develop a whole-space Nash inequality (Theorem 4.1.2). This easily yields the sharp time decay estimate

$$f(t,x,v) \lesssim \frac{\|f_{\text{in}}\|_{L^1}}{t^{2d}}, \tag{4.5}$$

for solutions of (4.1) posed on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d$. Actually, one can easily include a uniformly elliptic (rough) diffusion matrix in (4.1) in the arguments deriving (4.5). This is the content of Corollary 4.1.3. It is interesting to note that estimates of this form, suitably weighted, have been used in the parabolic setting to obtain Harnack inequalities and regularity in the classic work of Fabes and Stroock [30]. It is possible that this could provide a new method to understand estimates of the fundamental solution. See [58] for a related approach based on a kinetic Sobolev inequality and Moser's iteration.

4.1.3 Precise statements of main results: boundary behavior on the half-space

Theorem 4.1.1. *Suppose that f solves* (4.1)*. There is a constant $\alpha > 0$ and a nonnegative smooth function μ bounded by 1, satisfying*

$$\mu(t,x,v) = \mu(t,x_1,v_1) \approx \begin{cases} 0 & \text{if } v_1 < \alpha\sqrt{t} \text{ or } x_1 \geq \frac{|v_1|^3}{\alpha}, \\ 1 & \text{if } \alpha t v_1 \leq x_1 \leq v_1^3 \alpha, \\ e^{-\frac{\alpha t v_1}{x_1}} & \text{if } v_1 \geq \alpha\sqrt{t} \text{ and } x_1 \leq \alpha t v_1, \end{cases}$$

such that f may be decomposed as

$$f(t,x,v) = \varphi(x,v) h_1(t,x,v) + t^{1/4} \mu(x,v) h_2(x,v),$$

where, for $i = 1, 2$,

$$\|h_i(t,\cdot,\cdot)\|_{L^\infty(\mathbb{H}^d \times \mathbb{R}^d)} \lesssim \frac{1}{t^{2d+1/2}} \Big(\int f_{\text{in}} \widetilde{\varphi} \, dx \, dv + t^{1/4} \int f_{\text{in}} \widetilde{\mu} \, dx \, dv \Big), \tag{4.6}$$

where $\widetilde{\mu}(t,x_1,v_1) = \mu(t,x_1,-v_1)$ (see Section 4.1.6).

Theorem 4.1.1 is quite a bit to digest, so let us discuss it briefly. First, μ is defined in Lemma 4.4.3 (note: $\widetilde{\mu}(t,x,v) = \mu(t,x,-v)$ in Lemma 4.4.3).

Second, let us consider the simple case where f_{in} is compactly supported. Then, for t sufficiently large, the right hand side of (4.6) reduces to

$$\frac{1}{t^{2d+1/2}} \int f_{\text{in}} \widetilde{\varphi} \, dx \, dv.$$

Let us also only consider here the case $v_1 \leq \alpha\sqrt{t}$.

In this case, we see the following behavior near the "grazing set" $x_1 = v_1 = 0$: if $0 < x_1, |v_1| \ll 1$,

$$f(t,x,v) = \varphi(x,v) h_1(x,v) \lesssim \frac{\varphi(x,v)}{t^{2d+1/2}}.$$

Using (4.2), we see precisely the $C_x^{1/6} C_v^{1/2}$-regularity at $(t,0,0)$.

Next, consider the behavior near γ_-: fix any $v_1 \in (0, \alpha\sqrt{t})$ and take $0 < x_1 \ll 1$. Similarly to the above, we find

$$f(t,x,v) \lesssim \frac{\varphi(x,v)}{t^{2d+1/2}} \approx \frac{x_1}{v_1^{5/2} t^{2d+1/2}} e^{-\frac{v_1^3}{9x_1}}.$$

In other words, we recover a precise form of the super-polynomial decay observed by Silvestre in [82]. It should be noted that Silvestre considers a much more irregular model than (4.1).

The case when $v_1 \geq \alpha\sqrt{t}$ is essentially the same, although with the addition of an exponentially decaying (in v_1/x_1) term due to μ. Thus, just as in the previous case, we see "fast" decay in v_1/x_1.

As we mentioned above, the bottleneck to regularity up to the boundary is precisely in understanding the decay of f as $x_1 \to 0$. As such, it is straightforward to use interior regularity estimates, suitably scaled, to deduce that

$$f \in C_{\text{kin}}^{1/2} \approx C_t^{1/2} C_x^{1/6} C_v^{1/2}$$

from Theorem 4.1.1; see [48, 49] for one approach to this. We omit the details. Since it is not the main focus of this work, we also do not clarify precisely the spaces C_{kin}^{α} beyond the rough statement above.

Finally, let us discuss the meaning and necessity of the μ and $\widetilde{\mu}$ terms. As referenced in the discussion of (4.3), they arise due to the "isolated" region

$$I_t = \{(x,v) : v_1 \leq -O(\sqrt{t}), O(t) v_1 \leq x_1 \leq v_1^2\}. \tag{4.7}$$

This set contains particles that are too far from the outgoing boundary γ_+ to travel there by transport in time t and are too far from the incoming boundary γ_- to have made it there following transport for time t and then making "jump" in velocity of size $O(\sqrt{t})$. The "allowed" jump size is determined by scaling, although it comes up in more concrete ways in our arguments.

Given this isolation, one expects the L^1-norm of f on I_t to be roughly constant for times $[0, t]$. From a microscopic point of view, this says that the density of particles in I_t is roughly constant. Intuitively, particles can leave I_t in two ways. First, a particle can make a velocity jump, leaving I_t through the top. Here $\widetilde{\varphi}$ is "large" and we can control this quantity with a term of the form

$$\frac{1}{t^{1/4}} \int f \widetilde{\varphi} \, dx \, dv = \frac{1}{t^{1/4}} \int f_{\text{in}} \widetilde{\varphi} \, dx \, dv$$

(recall (4.4)). Let us note that the time scaling is not obvious at this point. Second, a particle can follow transport and leave I_t through the *left* (because $v_1 < 0$). This is accounted for by the exponential part of $\widetilde{\mu}$, which is the appropriate density for these dynamics. Indeed,

$$(\partial_t - v \cdot \nabla_x - \Delta_v) e^{\frac{\alpha t v_1}{x_1}} \leq 0$$

for $x_1 \leq -O(t) v_1$ and $v_1 \leq -O(\sqrt{t})$.

Previous results

The closest works to ours are those of Hwang, Jang, and Velázquez [49] and Hwang, Jang, and Jung [48]; see also [50]. In these remarkable works, the authors prove many results, the most relevant to the current work being the $C_x^{\alpha/3} C_v^{\alpha}$-regularity of f for any $\alpha < 1/2$ given $f_{\text{in}} \in L^1 \cap L^\infty$. They prove that the decay rate at the boundary controls the regularity. To understand the decay rate, they construct highly nontrivial supersolutions by a clever change-of-variables and a careful patching of special functions. Our approach is quite different than their comparison principle based one, and one advantage is that we are able to identify the precise regularity, time decay, and controlling quantities (the $L^1_{\widetilde{\varphi}}$ and $L^1_{\widetilde{\mu}}$-norms of f_{in}) of the boundary behavior.

A more general approach is given by the De Giorgi methods of Silvestre [82] and Zhu [89]. Allowing rough coefficients in (4.1), these works obtain C^α_{kin} estimates of f, where α depends on the bounds of the coefficients. Silvestre also observed that, as $(x, v) \to (0, v_+)$ with $v_+ > 0$, $f(t, x, v) \lesssim x^p$ for any $p > 0$. As discussed above, we obtain a precise version of this. We also mention the recent preprint [47].

Let us finally note that hypocoercivity is another approach to overcoming the lack of diffusion in x for kinetic equations. We point out Villani's classic memoir [86] for a discussion of this topic; however, this area remains quite active. See, for example, [9, 10, 13]. That approach is quite different from our own.

Generalizations

It is clear that, for a general convex domain Ω_x, our results immediately give, via the comparison principle, the upper bound in Theorem 4.1.1 when (4.1) is posed on $\mathbb{R}_+ \times \Omega_x \times \mathbb{R}^d_v$. One need only rotate and translate $\{0\} \times \mathbb{R}^{d-1}$ to be a supporting hyperplane of $\partial\Omega_x$.

A more interesting question is how to generalize the results to the case of a general nonconvex domain Ω or the case with nonconstant coefficients

$$(\partial_t + v \cdot \nabla_x) f = \nabla_v (a \nabla_v f) + (\text{lower order terms}). \tag{4.8}$$

Let us focus on the latter as the former is, in some sense, a subcase of the after applying a suitable boundary flattening change of coordinates.

If $a \equiv \mathrm{Id}$, then our results above are immediately applicable to obtain $x_1^{1/6}$ and $v_1^{1/2}$ decay near $x_1 = 0 = v_1$. The only difference is that the lower order terms may cause norm growth, so that the t^{2d+1} term in the numerator of Theorem 4.1.1 may be changed.

When $a \neq \mathrm{Id}$ and a is sufficiently smooth, a change of variables and a rescaling takes a to the identity plus a small perturbation, locally. This is a typical technique in the proof of Schauder estimates (see, e.g., [44, Section 2.2]). In principle, one should be able to use this to recover the $x_1^{1/6}$ and $v_1^{1/2}$ decay estimates in Theorem 4.1.1.

When a is "rough," one does not expect the $x_1^{1/6}$ and $v_1^{1/2}$ decay to hold by analogy with (divergence form) elliptic equations. In this case, the results of Silvestre [82] and Zhu [89] are likely the best one can hope for: C^α_{kin}-regularity up to the boundary with α depending on the ellipticity bounds of a.

Let us point out that an advantage to our approach is the boundary behavior of *generic* solutions reduces to understanding the boundary behavior of a *single* solution to each of the equation (4.1) and the adjoint equation (4.10). Here, we use the steady solution; however, significantly less is actually required. Indeed, we only use mild control of the asymptotic growth of $\widetilde{\varphi}$ in certain regimes (e.g., Lemma 4.4.2) and that the growth of

$$\int f(t,x,v)\widetilde{\varphi}(x,v)\, dx\, dv$$

is controlled in time. Thus, in the general case (4.8), we need only find g with the appropriate boundary behavior and asymptotic growth in $\mathcal{N}_R$ such that

$$\int f(t,x,v)g(t,x,v)\,dx\,dv$$

(at most) grows in a controlled way. This last requirement is true of any function g such that

$$(\partial_t + v\cdot\nabla_x + \nabla_v \cdot a\nabla_v)g \lesssim g.$$

4.1.4 Precise statements of main results: Nash inequalities and the whole space case

As we discuss in Section 4.2, we obtain the main functional inequality (Lemma 4.4.1) for Theorem 4.1.1 by interpolating between boundary Poincaré-type inequalities and a localized Nash inequality. The localized Nash inequality may be of independent interest, so we state it here. Let us note that the kinetic notation $\delta_\cdot$, $\cdot^{-1}$, and H^1_{kin} are defined in Section 4.3.

Theorem 4.1.2. *Fix $s_0 > 0$ and sets $\Omega_1, \Omega_2 \subset \mathbb{R}_+ \times \mathbb{R}^{2d}$ such that there is a bounded open set B with*

$$\Omega_1 \circ (\delta_s B)^{-1} \subset \Omega_2 \qquad \text{for all } s \in [0, s_0].$$

Then, for any $g \in H^1_{\text{kin}}$ and $s \in (0, s_0]$, we have

$$\|g\|^2_{L^2(\Omega_1)} \lesssim s[\![g]\!]_{H^1_{\text{kin}}(\Omega_2)}\|g\|_{L^2(\Omega_2)} + \frac{1}{s^{4d+2}}\|g\|^2_{L^1(\Omega_2)}.$$

The implied constant depends only on the choice of B and the dimension.

With this in hand, we can immediately deduce a simple time-decay estimate for the whole-space kinetic Fokker-Planck equation. This estimate is not new; one can derive it from existing results on fundamental solutions; see, e.g., [6, 58], although these proofs are quite different from our own. We only include it here because it is essentially immediate from Theorem 4.1.2. It is not our main interest in this study.

Corollary 4.1.3. *Suppose that a is a symmetric, uniformly elliptic matrix:*

$$|\xi|^2 \lesssim \xi\cdot a(t,x,v)\xi \qquad \text{for all } (t,x,v) \in \mathbb{R}_+ \times \mathbb{R}^{2d} \text{ and } \xi \in \mathbb{R}^d.$$

If f is a nonnegative solution to

$$\begin{cases} (\partial_t + v \cdot \nabla_x) f = \nabla_v \cdot (a \nabla_v f) & \text{in } \mathbb{R}_+ \times \mathbb{R}^{2d}, \\ f = f_{\mathrm{in}} & \text{on } \{0\} \times \mathbb{R}^{2d}, \end{cases} \tag{4.9}$$

then

$$f(t,x,v) \lesssim \frac{1}{t^{2d}} \int f_{\mathrm{in}} \, dx \, dv.$$

If one includes lower terms such as $b \cdot \nabla_v f + cf$ in (4.9), the bounds above will hold with (possibly) an additional exponentially growing in t factor depending only on $\|c\|_\infty$ and $\|b\|_\infty$.

Finally, we note that the well-posedness of (4.1) and (4.9) with merely weighted L^1_w initial data follows simply using ideas in [49, 89] and standard approximation schemes. By the established regularity theory, solutions will be classical in the interior (and up to the boundary in x) and continuous in time up to $t = 0$ in L^1_w. As such, we omit further discussion of this.

4.1.5 Organization of the chapter

To aid the reader, we give a discussion of the general strategy of the proof in the parabolic setting in Section 4.2. It is here that we also give an indication of the main difficulties in this chapter.

The main functional analysis and group theory setup that is appropriate for kinetic equations is given in Section 4.3.

The proof of Theorem 4.1.1 occurs in Sections 4.4 and 4.5. The former contains the proof of Theorem 4.1.1 subject to a few inequalities that are stated there. The main inequality stated in Section 4.4 (Lemma 4.4.1) relies on a decomposition of $\mathbb{H}^d \times \mathbb{R}^d$ into a "Nash" region $\mathcal{N}_R$ and two "Poincaré" regions $\mathcal{P}_R$ and $\mathcal{O}_R$. See Figure 4.1. This main inequality, proved in Section 4.5, follows by establishing a localized Nash inequality in $\mathcal{N}_R$ and Poincaré-type inequalities in $\mathcal{P}_R$ and $\mathcal{O}_R$. These proofs are also contained in Section 4.5.

The construction of μ occurs in Section 4.6, and several technical lemmas are proved in Section 4.7.

Finally, the whole space case is briefly considered in Section 4.8.

4.1.6 Notation

We use z to denote a generic point (t, x, v). When z is decorated with notation, the coordinates inherit that decoration; e.g., $z' = (t', x', v')$.

We write $A \lesssim B$ is $A \le CB$ for a constant C depending only on dimension. We write $A \approx B$ if $A \lesssim B$ and $B \lesssim A$.

In order to clearly define when we use the dynamics associated to (4.1), we reserve f for its solutions and use g (or other letters) for any generic element of H^1_{kin}.

Whenever the domain of integration is not specified, it is assumed to be in $\mathbb{H}^d \times \mathbb{R}^d$ if it is an integral with respect to $dvdx$, $\mathbb{H}^d$ if it is an integral with respect to dx, or $\mathbb{R}^d$ if it is an integral with respect to dv.

We write $v = (v_1, \overline{v})$, where $\overline{v} \in \mathbb{R}^{d-1}$. Similarly, $x = (x_1, \overline{x})$. We use the tilde to denote reflection in v:

$$\widetilde{f}(t, x, v) = f(t, x, -v).$$

This is defined similarly for functions that depend only on (x, v) or only on v. We use the star to denote taking the adjoint of an operator; that is A^* is the adjoint of an operator A. Note that this is some overlap here because the adjoint equation of (4.1) is

$$\begin{cases} (\partial_t - v \cdot \nabla_x)\widetilde{f} = \Delta_v \widetilde{f} & \text{in } \mathbb{R}_+ \times \mathbb{H}^d \times \mathbb{R}^d, \\ \widetilde{f}(t, 0, v) = 0 & \text{on } \mathbb{R}_+ \times \gamma_+, \\ \widetilde{f}(0, \cdot, \cdot) = \widetilde{f}_{\text{in}} & \text{in } \mathbb{H}^d \times \mathbb{R}^d \end{cases} \tag{4.10}$$

whose solution is $\widetilde{f}$ if f solves (4.1).

We sometimes use Y as shorthand for the transport operator:

$$Y = \partial_t + v \cdot \nabla_x.$$

While there are some downsides to this notation – it is opaque and it suppresses the dependence on v – it simplifies many expressions significantly and it follows a standard convention.

4.2 The strategy of the proof

4.2.1 Boundary behavior for the heat equation

Let us recall a simple approach to understanding the boundary behavior for the heat equation in one dimension:

$$\begin{cases} h_t = h_{xx} & \text{in } \mathbb{R}_+ \times \mathbb{R}_+, \\ h(t,0) = 0 & \text{for all } t > 0, \\ h(0,x) = h_{\text{in}}(x) & \text{for all } x > 0. \end{cases} \tag{4.11}$$

This will give the basic outline of our proof for the kinetic Fokker-Planck equation (4.1).

We observe that the equation above is formally self-adjoint and x is a steady solution to it; hence,

$$\frac{d}{dt}\int_0^\infty xh\,dx = \int_0^\infty xh_{xx}\,dx = 0. \tag{4.12}$$

Next, we notice the energy equality

$$\frac{d}{dt}\int_0^\infty h^2\,dx = -\int_0^\infty |h_x|^2\,dx. \tag{4.13}$$

In the whole space, it suffices to use the Nash inequality,

$$\Big(\int g^2\,dx\Big)^3 \lesssim \Big(\int |g_x|^2\,dx\Big)\Big(\int g\,dx\Big)^4 \qquad \text{for all } g \geq 0,$$

to control the right hand side of (4.13). However, we need to use the added information in (4.12).

To this end, we fix an arbitrary $R > 0$, apply the Poincaré inequality on $(0,R)$ and the Nash inequality on (R,∞): for any g,

$$\int_0^\infty g^2\,dx \lesssim R^2\int_0^R |g_x|^2\,dx + \Big(\int_R^\infty |g_x|^2\,dx\Big)^{1/3}\Big(\int_R^\infty g\,dx\Big)^{4/3}. \tag{4.14}$$

Here we are assuming that the Nash inequality can be localized. The usual proof using the Fourier transform does not allow this, but it is not difficult to develop a different proof that does. Then we add an x/R factor to the L^1-term to obtain

$$\int_0^\infty g^2\,dx \lesssim R^2\int_0^R |g_x|^2\,dx + \frac{1}{R^{4/3}}\Big(\int_R^\infty |g_x|^2\,dx\Big)^{1/3}\Big(\int_R^\infty xg\,dx\Big)^{4/3}. \tag{4.15}$$

Optimizing in R yields

$$\int_0^\infty g^2\,dx \lesssim \Big(\int_0^\infty |g_x|^2\,dx\Big)^{3/5}\Big(\int_0^\infty xg\,dx\Big)^{4/5}. \tag{4.16}$$

Applying (4.16) to h and folding it into (4.13), we deduce

$$\frac{d}{dt}\int_0^\infty h^2\,dx \lesssim -\frac{\big(\int h^2\,dx\big)^{5/3}}{\big(\int xh\,dx\big)^{4/3}} = -\frac{\big(\int h^2\,dx\big)^{5/3}}{\big(\int xh_{\text{in}}\,dx\big)^{4/3}}, \tag{4.17}$$

where the equality is due to (4.12). Solving this differential inequality gives us the desired $L^1_x \to L^2$ bound:

$$\Big(\int_0^\infty h(t,x)^2\,dx\Big)^{1/2} \lesssim \frac{1}{t^{3/4}}\int_0^\infty xh_{\text{in}}\,dx.$$

Letting $S_t : L^1_x \to L^2$ be the solution operator to (4.11), this translates to

$$\|S_t\|_{L^1_x\to L^2} \lesssim \frac{1}{t^{3/4}}.$$

On the other hand, the adjoint operator $S_t^* : L^2 \to L^\infty_{1/x}$ is also a solution operator to (4.11) because (4.11) is formally self-adjoint and must satisfy

$$\|S_t^*\|_{L^2\to L^\infty_{1/x}} = \|S_t\|_{L^1_x\to L^2} \lesssim \frac{1}{t^{3/4}}.$$

Hence, we have

$$\|h(t)\|_{L^\infty_{1/x}} = \|S^*_{t/2}S_{t/2}h_{\text{in}}\|_{L^\infty_{1/x}} \lesssim \frac{1}{t^{3/4}}\|S_{t/2}h_{\text{in}}\|_{L^2} \lesssim \frac{1}{t^{3/4}}\frac{1}{t^{3/4}}\|h_{\text{in}}\|_{L^1_x}.$$

In other words,

$$h(t,x) \lesssim \frac{x}{t^{3/2}}\int yh_{\text{in}}\,dy,$$

which provides the desired (sharp) boundary regularity.

4.2.2 Basic ideas in the kinetic setting

The energy equality and the H^1_{kin}-norm

Let us point out the basic changes that must occur to put the above plan into action. First, we already see a difference in the energy equality for (4.1):

$$\frac{1}{2}\frac{d}{dt}\int f^2\,dx\,dv + \int |\nabla_v f|^2\,dx\,dv + \int_{\gamma_+} |v_1| f\,d\overline{x}\,dv = 0, \tag{4.18}$$

where $x = (x_1, \overline{x})$. One might be tempted to drop the boundary term above since it has a "good" sign; however, we see below that this is not possible.

Using the definition of the H^1_{kin}-norm in (4.21), we immediately obtain, from (4.1),

$$[\![f]\!]_{H^1_{\text{kin}}([T_1,T_2]\times\mathbb{H}^d\times\mathbb{R}^d)} \approx \|\nabla_v f\|_{L^2([T_1,T_2]\times\mathbb{H}^d\times\mathbb{R}^d)}. \tag{4.19}$$

In this sense, we immediately obtain bounds on the H^1_{kin}-norm of f by integrating (4.18) in time.

At this point, we notice our first roadblock to the strategy above: the H^1_{kin}-norm involves a time integral, meaning that any inequality following from a Nash-type inequality will involve time integrals. Thus, no differential inequality, such as (4.17) is possible. This, however, is not too difficult to overcome – it essentially amounts to using the integral form of Grönwall's inequality instead of the differential form.

The Poincaré bound

Next, after determining the appropriate notion of distance, we may start to follow the decomposition in (4.14). First, we can define the set $\mathcal{P}_R$ of points (x, v) within distance R to the boundary γ_- on which we have zero boundary data. See Figure 4.1. This requires some technical care, but follows a general method of proving the Poincaré inequality by integrating Yf and $\nabla_v f$ along a path starting on γ_-. Here we are able to follow the ideas of [2] to obtain an inequality *like*

$$\|f\|_{L^2([T,T+R]\times\mathcal{P}_R)} \lesssim R\|f\|_{H^1_{\text{kin}}([T,T+O(R)]\times\mathcal{P}_{2R})}. \tag{4.20}$$

See Proposition 4.5.1 for the actual inequality.

The outgoing region

Next, by analogy with the heat equation, one might hope to have a Nash inequality on $\mathcal{P}_R^c$ and follow the step (4.15) in which the steady solution is brought into the integral up to an R factor. For this, we would need $\widetilde{\varphi} \gtrsim R^p$, for some p, on $\mathcal{P}_R^c$. In view of (4.2), $\widetilde{\varphi}$ is exponentially small when $x_1 \ll -v_1^3$. Hence, this is not immediately possible.

This leads us to the observation that many of the particles in $\mathcal{P}_R^c$ leave the domain through γ_+. Defining $\mathcal{O}_R$ to be these outgoing particles (over a time interval of size $O(R)$),

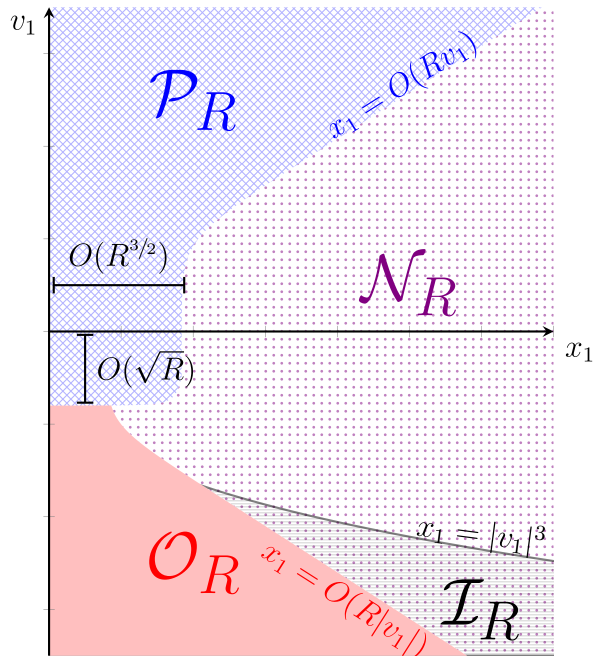

Figure 4.1: A cartoon picture of each of the key domains. The Poincaré region $\mathcal{P}_R$ is the blue crosshatched region, the Nash region $\mathcal{N}_R$ is the violet dotted region, and the outgoing region $\mathcal{O}_R$ is the red shaded region. The subregion $\mathcal{I}_R \subset \mathcal{N}_R$ is the black striped region. The rough asymptotics of the boundaries separating each region are given as well.

we can argue as in the Poincaré case to obtain a similar inequality to (4.20) that includes the boundary term from (4.18).

It is easy to see $\mathcal{O}_R$ is approximately those particles such that $v_1 \leq -O(\sqrt{R})$ and $x_1 \leq -O(R)v_1$. The latter reflects that particles can be taken to the boundary by pure transport over time $O(R)$.

The Nash region

At this point, we have no choice but to take the Nash inequality on the set $\mathcal{N}_R$ of points greater than distance R from $\gamma_\pm$. See Theorem 4.1.2 and Proposition 4.5.2. The issue is in connecting the L^1-norm that appears there with $\widetilde{\varphi}$. When $x_1 \geq -O(R)v_1^3$, we have $\widetilde{\varphi} \gtrsim R^{1/4}$, and we can argue exactly as in (4.15). This, however, is not the entirety of $\mathcal{N}_R$.

We prove the Nash inequality via an interpolation argument. It proceeds by suitably smoothing g to obtain g_ε, writing

$$\|g\|_{L^2} \le \|g_\varepsilon\|_{L^2} + \|g_\varepsilon - g\|_{L^2},$$

bounding the first term by the L^1-norm of g via a kinetic Young's inequality, and then bounding the second term by the H^1_{kin}-norm of g. This second term requires some technical care due to the H_v^{-1}-H_v^1 pairing in the H^1_{kin}-norm (see (4.21)). The result Proposition 4.5.2 follows by varying ε.

The isolated region

This leaves the isolated region $\mathcal{I}_R \subset \mathcal{N}_R$ of points $-O(R)v_1 \le x_1 \le -v_1^3$, on which $\widetilde{\varphi}$ is small. This is the region where, phenomenologically, the behavior of (4.1) is most different from (4.11). In the other regions, the computations, while technically more complicated, bore some resemblance towards their analogues for the heat equation.

This region and its role was discussed around (4.3) and (4.7). There it is pointed out that no inequality is possible purely using $\widetilde{\varphi}$. As mentioned there, we overcome this by the construction of a weight μ_R that encapsulates the movement of particles into and out of $\mathcal{I}_R$. We summarize by noting that we get an inequality like

$$\|f\|_{L^1([T,T+R]\times\mathcal{I}_R)} \lesssim R^{3/4} \int f_{\text{in}} \widetilde{\varphi}\, dx\, dv + R^{3/4} \int f_{\text{in}} \widetilde{\mu}_R\, dx\, dv.$$

See Lemma 4.4.3 and (4.32).

4.3 Kinetic functional analysis

4.3.1 The functional space H^1_{kin}

Let us define the space

$$H^1_{\text{kin},0}((T_1,T_2)\times\Omega\times\mathbb{R}^d) = \{f \in H^1_{\text{kin}}((T_1,T_2)\times\Omega) : f(t,x,v) = 0 \ \text{ if } (x,v) \in \partial_{\text{kin}}\Omega\}.$$

where

$$\partial_{\text{kin}}\Omega = \{(x,v) : x \in \partial\Omega, v\cdot\eta(x) < 0\},$$

and $\eta(x)$ is the outward pointing normal vector to $\partial\Omega$. We define the semi-norm on this by

$$[\![f]\!]_{H^1_{\text{kin}}} = \|\nabla_v f\|_{L^2} + \sup_{h\in H^1_{\text{kin}}, \|\nabla_v h\|_{L^2}=1} \int_{T_1}^{T_2} \int_{\Omega\times\mathbb{R}^d} (Yf)(t,x,v)h(t,x,v)\, dv\, dx\, dt. \tag{4.21}$$

In some sense, the last integral should really be understood as an H_v^{-1}-H_v^1 pairing in the v-variable that is equal to the integral if u and g are sufficiently smooth. We abuse notation, however, and simply write the integral. This is justified due to the density of smooth functions; see discussion in [2]. A norm on $H^1_{\text{kin},0}$ is obtained by including the L^2-norm as well. One can then construct $H^1_{\text{kin},0}$ as the closure of C_c^∞ functions under this norm.

Let us note that there is not accepted convention on the "correct" kinetic Sobolev space. There are several approaches to kinetic Besov and Sobolev spaces, e.g. [2, 36, 78]. We use the one proposed by Albritton, Armstrong, Mourrat, and Novack in [2] as it appears to pair well with the equation (4.1). Indeed, using (4.21), one immediately obtains an H^1_{kin}-bound from the energy equality (4.18) (see the discussion below (4.18)).

4.3.2 The Lie group structure, kinetic distance, and kinetic convolution

To aid the reader, let us review standard facts on the scaling and Lie group structure relevant to kinetic Fokker-Planck equations. This simplifies many arguments notationally and technically.

The equation (4.1) has a 2-3-1 scaling law; that is, it is invariant under dilations

$$\delta_r z = (r^2 t, r^3 x, rv).$$

Given z, z', we define

$$z \circ z' = (t + t', x + x' + t'v, v + v').$$

Roughly, this reflects the structure of (4.1) that allows mass to move by diffusion in v and by transport in (t, x). Indeed, if a unit of mass is at (x, v) and we move forward in time by t' units, our mass shifts to $x \mapsto x + t'v$. In fact, one sees that, for a fixed z_0, $\widetilde{f}(z) = f(z_0 \circ z)$ solves the first equation in (4.1). This is sometimes referred to as the Galilean invariance of (4.1) and is at the heart of why $\circ$ is the appropriate nothing of translation.

Clearly this action is invertible with

$$z^{-1} = (-t, -x + tv, -v),$$

whence

$$z^{-1} \circ \widetilde{z} = (\widetilde{t} - t, \widetilde{x} - x - (\widetilde{t} - t)v, \widetilde{v} - v) \quad \text{and}$$
$$z \circ \widetilde{z}^{-1} = (t - \widetilde{t}, x - \widetilde{x} - \widetilde{t}(v - \widetilde{v}), v - \widetilde{v}).$$

Note the group action is non-commutative but associative.

Given two sets $A, B \subset \mathbb{R}^{2d+1}$, we can analogously define the Lie action between them:

$$A \circ B = \{a \circ b : a \in A, b \in B\}. \tag{4.22}$$

and we also define the set of inverses

$$B^{-1} = \{b^{-1} : b \in B\}.$$

This will play a role in understanding how A and A_ε relate to each other integrals of the form

$$\int_{A_\varepsilon} u_\varepsilon \, dz \approx \int_A u \, dz,$$

where u_ε is defined via convolution of u with a compactly supported mollifier (see Lemma 4.A.1).

There are several norms and distances that one may choose. Here, we follow [54] and use

$$d_{\text{kin}}(z', z) = \| z^{-1} \circ z' \| \quad \text{and} \quad \| z \| = \min_{w \in \mathbb{R}^d} \left[\max \left\{ |t|^{1/2}, |x - tw|^{1/3}, |v - w|, |w| \right\} \right]. \tag{4.23}$$

We point out that this norm respects the 2-3-1 scaling and Galilean invariance of the equation (4.1). Indeed,

$$d_{\text{kin}}(\delta_r z_1, \delta_r z_2) = r d_{\text{kin}}(z_1, z_2) \quad \text{and} \quad d_{\text{kin}}(z \circ z_1, z \circ z_2) = d_{\text{kin}}(z_1, z_2).$$

An advantage to this choice of distance and norm is that

$$\| z \circ z' \| \leq \| z \| + \| z' \|,$$

so that the triangle inequality for d_{kin} holds:

$$\begin{aligned} d_{\text{kin}}(z_1, z_3) &= \| z_3^{-1} \circ z_1 \| = \| z_3^{-1} \circ z_2 \circ z_2^{-1} \circ z_1 \| \\ &\leq \| z_3^{-1} \circ z_2 \| + \| z_2^{-1} \circ z_1 \| = d_{\text{kin}}(z_2, z_3) + d_{\text{kin}}(z_1, z_2). \end{aligned} \tag{4.24}$$

One can easily check that d_{kin} is symmetric and positive definition, and, hence, it is a metric (see [53, Proposition 2.2]). Naturally, one defines the kinetic cylinders

$$Q_r(z_0) = \{z : t \leq t_0, d_{\text{kin}}(z, z_0) \leq r\}.$$

We will, importantly, consider the distance between a point and a set. This is defined in the traditional way:

$$\operatorname{dist}(S, z) = \inf_{s \in S} d_{\text{kin}}(s, z).$$

It is sometimes useful to use the obvious equality

$$\mathrm{dist}(S, z) = \inf\left\{\|z'\| : z \circ z' \in S\right\}. \tag{4.25}$$

Indeed, for every $s \in S$, we can take $z' = z^{-1} \circ s$, whence $\|z'\| = d_{\mathrm{kin}}(s, z)$.

Finally, we define the kinetic convolution:

$$(f * g)(z) = \int f(\tilde{z}) g(\tilde{z}^{-1} \circ z)\, d\tilde{z}. \tag{4.26}$$

We note that this conflicts with the standard notation for convolution; however, as that does not appear in this work, there is no risk of confusion. We sometimes convolve f and g where g has no time dependence. In this case, we abuse notation and denote it the same way. We note that, for any i,

$$\partial_{v_i}(f * g) = f * (\partial_{v_i} g) \qquad \text{and} \qquad Y(f * g) = f * (Yg). \tag{4.27}$$

We refer to [54, Section 3.3] and [82, Section 2] for more in-depth discussion.

4.4 Statement of the main propositions and proof of the main theorem

4.4.1 The Poincaré, Nash, and outgoing regions

We first decompose $\mathbb{H}^d \times \mathbb{R}^d$ into natural subdomains on which different functional inequalities hold. Let

$$\begin{cases} \mathcal{P}_R &= \{(x, v) \in \mathbb{H}^d \times \mathbb{R}^d : \mathrm{dist}(\mathbb{R} \times \gamma_-, (0, x, v)) \leq \sqrt{R}\}, \\ \mathcal{O}_R &= \{(x, v) \in (\mathbb{H}^d \times \mathbb{R}^d) \setminus \mathcal{P}_R : \mathrm{dist}(\mathbb{R} \times \gamma_+, (0, x, v)) \leq \sqrt{R/10}\}, \qquad \text{and} \\ \mathcal{N}_R &= \{(x, v) \in \mathbb{H}^d \times \mathbb{R}^d : \mathrm{dist}(\mathbb{R} \times \partial\mathbb{H}^d \times \mathbb{R}^d, (0, x, v)) \geq \sqrt{R/10}\}. \end{cases} \tag{4.28}$$

The reason for the difference in choice of distance for $\mathcal{O}_R$ and $\mathcal{N}_R$ is technical and related to the fact that we want v_1 to be bounded away from zero when $(x, v) \in \mathcal{O}_R$.

Clearly $\mathcal{P}_R \cup \mathcal{O}_R \cup \mathcal{N}_R$ is a decomposition of $\mathbb{H}^d \times \mathbb{R}^d$. In the proof we handle the estimates on each set separately. Along these lines, we require cutoff functions with nice scaling properties for each set. For this, it is useful to note that

$$\mathcal{P}_R = \delta_{1/\sqrt{R}} P_1, \quad \mathcal{O}_R = \delta_{1/\sqrt{R}} O_1, \quad \text{and} \quad \mathcal{N}_R = \delta_{1/\sqrt{R}} N_1. \tag{4.29}$$

It is sometimes helpful to keep in mind that we eventually choose $R = O(t)$. In this sense, $\mathcal{O}_R$ and $\mathcal{P}_R$ are, roughly, the sets in which transport can connect (x, v) to the boundaries γ_- and γ_+, respectively, in time $O(t)$. The one subtlety is that, for $\mathcal{P}_R$, we allow "jumps" in velocity of size $O(\sqrt{R})$, while we use pure transport in $\mathcal{O}_R$.

4.4.2 The main propositions and lemmas

The proof of Theorem 4.1.1 involves combining two estimates: the first is a general functional inequality that holds for any $g \in H^1_{\text{kin}}$, the second is a bound on how solutions f to (4.1) have far to the "bottom right," that is, where $x_1 \gg 1$ and $v_1 \ll -1$ (the isolated region $\mathcal{I}_R$ in Figure 4.1). This latter region is, on the time scale $t \approx R$, isolated from the boundaries, but is not on the infinite time scales on which $\widetilde{\varphi}$ and φ are defined. We break these into separate estimates at this point because both may be of an independent interest.

Let us state our general functional inequality here. It arises by, roughly, combining Poincaré type inequalities for $x_1 \ll \max\{|v_1|^3, |v_1|\}$ with the Nash inequality (Proposition 4.5.2) when $x_1 \gg \max\{|v_1|^3, |v_1|\}$. Its proof is in Section 4.5.

Lemma 4.4.1. *Fix $R, \delta > 0$. Suppose that $g \in H^1_{\text{kin},0}\left([T_1 - 2R, T_2] \times \mathbb{H}^d \times \mathbb{R}^d\right)$ with $T_1 \geq 2R$. Then*

$$\int_{T_1}^{T_2} \int g(z)^2 \, dz - \delta \int_{T_1 - 2R}^{T_2} \int g(z)^2 \, dz$$
$$\lesssim \frac{R}{\delta} [\![g]\!]^2_{H^1_{\text{kin}}([T_1 - 2R, T_2 + R] \times \mathbb{H}^d \times \mathbb{R}^d)} + R \int_{T_1}^{T_2 + R} \int_{\mathbb{R}^{d-1} \times \widetilde{\mathbb{H}^d}} |v_1| g(t, (0, \overline{x}), v)^2 \, d\overline{x} \, dv dt$$
$$+ \frac{1}{R^{2d+1}} \|g\|^2_{L^1([T_1 - 2R, T_2] \times \mathcal{N}_{R/2})}.$$

The first two terms on the right hand side are exactly as we would expect for the energy equality (4.18) associated to solutions of (4.1). The last term, however, is not what we desire because it does not include the steady solution to the adjoint equation $\widetilde{\varphi}$. On a portion of $\mathcal{N}_R$, we can "sneak in" a factor of $\widetilde{\varphi}/R^{1/4}$. Indeed, using (4.2), we can deduce the following lemma, whose proof is in Section 4.7:

Lemma 4.4.2. *Fix $R > 0$. If $(x, v) \in \mathcal{N}_R$ and $x_1 \geq -v_1^3$, then*

$$\widetilde{\varphi}(x, v) \gtrsim R^{1/4}.$$

Thus, on this subdomain, we can always replace the L^1-norm of g with $R^{-1/4}\|g\varphi^*\|_{L^1}$, which is a quantity conserved by the equation (4.1).

On the other hand, when $x \leq -v^3$, we have no lower bound on φ^* and can not appeal to $\|f\varphi^*\|_{L^1}$. This region is "too far" from the boundary to be influenced by it on a time-scale $t = O(R)$. Thus, we have the following estimate that quantifies how isolated it is.

Lemma 4.4.3. *Fix $R > 0$. Let f be a solution to* (4.1). *There exists a nonnegative function $\widetilde{\mu}_R \lesssim 1$ such that*

$$\begin{aligned} &\widetilde{\mu}_R(x,v) = 1 \qquad \text{if } (x,v) \in \mathcal{N}_R \cap \{x_1 \leq -v_1^3\}, \\ &\widetilde{\mu}_R(x,v) \lesssim \begin{cases} 0 & \text{if } v_1 > -\frac{1}{2}\sqrt{\frac{R}{10}} \text{ or } x_1 \geq 2|v_1|^3, \\ e^{\frac{Rv_1}{10x_1}} & \text{if } v_1 \leq -\sqrt{\frac{R}{10}} \text{ and } x_1 \leq -\frac{Rv_1}{10}, \end{cases} \end{aligned} \tag{4.30}$$

and

$$(\Delta_v + v\cdot\nabla_x)\widetilde{\mu}_R \lesssim \frac{1}{R^{5/4}}\widetilde{\varphi}. \tag{4.31}$$

As discussed in the introduction, the construction of this cutoff-type function requires some care as it has to encode the physics of the situation – particles in the $x \approx -Rv_1$ region will exit the region $\mathcal{N}_R \cap \{x_1 \leq -v_1^3\}$ in R units of time. That said, the proof of Lemma 4.4.3 is rather tedious, so we relegate it to Section 4.6.

Before continuing on, let us note that the $R/10$ is somewhat arbitrary. It comes from the $R/10$ taken in the definition of $\mathcal{N}_R$, which is mainly taken for convenience. This can certainly be improved, although it is not clear exactly what the optimal exponential decay rate is.

We now combine all estimate into one that will be the main functional inequality in the proof of Theorem 4.1.1.

Proposition 4.4.4. *Fix $T_1, T_2, R, \delta > 0$ with $T_2 > T_1 > 2R$. Suppose that $f \in H^1_{\mathrm{kin},0}\left([0,T_2]\times\mathbb{H}^d\times\mathbb{R}^d\right)$ solves* (4.1). *Then*

$$\begin{aligned} &\int_{T_1}^{T_2}\int f(z)^2\,dz - \delta\int_{T_1-2R}^{T_2}\int f(z)^2\,dz \\ &\quad\lesssim \frac{R}{\delta}[\![f]\!]^2_{H^1_{\mathrm{kin}}([T_1-2R,T_2+R]\times\mathbb{H}^d\times\mathbb{R}^d)} + R\int_{T_1}^{T_2+R}\int_{\mathbb{R}^{d-1}\times\widetilde{\mathbb{H}^d}}|v_1|f(t,(0,\overline{x}),v)^2\,d\overline{x}\,dv dt \\ &\qquad + \Big(\frac{(T_2-T_1)^4}{R^{2d+7/2}} + \frac{1}{R^{2d-1/2}}\Big)\Big(\int f_{\mathrm{in}}\widetilde{\varphi}\,dx\,dv\Big)^2 + \frac{(T_2-T_1)^2}{R^{2d+1}}\Big(\int f_{\mathrm{in}}\widetilde{\mu}_R\,dx\,dv\Big)^2. \end{aligned}$$

Proof. For convenience, let us write

$$\mathcal{N}_{R/2} = \mathcal{W}_{R/2} \cup \mathcal{I}_{R/2},$$

where

$$\mathcal{I}_{R/2} = \{(x,v) \in \mathcal{N}_{R/2} : \ v_1 \leq 0, \ x_1 \leq -v_1^3\} \quad \text{and} \quad \mathcal{W}_{R/2} = \mathcal{N}_{R/2} \setminus \mathcal{I}_{R/2}.$$

Here $\mathcal{I}_{R/2}$ is the "isolated region," where the effects of the boundary have not "yet" been felt. In this region, we use Lemma 4.4.3. Its complement, $\mathcal{W}_{R/2}$, is the "weighted region", where $\widetilde{\varphi}$ can be included directly in the integral via Lemma 4.4.2.

First, we note that, by Lemma 4.4.3,

$$\begin{aligned}
\frac{d}{dt}\int f(t,x,v)\widetilde{\mu}_R(x,v)\,dx\,dv &= \int \left[(\Delta_v - v\cdot\nabla_x) f(t,x,v)\right]\widetilde{\mu}_R\,dx\,dv \\
&= \int f(t,x,v)\left(\Delta_v + v\cdot\nabla_x\right)\widetilde{\mu}_R\,dx\,dv \\
&\lesssim \frac{1}{R^{5/4}}\int f(t,x,v)\widetilde{\varphi}(x,v)\,dx\,dv = \frac{1}{R^{5/4}}\int f_{\text{in}}(x,v)\widetilde{\varphi}(x,v)\,dx\,dv.
\end{aligned}$$

The last equality holds by (4.4). We deduce that

$$\begin{aligned}
\int_{\mathcal{I}_{R/2}} f(t,x,v)\,dx\,dv &\leq \int f(t,x,v)\widetilde{\mu}_R(x,v)\,dx\,dv \\
&\lesssim \frac{t}{R^{5/4}}\int f_{\text{in}}\widetilde{\varphi}\,dx\,dv + \int f_{\text{in}}(x,v)\widetilde{\mu}_R(x,v)\,dx\,dv.
\end{aligned}$$

Next, we use Lemma 4.4.2 to find

$$\int_{\mathcal{W}_{R/2}} f(t,x,v)\,dx\,dv \lesssim \frac{1}{R^{1/4}}\int_{\mathcal{W}_{R/2}} f(t,x,v)\widetilde{\varphi}\,dx\,dv = \frac{1}{R^{1/4}}\int_{\mathcal{W}_{R/2}} f_{\text{in}}\widetilde{\varphi}\,dx\,dv.$$

In total, we deduce that

$$\begin{aligned}
\|f\|_{L^1([T_1-2R,T_2]\times\mathcal{N}_{R/2})} \lesssim& \frac{(T_2-T_1)^2}{R^{5/4}}\int f_{\text{in}}\widetilde{\varphi}\,dx\,dv + (T_2-T_1)\int f_{\text{in}}(x,v)\widetilde{\mu}_R(x,v)\,dx\,dv \\
&+ R^{3/4}\int f_{\text{in}}\widetilde{\varphi}\,dx\,dv.
\end{aligned} \tag{4.32}$$

The combination of this inequality with Lemma 4.4.1 completes the proof. □

4.4.3 Proof of the main result: Theorem 4.1.1

Proof. The proof takes several steps. All but the last aim for a weighted $L^1 \to L^2$ type estimate. The last step bootstraps that to a weighted $L^2 \to L^\infty$ type estimate.

 # Step one: setting notation. For ease, let us denote the "energy" and "dissipation" as

$$E(t) = \int f(t,x,v)^2 \, dx \, dv \quad \text{and} \quad D(t) = \int |\nabla_v f(t,x,v)|^2 \, dx \, dv. \tag{4.33}$$

Although physically it is not correct to call E the energy, we abuse terminology and do so in analogy with work for parabolic equations. It is also useful to set notation for the boundary term

$$B(t) = \int_{\gamma_+} |v_1| f(t,(0,\overline{x}),v) \, d\overline{x} \, dv.$$

Applying (4.18) yields, for any nonnegative $t_1 < t_2$,

$$E(t_2) + \int_{t_1}^{t_2} \left(D(s) + B(s)\right) \, ds \leq E(t_1). \tag{4.34}$$

We see that E is decreasing.

 # Step two: applying the weighted Nash inequality Proposition 4.4.4. With $\delta, \epsilon \in (0, {}^1\!/_{100})$ be to chosen, we let

$$R = \varepsilon t, \quad T_1 - 2R = {}^t\!/_2, \quad \text{and} \quad T_2 + R = t.$$

Then, in the notation above and in view of the correspondence (4.19) between D and the H^1_{kin}-norm, Proposition 4.4.4 yields

$$\begin{aligned}
&\int_{\frac{1+4\varepsilon}{2}t}^{(1-\varepsilon)t} E(s)\, ds - \delta \int_{{}^t\!/_2}^{(1-\varepsilon)t} E(s)\, ds \\
&\qquad \lesssim \frac{\varepsilon t}{\delta} \int_{{}^t\!/_2}^{t} \left(D(s) + B(s)\right) \, ds + \frac{1}{\varepsilon^{2d+{}^7\!/_2} t^{2d-{}^1\!/_2}} \Big(\int f_{\text{in}} \widetilde{\varphi} \, dx \, dv \Big)^2 \\
&\qquad\qquad + \frac{1}{\varepsilon^{2d+1} t^{2d-1}} \Big(\int f_{\text{in}} \widetilde{\mu}_{\varepsilon t} \, dx \, dv \Big)^2.
\end{aligned}$$

Combining this with (4.34), we see that

$$\begin{aligned}
&\int_{\frac{1+4\varepsilon}{2}t}^{(1-\varepsilon)t} E(s)\, ds - \delta \int_{{}^t\!/_2}^{(1-\varepsilon)t} E(s)\, ds \\
&\qquad \lesssim \frac{\varepsilon t}{\delta} E({}^t\!/_2) + \frac{1}{\varepsilon^{2d+{}^7\!/_2} t^{2d-{}^1\!/_2}} \Big(\int f_{\text{in}} \widetilde{\varphi} \, dx \, dv \Big)^2 + \frac{1}{\varepsilon^{2d+1} t^{2d-1}} \Big(\int f_{\text{in}} \widetilde{\mu}_{\varepsilon t} \, dx \, dv \Big)^2.
\end{aligned} \tag{4.35}$$

 # Step three: setting up a "first touching" argument. Fix $\bar{\alpha}, \overline{\beta} \gg 1$ be constants to be chosen, and let

$$\alpha = \overline{\alpha}\Big(\int f_{\rm in}\widetilde{\varphi}\, dx\, dv\Big)^2 \quad \text{and} \quad \beta = \overline{\beta}\Big(\int f_{\rm in}\widetilde{\mu}_{\varepsilon t_0}\, dx\, dv\Big)^2. \tag{4.36}$$

Define

$$t_0 = \sup\Big\{t : E(s) \le \frac{\alpha}{t^{2d+1/2}} + \frac{\beta}{t^{2d}}\Big\}.$$

Up to approximation, we may assume that $f_{\rm in}$ is smooth and compactly supported, whence $t_0 > 0$. Our goal is to show that $t_0 = \infty$. Hence, we argue by contradiction assuming that t_0 is finite.

Let us note that, we immediately have, from the definition of t_0,

$$E(s) \lesssim \frac{\alpha}{t^{2d+1/2}} + \frac{\beta}{t^{2d}} \qquad \text{for all } s \in [t_0/4, t_0]. \tag{4.37}$$

We use this frequently in the sequel.

 # Step four: obtaining a contradiction to the definition of t_0. Moving the negative integral term on the left hand side of (4.35) to the right hand side and applying the definition (4.37) of t_0, we deduce that

$$\begin{aligned} t_0 E\left((1-\varepsilon)t_0\right) &\lesssim \frac{\varepsilon t}{\delta} E(t/2) + \delta \int_{t/2}^{(1-\varepsilon)t} E(s)\, ds + \frac{1}{\varepsilon^{2d+7/2} t^{2d-1/2}}\Big(\int f_{\rm in}\widetilde{\varphi}\, dx\, dv\Big)^2 \\ &\quad + \frac{1}{\varepsilon^{2d+1} t^{2d-1}}\Big(\int f_{\rm in}\widetilde{\mu}_{\varepsilon t}\, dx\, dv\Big)^2 \\ &\lesssim \left(\frac{\varepsilon}{\delta} + \delta\right) \frac{\alpha}{t_0^{2d-1/2}} + \left(\frac{\varepsilon}{\delta} + \delta\right) \frac{\beta}{t_0^{2d-1}} \\ &\quad + \frac{1}{\varepsilon^{2d+7/2} t_0^{2d-1/2}}\Big(\int f_{\rm in}\widetilde{\varphi}\, dx\, dv\Big)^2 + \frac{1}{\varepsilon^{2d+1} t_0^{2d-1}}\Big(\int f_{\rm in}\widetilde{\mu}_{\varepsilon t_0}\, dx\, dv\Big)^2. \end{aligned}$$

Recalling the definition (4.36) of α and β, we find

$$E\left((1-\varepsilon)t_0\right) \le \frac{\alpha}{t_0^{2d+1/2}}\left(\frac{\varepsilon}{\delta} + \delta + \frac{1}{\overline{\alpha}\varepsilon^{2d+7/2}}\right) + \frac{\beta}{t_0^{2d}}\left(\frac{\varepsilon}{\delta} + \delta + \frac{1}{\overline{\beta}\varepsilon^{2d+1}}\right).$$

Recalling again that E is decreasing, we have

$$E((1-\varepsilon)t_0) \ge E(t_0) = \frac{\alpha}{t_0^{2d+1/2}} + \frac{\beta}{t_0^{2d}}.$$

In summary,

$$\frac{\alpha}{t_0^{2d+1/2}}+\frac{\beta}{t_0^{2d}}\leq\frac{\alpha}{t_0^{2d+1/2}}\left(\frac{\varepsilon}{\delta}+\delta+\frac{1}{\overline{\alpha}\varepsilon^{2d+7/2}}\right)+\frac{\beta}{t_0^{2d}}\left(\frac{\varepsilon}{\delta}+\delta)+\frac{1}{\overline{\beta}\varepsilon^{2d+1}}\right).$$

Choosing ε small, then δ small (depending on ε), and then choosing $\bar{\alpha}$ and $\overline{\beta}$ large (depending on both ε and δ, we obtain a contradiction.

It follows that $t_0=\infty$ and, hence, for all $t>0$,

$$E(t)\leq\frac{\bar{\alpha}}{t^{2d+1/2}}\Big(\int f_{\mathrm{in}}\widetilde{\varphi}\,dx\,dv\Big)^2+\frac{\bar{\beta}}{t^{2d}}\Big(\int f_{\mathrm{in}}\widetilde{\mu}_{\varepsilon t}\,dx\,dv\Big)^2. \tag{4.38}$$

\# **Step five: some functional analysis and the conclusion.** The inequality (4.38) implies that the solution operator of (4.1)

$$S_t:X_t\to L^2(\mathbb{H}^d\times\mathbb{R}^d)$$

is well-defined and bounded. In other words, $S_tf_{\mathrm{in}}=f(t)$ if $f_{\mathrm{in}}\in X_t$. Here, we define the Banach space

$$X_t=L^1_{\widetilde{\varphi}}\cap\big(t^{-1/4}L^1_{\widetilde{\mu}_{\varepsilon t}}\big)=\Big\{h\in L^1_{\mathrm{loc}}(\mathbb{H}^d\times\mathbb{R}^d):\int|h|(\widetilde{\varphi}+\widetilde{\mu}_{\varepsilon t})\,dx\,dv<\infty\Big\}$$

with the norm

$$\|h\|_{X_t}=\int|h|\widetilde{\varphi}\,dx\,dv+t^{1/4}\int|h|\widetilde{\mu}_{\varepsilon t}\,dx\,dv.$$

Hence, (4.38) translates to the bound

$$\|S_t\|_{X_t\to L^2}\lesssim\frac{1}{t^{d+1/4}}. \tag{4.39}$$

By using the fact that $\widetilde{f}$ solves (4.10) if f solves (4.1), we also obtain the bound

$$\|\widetilde{S}_t\|_{\widetilde{X}_t\to L^2}\lesssim\frac{1}{t^{d+1/4}}.$$

where

$$\widetilde{S}_t:\widetilde{X}_t\to L^2(\mathbb{H}^d\times\mathbb{R}^d)$$

is the solution operator of (4.10) and

$$\widetilde{X}_t=L^1_{\varphi}\cap\big(t^{-1/4}L^1_{\mu_{\varepsilon t}}\big)=\Big\{h\in L^1_{\mathrm{loc}}(\mathbb{H}^d\times\mathbb{R}^d):\int|h|(\varphi+\mu_{\varepsilon t})\,dx\,dv<\infty\Big\}$$

with the norm

$$\|h\|_{\widetilde{X}_t} = \int |h|\varphi\, dx\, dv + t^{1/4} \int |h|\mu_{\varepsilon t}\, dx\, dv.$$

Let us note that since (4.1) and (4.10) are adjoint to one another,

$$\widetilde{S}_t^* : L^2 \to X_t^*$$

is also a solution operator to (4.1). By standard results on adjoint operators, we deduce that

$$\|\widetilde{S}_t^*\|_{L^2 \to \widetilde{X}_t^*} = \|\widetilde{S}_t\|_{\widetilde{X}_t \to L^2} \lesssim \frac{1}{t^{d+1/4}}. \tag{4.40}$$

It is easy to identify $\widetilde{X}_t^*$ as

$$\widetilde{X}_t^* = L^\infty_{1/\varphi} + t^{1/4} L^\infty_{1/\mu_{\varepsilon t}} = \{h : h = \varphi h_1 + t^{1/4}\mu_{\varepsilon t} h_2,\ h_i \in L^\infty\},$$

with the norm

$$\|h\|_{\widetilde{X}_t^*} = \inf_{h = \varphi h_1 + t^{1/4}\mu_{\varepsilon t} h_2} \left(\|h_1\|_{L^\infty} + \|h_2\|_{L^\infty}\right).$$

We now conclude using the semigroup property

$$f(t) = \widetilde{S}_{t/2}^* S_{t/2} f_{\text{in}}.$$

Indeed, recalling (4.39) and (4.40),

$$\|f(t)\|_{\widetilde{X}_{t/2}^*} = \|\widetilde{S}_{t/2}^* S_{t/2} f_{\text{in}}\|_{\widetilde{X}_t^*} \lesssim \frac{1}{t^{d+1/4}} \|S_{t/2} f_{\text{in}}\|_{L^2} \lesssim \frac{1}{t^{2d+1/2}} \|f_{\text{in}}\|_{X_{t/2}}.$$

The proof is finished after unpacking the definitions of the norms. □

4.5 The main functional inequality: Lemma 4.4.1

We decompose $\mathbb{H}^d \times \mathbb{R}^d$ into three regions $\mathcal{P}_R$, $\mathcal{N}_R$, and $\mathcal{O}_R$, depending on the influence of the boundary. Recall that these are defined in (4.28). We state L^2-estimates on each regime, postponing their proofs until Section 4.5.3. In Section 4.5.2, we combine these estimates to prove Lemma 4.4.1.

4.5.1 The functional inequalities on each region

We begin by stating the Poincaré-type inequality. This inequality is on the portion of the domain that is "close" to the boundary γ_-, where particles are absorbed. It quantifies this effect.

Proposition 4.5.1 (Poincaré inequality on the incoming region $\mathcal{P}_R$)**.** *Fix any positive numbers R, T_1, and T_2 such that $2R < T_1 < T_2$. Suppose that $g \in H^1_{\text{kin},0}([T_1 - 2R, T_2] \times \mathbb{H}^d \times \mathbb{R}^d)$. Then*

$$\begin{aligned}\|g\|^2_{L^2([T_1,T_2]\times\mathcal{P}_R)} \lesssim & R[\![g]\!]^2_{H^1_{\text{kin}}([T_1-2R,T_2]\times\mathbb{H}^d\times\mathbb{R}^d)} \\ &+ \sqrt{R}[\![g]\!]_{H^1_{\text{kin}}([T_1-2R,T_2]\times\mathbb{H}^d\times\mathbb{R}^d)}\|g\|_{L^2([T_1-2R,T_2]\times\mathbb{H}^d\times\mathbb{R}^d)}.\end{aligned} \tag{4.41}$$

Let us note that the norms on the right hand side can be localized to P_{cR}, for an appropriate $c > 1$, with some extra care in the proof. We opt for simplicity here.

Next, we state the Nash-type inequality. This inequality is on the portion of the domain that is "far" from all boundaries. It quantifies the fact that the evolution of (4.1) is on $\mathbb{R} \times \mathbb{H}^d \times \mathbb{R}^d$ is essentially the same it would be on $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$. Let us make note that $\mathcal{N}_{R/2}$ is *larger* than $\mathcal{N}_R$.

Proposition 4.5.2 (Nash inequality on $\mathcal{N}_R$)**.** *Suppose that $g \in H^1_{\text{kin}}\left([T_1, T_2] \times \mathbb{H}^d \times \mathbb{R}^d\right)$, and fix $R, \varepsilon > 0$. If ε is sufficiently small,*

$$\begin{aligned}\int_{T_1}^{T_2} \|g\|^2_{L^2(\mathcal{N}_R)}\,dt \lesssim & \sqrt{\varepsilon R}[\![g]\!]_{H^1_{\text{kin}}([T_1-\varepsilon R,T_2]\times\mathcal{N}_{R/2})}\|g\|_{L^2([T_1-\varepsilon R,T_2]\times\mathcal{N}_{R/2})} \\ &+ \frac{1}{(\varepsilon R)^{2d+1}}\|g\|^2_{L^1([T_1-\varepsilon R,T_2]\times\mathcal{N}_{R/2})}.\end{aligned}$$

Finally, we state Poincaré-type inequality of a different flavor. Particles near the outgoing part of the boundary γ_+, will likely leave the domain. In the context of (4.1), this quantifies the effect of the boundary in (4.18). The following estimate quantifies that.

Proposition 4.5.3 (Poincaré inequality on the outgoing region $\mathcal{O}_R$)**.** *Fix any positive numbers R, T_1, and T_2 such that $2R < T_1 < T_2$. Suppose that $g \in H^1_{\text{kin}}([T_1 - 2R, T_2] \times \mathbb{H}^d \times \mathbb{R}^d)$. Then*

$$\begin{aligned}\|g\|^2_{L^2([T_1,T_2]\times\mathcal{O}_R)} \lesssim & R[\![g]\!]^2_{H^1_{\text{kin}}([T_1,T_2+R]\times\mathbb{H}^d\times\mathbb{R}^d)} \\ &+ \sqrt{R}[\![g]\!]_{H^1_{\text{kin}}([T_1,T_2+R]\times\mathbb{H}^d\times\mathbb{R}^d)}\|g\|_{L^2([T_1,T_2+2R]\times\mathbb{H}^d\times\mathbb{R}^d))} \\ &+ R\int_{T_1}^{T_2+R}\int_{\mathbb{R}^d}\int_{\mathbb{R}^{d-1}} |v_1| f(t,(0,\overline{x}),v)^2\,d\overline{x}\,dv\,dt.\end{aligned}$$

As we noted after Proposition 4.5.1, the norms on the right hand side of the inequality in Proposition 4.5.3 can be localized with some extra care.

We remind the reader that the proofs of these lemmas can be found in Section 4.5.3.

4.5.2 The proof of Lemma 4.4.1

Proof. It is clear that this is a simple consequence of Propositions 4.5.1 to 4.5.3 combined with Young's inequality. We omit the details. □

4.5.3 Establishing the Nash and Poincaré-type inequalities

The proof of the Poincaré-type inequality on $\mathcal{P}_R$

To begin, we first show that the Poincaré regime can be characterized in a simple way, depending on (x, v). This is useful in understanding "paths" from any point in $\mathcal{P}_R$ to the boundary γ_+ that are at the heart of the proof of Proposition 4.5.1. The proof is postpone to Section 4.7.

Lemma 4.5.4. *Suppose that $(x, v) \in \mathcal{P}_R$. Then*

$$x_1 \leq R \max\{v_1, 3\sqrt{R}\} \quad \text{and} \quad -2\sqrt{R} \leq v_1.$$

Additionally, if $(x, v) \in \mathcal{O}_R$, then

$$\frac{x_1}{|v_1|} \leq R.$$

We now proceed with our Poincaré-type estimate. For the proof of Proposition 4.5.1, let us make the convention that every norm is taken on $[T_1 - 2R, T_2] \times \mathbb{H}^d \times \mathbb{R}^d$ unless otherwise specified. For example, by writing

$$\|g\|_{L^2} \quad \text{we mean} \quad \|g\|_{L^2([T_1-2R,T_2]\times\mathbb{H}^d\times\mathbb{R}^d)}.$$

This saves significant space and does not cost clarity.

Proof of Proposition 4.5.1. Let us extend g to $\overline{g}$ by

$$\overline{g}(t, x, v) = \begin{cases} g(t, x, v) & \text{if } x_1 > 0, \\ 0 & \text{if } x_1 \leq 0 < v_1. \end{cases}$$

Take any mollifier: a nonnegative, smooth function ψ such that

$$\int \psi(z)\,dz = 1.$$

Up to translation and dilation, we may assume that

$$\operatorname{supp}\psi \subset (0,1)\times\mathbb{H}^d\times\widetilde{\mathbb{H}^d} = \{t\in(0,1), x_1>0, v_1<0\}.$$

Define, for any $s\in(0,\sqrt{\varepsilon R}]$,

$$g_s(z) = (\psi_s * \overline{g})(z) = \int \psi(\delta_{1/s}(\widetilde{z}))\overline{g}(\widetilde{z}^{-1}\circ z)\frac{d\widetilde{z}}{s^{4d+2}} = \int \psi(\widetilde{z})\overline{g}(\delta_s(\widetilde{z})^{-1}\circ z)\,d\widetilde{z},$$

where, for all z,

$$\psi_s(z) := \frac{1}{s^{4d+2}}\psi\left(\delta_{1/s}(z)\right) \qquad \text{satisfies} \qquad \int \psi_s\,dw\,dy\,ds = 1$$

We observe a few simple facts about g_s. First, after changing variables, we see that g_s is smooth. Second, from (4.27), it is clear that

$$\lim_{s\to 0}\|g_s - g\|_{H^1_{\text{kin}}} = 0.$$

Hence, we need only prove (4.41) for g_s. Finally, due to the choice of support of ψ, we see that g_s is well-defined on

$$[T_1-\varepsilon R, T_2]\times\Gamma := [T_1-\varepsilon R, T_2]\times\{x_1\geq 0 \text{ or both } x_1\leq 0 \text{ and } v_1\geq 0\},$$

and that

$$g_s(t,x,v) = 0 \qquad \text{for any } x_1\leq 0, v_1\geq 0.$$

We use the notation Γ here because this region, when $d=1$, looks approximately like a "backwards" Γ. See Figure 4.2.

Let $\chi_{\mathcal{P}_R}$ be a cutoff function for $\mathcal{P}_R$ such that

$$\chi_{\mathcal{P}_R}\equiv 1 \qquad \text{in } \mathcal{P}_R \quad \text{and} \quad \chi_{\mathcal{P}_R}\equiv 0 \qquad \text{in } \mathcal{P}^c_{3R/2} \tag{4.42}$$

while

$$\|\nabla_v\chi_{\mathcal{P}_R}\|_{L^\infty}\lesssim\frac{1}{\sqrt{R}} \quad \text{and} \quad \|\nabla_x\chi_{\mathcal{P}_R}\|_{L^\infty}\lesssim\frac{1}{R^{3/2}}. \tag{4.43}$$

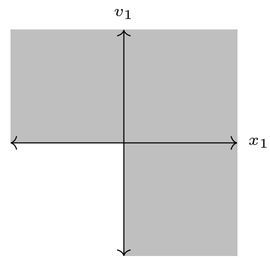

Figure 4.2: The region Γ. Notice that it looks like a backwards Γ.

This can easily be constructed when $R = 1$ and the general case follows by letting

$$\chi_{\mathcal{P}_R}(z) = \chi_{P_1}\left(\delta_{1/\sqrt{R}}(z)\right). \tag{4.44}$$

Notice that we use the scaling (4.29).

Fix any $(x, v) \in \operatorname{supp} \chi_{\mathcal{P}_R}$, and let

$$v_R = v + 10 e_1 \sqrt{R}$$

for succinctness. Here $e_1 = (1, 0, \ldots, 0)$ is the first canonical basis vector. Let us note that, if s is sufficiently small in a way depending only on ε and R,

$$g_s(x - 2R v_R, v_R) = 0$$

because, recalling that $(x, v) \in \mathcal{P}_{3R/2}$ and using Lemma 4.5.4,

$$x_1 - 2R v_1 - 20 R^{3/2} < 0 < v_1 + 10\sqrt{R}.$$

Hence, we may write

$$g_s(t, x, v)^2 = -2 \int_0^{10\sqrt{R}} (g_s\, \partial_{v_1} g_s)(t, x, v + r e_1)\, dr - 2 \int_0^{2R} (g_s\, Y g_s)(t - s, x - r v_R, v_R)\, dr.$$

We deduce that

$$\begin{aligned}
\int_{T_1}^{T_2} \int_{\mathcal{P}_R} g_s(t, x, v)^2\, dz &\le \int_{T_1}^{T_2} \int_{\mathbb{H}^d \times \mathbb{R}^d} g_s(t, x, v)^2 \chi_{\mathcal{P}_R}(x, v)^2\, dz \\
&= -2 \int_{T_1}^{T_2} \int_{\mathbb{H}^d \times \mathbb{R}^d} \Big(\int_0^{10\sqrt{R}} (g_s\, \partial_{v_1} g_s)(t, x, v + r e_1)\, dr \Big) \chi_{\mathcal{P}_R}(x, v)^2\, dt\, dx\, dv \\
&\quad - 2 \int_{T_1}^{T_2} \int_{\mathbb{H}^d \times \mathbb{R}^d} \Big(\int_0^{2R} (g_s\, Y g_s)(t - s, x - r v_R, v_R)\, dr \Big) \chi_{\mathcal{P}_R}(x, v)^2\, dt\, dx\, dv \\
&=: I_1 + I_2.
\end{aligned}$$

We estimate each term in turn.

Let us handle I_1 first as it is simpler. Then

$$\begin{aligned}
|I_1| &\approx \Big| \int_0^{10\sqrt{R}} \int_{T_1}^{T_2} \int_{\mathbb{H}^d\times\mathbb{R}^d} (g_s\,\partial_{v_1} g_s)(t,x,v)\chi_{\mathcal{P}_R}(x, v - re_1)^2\,dt\,dx\,dv\,dr \Big| \\
&\leq \int_0^{10\sqrt{R}} \|\partial_v g_s \chi_{\mathcal{P}_R}(\cdot, \cdot - re_1)\|_{L^2([T_1,T_2]\times\mathbb{H}^d\times\mathbb{R}^d)} \|g_s\chi_{\mathcal{P}_R}(\cdot,\cdot - re_1)\|_{L^2([T_1,T_2]\times\mathbb{H}^d\times\mathbb{R}^d)}\,dr \\
&\lesssim \sqrt{R}[\![g_s]\!]_{H^1_{\mathrm{kin}}} \|g_s\|_{L^2}.
\end{aligned} \tag{4.45}$$

We now consider I_2. We first change the order of integration and change variables:

$$\begin{aligned}
-\frac{1}{2}I_2 &= \int_0^{2R}\int_{T_1}^{T_2}\int_{\mathbb{H}^d\times\mathbb{R}^d} (g_s\,Y g_s)(t-r, x - rv_R, v_R)\chi_{\mathcal{P}_R}(x,v)^2\,dt\,dx\,dv\,dr \\
&= \int_0^{2R}\int_{T_1-r}^{T_2-r}\int_{\mathbb{H}^d\times\mathbb{R}^d} (g_s\,Y g_s)(t,x,v)\chi_{\mathcal{P}_R}(x + rv_R, v - 10\sqrt{R}e_1)^2\,dt\,dx\,dv\,dr.
\end{aligned}$$

In the second equality, we used that $g_s(t,x,v) \equiv 0$ for $x_1 < 0 < v_1$ and $\chi_{\mathcal{P}_R}(x, v - 10\sqrt{R}e_1) \equiv 0$ if $v_1 \leq 0$ (see Lemma 4.5.4 and (4.42)). We now use that H_v^{-1}-H_v^1 pairing of Yg_s with $g_s\chi_{\mathcal{P}_R}$. We find

$$\begin{aligned}
|I_2| &\lesssim \int_0^{2R} [\![g_s]\!]_{H^1_{\mathrm{kin}}} \|\nabla_v \left(g_s\chi_{\mathcal{P}_R}(\cdot + rv_R, \cdot - 10\sqrt{R}e_1)\right)\|_{L^2([T_1,T_2]\times\mathbb{H}^d\times\mathbb{R}^d)}\,dr \\
&\lesssim [\![g_s]\!]_{H^1_{\mathrm{kin}}} \int_0^{2R} \left(\|\nabla_v g_s\|_{L^2} + \|g_s\|_{L^2}\left(\frac{r}{R^{3/2}} + \frac{1}{\sqrt{R}}\right)\right)\,dr \\
&\lesssim R[\![g_s]\!]_{H^1_{\mathrm{kin}}}^2 + \sqrt{R}[\![g_s]\!]_{H^1_{\mathrm{kin}}}\|g_s\|_{L^2}.
\end{aligned} \tag{4.46}$$

where the second inequality follows from a simple computation of $\|\nabla_v \chi_{\mathcal{P}_R}(\cdot + sv_R, v_R)\|_{L^\infty}$ using (4.43). The combination of (4.45) and (4.46) finishes the proof. □

The proof of the Nash-type inequality on $\mathcal{N}_R$

Proof of Proposition 4.5.2. Our proof proceeds by an interpolation argument using a mollifier. With this in mind, take any compactly supported, nonnegative, smooth function ψ such that

$$\int \psi(z)\,dz = 1. \tag{4.47}$$

Up to translation and dilation, we may assume that

$$\operatorname{supp}\psi \subset \{z \in (1/2, 1)\times\mathbb{R}^{2d} : d_{\mathrm{kin}}(0,z) \leq 1\}. \tag{4.48}$$

For $\varepsilon \in (0,1)$ to be chosen and any $s \in (0, \sqrt{\varepsilon R}]$, define

$$g_s(z) = g * \psi_s(z) = \int g(z \circ \widetilde{z}^{-1}) \psi(\delta_{1/s}(\widetilde{z})) \frac{d\widetilde{z}}{s^{4d+2}} = \int g(z \circ \delta_s(\widetilde{z})^{-1}) \psi(\widetilde{z})\, d\widetilde{z},$$

where, for all z,

$$\psi_s(z) := \frac{1}{s^{4d+2}} \psi\left(\delta_{1/s}(z)\right) \qquad \text{satisfies} \qquad \int \psi_s \, dw\, dy\, ds = 1$$

by (4.47) and a standard change of variables. Clearly, $g_s \to g$ as $s \to 0$.

We note that the order of convolution does not matter here; in our arguments, only the scaling plays an important role. Indeed, one could argue similarly using $g_s = \psi_s * g$ instead.

For later, we note that, recalling the definition in (4.24),

$$(\operatorname{supp} \psi_s)^{-1} \subset \{z \in [0,s] \times \mathbb{R}^{2d} : d_{\mathrm{kin}}(0,z) \leq s\}^{-1} = \overline{Q}_s. \tag{4.49}$$

In order to localize g to the domain $\mathcal{N}_R$, we use a cutoff function $\chi_{\mathcal{N}_R}$ such that

$$\chi_{\mathcal{N}_R} \equiv 1 \qquad \text{in } \mathcal{N}_R \quad \text{and} \quad \chi_{\mathcal{N}_R} \equiv 0 \qquad \text{in } \mathcal{N}^c_{3R/4}, \tag{4.50}$$

while

$$\|\nabla_v \chi_{\mathcal{N}_R}\|_{L^\infty} \lesssim \frac{1}{\sqrt{R}} \quad \text{and} \quad \|\nabla_x \chi_{\mathcal{N}_R}\|_{L^\infty} \lesssim \frac{1}{R^{3/2}}. \tag{4.51}$$

This can be constructed exactly as in (4.42)-(4.44) for $\chi_{\mathcal{P}_R}$.

Before embarking on the estimate, let us understand the supports of the various functions. First, clearly, up to decreasing ε,

$$\operatorname{supp}\left(\chi_R \, g_{\sqrt{\varepsilon R}}\right) \subset \mathcal{N}_{3R/4}$$

and

$$\int_{T_1}^{T_2} \int_{\mathcal{N}_R} g^2 \, dz \leq \int_{\Omega_R} \chi_R \, g^2 \, dz$$

due to (4.50)-(4.51). Here we have made the change of notation to

$$\Omega_R := [T_1 - \varepsilon R, T_2] \times \mathcal{N}_{R/2}$$

for simplicity. Hence, we have

$$\int_{T_1}^{T_2} \int_{\mathcal{N}_R} g^2 \, dz \leq \int_{T_1}^{T_2} \int \chi_R g^2_{\sqrt{\varepsilon R}} \, dz + \int_{T_1}^{T_2} \int \chi_R (g^2 - g^2_{\sqrt{\varepsilon R}}) \, dz =: I_1 + I_2. \tag{4.52}$$

Our goal is to show that

$$I_1 \lesssim \frac{1}{(\varepsilon R)^{2d+1}} \|g\|^2_{L^1(\Omega_R)}, \tag{4.53}$$

and

$$I_2 \lesssim \sqrt{\varepsilon R} [\![g]\!]_{H^1_{\mathrm{kin}}(\Omega_R)} \|g\|_{L^2(\Omega_R)}. \tag{4.54}$$

Indeed, this would complete the proof.

 # Step one: applying Young's convolution inequality to I_1. We begin by analyzing I_1, which is the simpler of the two cases. Here, we simply use the kinetic version of Young's inequality for convolutions: Lemma 4.A.1. Indeed, we have

$$I_1 \le \|g * \psi_{\sqrt{\varepsilon R}}\|^2_{L^2([T_1,T_2]\times \mathcal{N}_{3R/4})} \le \|g\|^2_{L^1(([T_1\times T_2]\times \mathcal{N}_{3R/4})\circ Q_{\sqrt{\varepsilon R}})} \|\psi_{\varepsilon R}\|^2_{L^2},$$

where we used (4.49) to analyze the support of $\psi_{\sqrt{\varepsilon R}}$. Using only the definition of $N.$ and the triangle inequality for d_{kin}, it is easy to see that, up to decreasing ε, we have

$$\Big(\big([T_1, T_2] \times \mathcal{N}_{3R/4}\big) \circ Q_{\sqrt{\varepsilon R}}\Big) \subset [T_1 - \varepsilon R, T_2] \times \mathcal{N}_{R/2} = \Omega_R,$$

as desired. Additionally, a straightfoward computation yields

$$\|\psi_{\sqrt{\varepsilon R}}\|^2_{L^2} \lesssim \frac{1}{(\varepsilon R)^{2d+1}}.$$

Hence, (4.53) is proved.

 # Step two: rewriting I_2 as a series of integrals. Let us alter our notation:

$$g_{a,b,c}(z) := \int \underbrace{g\big(z \circ (0, 0, a(-\widetilde{v} + {}^{\widetilde{x}}\!/\!_{\widetilde{t}})) \circ (-b\widetilde{t}, 0, 0) \circ (0, 0, -{}^{c\widetilde{x}}\!/\!_{\widetilde{t}})\big)}_{=: g_{a,b,c}(z;\widetilde{z})} \psi(\widetilde{z})\, d\widetilde{z}.$$

In the sequel, this is useful because the first and third group actions correspond to shifts in v, which are represented in the H^1_{kin}-norm by the L^2-norm of $\nabla_v g$, while the second group action corresponds to a shift along transport, which is represented in the H^1_{kin}-norm by the $L^2_{t,x} H^{-1}_v$-norm of Yg.

We clearly have

$$g_{0,0,0}(z) = g(z) \quad \text{and} \quad g_{s,s^2,s}(z) = g_s(z).$$

The fundamental theorem of calculus then yields

$$\begin{aligned} g(z)^2 - g_{\sqrt{\varepsilon R}}(z)^2 = &- 2\int_0^{\sqrt{\varepsilon R}} g_{a,0,0}(z)\partial_a g_{a,0,0}(z)\,da \\ &- 2\int_0^{\varepsilon R} g_{\sqrt{\varepsilon R},b,0}(z)\partial_b g_{\sqrt{\varepsilon R},b,0}(z)\,db \\ &- 2\int_0^{\sqrt{\varepsilon R}} g_{\sqrt{\varepsilon R},\varepsilon R,c}(z)\partial_c g_{\sqrt{\varepsilon R};\varepsilon R,c}(z)\,dc. \end{aligned} \tag{4.55}$$

Let us write I_2, defined in (4.52), as

$$I_2 = I_{21} + I_{22} + I_{23},$$

where each of I_{2k} above corresponds, respectively, to a term in (4.55).

 # Step three: bound I_{21}. This case is simple. It follows by, in turn, directly computing the a derivative in (4.55), changing the order of integration, using the Cauchy-Schwarz inequality, and noticing that

$$\left| -\widetilde{v} + \frac{\widetilde{x}}{\widetilde{t}} \right| \lesssim 1,$$

due to the choice (4.48) of the support of ψ. Indeed, we find:

$$\begin{aligned} I_{21} &= -2\int_{T_1}^{T_2}\int \chi_R(z)\int_0^{\sqrt{\varepsilon R}}\int g_{a,0,0}(z;\widetilde{z})\partial_a g_{a,0,0}(z;\widetilde{z})\psi(\widetilde{z})\,d\widetilde{z}\,da\,dz \\ &= -2\int_{T_1}^{T_2}\int \chi_R(z)\int_0^{\sqrt{\varepsilon R}}\int g\left(t,x,v+a\left(-\widetilde{v}+\frac{\widetilde{x}}{\widetilde{t}}\right)\right) \\ &\qquad\qquad \left(-\widetilde{v}+\frac{\widetilde{x}}{\widetilde{t}}\right)\cdot\nabla_v g\left(t,x,v+a\left(-\widetilde{v}+\frac{\widetilde{x}}{\widetilde{t}}\right)\right)\psi(\widetilde{z})\,d\widetilde{z}\,da\,dz \\ &= -2\int_0^{\sqrt{\varepsilon R}}\int\int_{T_1}^{T_2}\int \chi_R(x,v+a\widetilde{v}-\frac{a\widetilde{x}}{\widetilde{t}})g(z)\left(-\widetilde{v}+\frac{\widetilde{x}}{\widetilde{t}}\right)\cdot\nabla_v g(z)\psi(\widetilde{z})\,dz\,d\widetilde{z}\,da \\ &\lesssim \int_0^{\sqrt{\varepsilon R}}\int\int_{T_1}^{T_2}\int \|g\|_{L^2(\Omega_R)}\|\nabla_v g\|_{L^2(\Omega_R)}\psi(\widetilde{z})\,d\widetilde{z}\,da \\ &\le \sqrt{\varepsilon R}\|g\|_{L^2(\Omega_R)}[\![g]\!]_{H^1_{\text{kin}}(\Omega_R)}. \end{aligned} \tag{4.56}$$

In the first inequality, we used that

$$[T_1,T_2]\times\operatorname{supp}_{x,v}\chi_R(\cdot,\cdot+a\widetilde{v}-{}^{a\widetilde{x}}\!/\!_{\widetilde{t}}) \subset [T_1-\varepsilon R, T_2]\times\mathcal{N}_{R/2} = \Omega_R. \tag{4.57}$$

The inclusion in the time variable is simply due to enlarging the domain. We show the inclusion of the spatial and velocity variables by using the triangle inequality, the fact that $|a| \le \sqrt{\varepsilon R}$, and by decreasing ε if necessary. Indeed, fix any point

$$(x, v) \in \operatorname{supp} \chi_R(\cdot, \cdot + a\widetilde{v})$$

and any point $z_\partial \in \mathbb{R} \times \partial\mathbb{H}^d \times \mathbb{R}^d$. Then, by triangle inequality

$$\begin{aligned} d_{\mathrm{kin}}\left(z_\partial, (0, x, v)\right) &\ge d_{\mathrm{kin}}\left(z_\partial, (0, x, v + a\widetilde{v} - {}^{a\widetilde{x}}/\widetilde{t})\right) - d_{\mathrm{kin}}\left((0, x, v + a\widetilde{v} - {}^{a\widetilde{x}}/\widetilde{t}), (0, x, v)\right) \\ &\ge \sqrt{\frac{3R}{4}} - \|\!|(0, -x, -v) \circ (0, x, v + a\widetilde{v} - {}^{a\widetilde{x}}/\widetilde{t})\|\!| \\ &\ge \sqrt{\frac{3R}{4}} - \|\!|(0, 0, a\widetilde{v} - {}^{a\widetilde{x}}/\widetilde{t})\|\!| = \sqrt{\frac{3R}{4}} - a\|\!|(0, 0, \widetilde{v} - {}^{\widetilde{x}}/\widetilde{t})\|\!|. \end{aligned}$$

The conclusion then follows by using that $|a| \le \sqrt{\varepsilon R}$ and the choice (4.48) of support of ψ, which makes $\|\!|(0, 0, \widetilde{v} - {}^{\widetilde{x}}/\widetilde{t})\|\!| \lesssim 1$ uniformly over the support of ψ.

A bound of the type (4.54) then follows from applying Young's inequality to in the last line of (4.56).

Step four: bound I_{22}. This is the most difficult term as it involves because it requires arguing by the H_v^{-1}-H_v^1 pairing. To access this, we begin by directly computing the b derivative in (4.55), changing the order of integration, and then changing variables:

$$\begin{aligned} I_{22} &= 2\int_{T_1}^{T_2}\int \chi_R(z) \int_0^{\varepsilon R}\int g_{\sqrt{\varepsilon R}, b, 0}(z; \widetilde{z}) \partial_b g_{\sqrt{\varepsilon R}, b, 0}(z; \widetilde{z}) \psi(\widetilde{z})\, d\widetilde{z}\, da\, dz \\ &= 2\int_0^{\varepsilon R}\int \widetilde{t}\psi(\widetilde{z}) \int_{T_1}^{T_2}\int \chi_R(z) g\left[t - b\widetilde{t}, x - b\widetilde{t}\left(v + \sqrt{\varepsilon R}(-\widetilde{v} + {}^{\widetilde{x}}/\widetilde{t})\right), v + \sqrt{\varepsilon R}(-\widetilde{v} + {}^{\widetilde{x}}/\widetilde{t})\right] \\ &\qquad \times (Yu)\left[t - b\widetilde{t}, x - b\widetilde{t}\left(v + \sqrt{\varepsilon R}(-\widetilde{v} + {}^{\widetilde{x}}/\widetilde{t})\right), v + \sqrt{\varepsilon R}(-\widetilde{v} + {}^{\widetilde{x}}/\widetilde{t})\right]\, dz\, d\widetilde{z}\, da \\ &= 2\int_0^{\varepsilon R}\int \widetilde{t}\psi(\widetilde{z}) \int_{T_1 - b\widetilde{t}}^{T_2 - b\widetilde{t}}\int \chi_R\left[x + b\widetilde{t}\left(v + \sqrt{\varepsilon R}(-\widetilde{v} + {}^{\widetilde{x}}/\widetilde{t})\right), v + \sqrt{\varepsilon R}(-\widetilde{v} + {}^{\widetilde{x}}/\widetilde{t})\right] \\ &\qquad\qquad g(z)(Yu)(z)\, dz\, d\widetilde{z}\, da. \end{aligned}$$

We momentarily simplify the notation for the cutoff term χ_R, letting

$$\chi_R(x, v; b; \widetilde{z}) = \chi_R\left(x + b\widetilde{t}\left(v + \sqrt{\varepsilon R}(-\widetilde{v} + {}^{\widetilde{x}}/\widetilde{t})\right), v + \sqrt{\varepsilon R}(-\widetilde{v} + {}^{\widetilde{x}}/\widetilde{t})\right).$$

Next, arguing as in the justification of (4.57), we see that

$$[T_1 - b\widetilde{t}, T_2 - b\widetilde{t}] \times \operatorname{supp}_{x,v} \chi_R(\cdot, \cdot; b; \widetilde{z}) \subset \Omega_R,$$

for any $\widetilde{z} \in \operatorname{supp}\psi$ and any $b \in [0, \varepsilon R]$. We recall the choice of support of ψ in (4.48). Hence, using the H_v^{-1}-H_v^1 pairing, we have

$$I_{22} \lesssim \int_0^{\varepsilon R} \int \psi(\widetilde{z}) \, \|\nabla_v \left(g \, \chi_R(\cdot, \cdot; b; \widetilde{z})\right)\|_{L^2(\Omega_R)} \, [\![g]\!]_{H^1_{\text{kin}}(\Omega_R)} \, d\widetilde{z} \, da.$$

A direct computation using (4.51), it is easy to see that

$$\begin{aligned}\|\nabla_v \left(g \, \chi_R(\cdot, \cdot; b; \widetilde{z})\right)\|_{L^2(\Omega_R)} &\lesssim \|\nabla_v g\|_{L^2(\Omega_R)} + \left(\frac{b}{R^{3/2}} + \frac{1}{\sqrt{R}}\right) \|g\|_{L^2(\Omega_R)} \\ &\lesssim [\![g]\!]_{H^1_{\text{kin}}(\Omega_R)} + \left(\frac{b}{R^{3/2}} + \frac{1}{\sqrt{R}}\right) \|g\|_{L^2(\Omega_R)} \, .\end{aligned}$$

Hence,

$$\begin{aligned}I_{22} &\lesssim \int_0^{\varepsilon R} \int \psi(\widetilde{z}) \left([\![g]\!]_{H^1_{\text{kin}}(\Omega_R)} + \left(\frac{b}{R^{3/2}} + \frac{1}{\sqrt{R}}\right) \|g\|_{L^2(\Omega_R)} \right) [\![g]\!]_{H^1_{\text{kin}}(\Omega_R)} \, d\widetilde{z} \, db \\ &\lesssim \varepsilon R [\![g]\!]_{H^1_{\text{kin}}(\Omega_R)} + \sqrt{\varepsilon R} \, \|g\|_{L^2(\Omega_R)} \, [\![u]\!]_{H^1_{\text{kin}}(\Omega_R)}.\end{aligned}$$

Again, the proof is then finished after applying Young's inequality.

Step five: bound I_{23}. The proof of this is exactly the same as the proof of the bound of I_{21} in step two. Indeed, this is only a shift in v, which is the simplest case. As such, we omit it. This concludes the proof of (4.48) and, thus, Proposition 4.5.2. $\square$

The proof of the Poincaré-type inequality on $\mathcal{O}_R$

While $\mathcal{P}_R$ and $\mathcal{N}_R$ are, respectively, increasing and decreasing in R, $\mathcal{O}_R$ is not monotonic in R. This monotonicity was useful in constructing cutoff functions. In this case, we must define, for any $R' > 0$,

$$\Theta_{R'} = \{(x, v) \in (\mathbb{H}^d \times \mathbb{R}^d) \setminus \mathcal{P}_{R'/2} : \operatorname{dist}(\mathbb{R} \times \gamma_+) \leq \sqrt{R'/10}\}.$$

In the proof of Proposition 4.5.3, we use a cutoff function on that is one on $\mathcal{O}_R$ and zero on Θ^c_{2R}. A key aspect of the proof is working with the ratio $x_1/|v_1|$, so we state a lemma to bound that now. The proof is postponed to Section 4.7.

Lemma 4.5.5. *Fix any $(x, v) \in \Theta_R$. Then*

$$\frac{x_1}{|v_1|} \leq \frac{R}{4}.$$

For the proof of Proposition 4.5.3, let us make the convention that every norm is taken on $[T_1, T_2 + R] \times \mathbb{H}^d \times \mathbb{R}^d$ unless otherwise specified. For example, by writing

$$\|g\|_{L^2} \quad \text{we mean} \quad \|g\|_{L^2([T_1,T_2+R]\times\mathbb{H}^d\times\mathbb{R}^d)}.$$

This saves significant space and does not cost clarity.

Proof of Proposition 4.5.3. We extend g to $\overline{g}$ on $[T_1, T_2 + R] \times \mathbb{R}^d \times (\mathbb{R}_- \times \mathbb{R}^{d-1})$ as follows:

$$\overline{g}(t,x,v) = \begin{cases} f(t,x,v) & \text{if } x_1 \geq 0, \\ f(t - {}^{x_1}/_{v_1}, (0,\overline{x}), v) & \text{if } x_1 < 0, \end{cases}$$

where $\overline{x} \coloneqq (x_2, \dots, x_d)$. Notice that

$$Y\overline{f} = 0 \qquad \text{for } x_1 < 0. \tag{4.58}$$

Next, we take a cutoff function $\chi_{\mathcal{O}_R}$ such that

$$\chi_{\mathcal{O}_R} \equiv 1 \quad \text{in } \mathcal{O}_R, \qquad \text{and} \qquad \chi_{\mathcal{O}_R} \equiv 0 \quad \text{in } \Theta_{2R},$$

while

$$\|\nabla_v \chi_{\mathcal{O}_R}\|_{L^\infty} \lesssim \frac{1}{\sqrt{R}} \quad \text{and} \quad \|\nabla_x \chi_{\mathcal{O}_R}\|_{L^\infty} \lesssim \frac{1}{R^{3/2}}. \tag{4.59}$$

This can be constructed easily using the scaling properties in (4.29); see the discussion around (4.42)-(4.43).

By the fundamental theorem of calculus we have that, for any $z \in [T_1, T_2] \times \Theta_{2R}$,

$$\begin{aligned} g(t,x,v)^2 &= -2\int_0^{\frac{x_1}{|v_1|}} g(t+r, x+rv, v) Y g(t+r, x+rv, v)\, dr + g(t + {}^{x_1}/_{|v_1|}, (0,\overline{x}), v)^2 \\ &= -2\int_0^R \overline{g}(t+r, x+rv, v) Y \overline{g}(t+r, x+rv, v)\, dr + g(t + {}^{x_1}/_{|v_1|}, (0,\overline{x}), v)^2, \end{aligned} \tag{4.60}$$

where the second equality follows from (4.58) and Lemma 4.5.4. Therefore,

$$\begin{aligned} \int_{T_1}^{T_2}\int_{\mathcal{O}_R} g(z)^2\, dz \leq& \int_{T_1}^{T_2}\int_{\mathbb{H}^d\times\mathbb{R}^d} \chi_{\mathcal{O}_R} g(z)^2\, dz \\ =& -2\int_{T_1}^{T_2}\int_{\mathbb{H}^d\times\mathbb{R}^d}\int_0^R g(t+r, x+rv, v) Y\overline{g}(t+r, x+rv, v)\, dr\, dx\, dv\, dt \\ &+ \int_{T_1}^{T_2}\int_{\mathbb{H}^d\times\mathbb{R}^d} \chi_{\mathcal{O}_R}(x,v) f(t + {}^{x_1}/_{|v_1|}, (0,\overline{x}), v)^2\, dx\, dv\, dt = I_1 + I_2, \end{aligned}$$

where we passed from the first to the second line rewriting $f(t,x,v)^2$ according to (4.60).

We first see, by a change of variables, that

$$\begin{aligned} I_1 =& -2\int_0^R \int_{T_1+r}^{T_2+r} \int_{\mathbb{R}^d} \int_{\mathbb{R}_- \times \mathbb{R}^{d-1}} \chi_{\mathcal{O}_R}(x - rv, v) \overline{g}(t,x,v) Y \overline{g}(t,x,v)\, dx\, dv\, dt\, dr \\ =& -2\int_0^R \int_{T_1+r}^{T_2+r} \int_{\mathbb{H}^d \times \mathbb{R}^d} \chi_{\mathcal{O}_R}(x - rv, v) g(t,x,v) Y g(t,x,v)\, dx\, dv\, dt\, dr. \end{aligned}$$

In the second equality, we used (4.58) again to reduce the domain of the integral.

Let $\bar{\chi}_{\mathcal{O}_R}(r,x,v) = \chi_{\mathcal{O}_R}(x - rv, v)$. Then, by the H_v^{-1}-H_v^1 pairing, we get

$$\begin{aligned} |I_1| \lesssim& \int_0^R \|\nabla_v \left(\bar{\chi}_{\mathcal{O}_R} g\right)\|_{L^2([T_1+r,T_2+r]\times\mathbb{H}^d\times\mathbb{R}^d)} [\![g]\!]_{H^1_{\text{kin}}([T_1+r,T_2+r]\times\mathbb{H}^d\times\mathbb{R}^d)} dr \\ \lesssim& [\![f]\!]_{H^1_{\text{kin}}} \int_0^R \left(\left(\frac{1}{\sqrt{R}} + \frac{r}{R^{3/2}} \right) \|g\|_{L^2} + \|\chi_{\mathcal{O}_R} \nabla g\|_{L^2} \right) dr \\ \lesssim& \sqrt{R} [\![g]\!]_{H^1_{\text{kin}}} \|f\|_{L^2} + R [\![g]\!]^2_{H^1_{\text{kin}}}. \end{aligned}$$

In the second line follows from estimates (4.59).

We now estimate I_2. Changing variables yields

$$\begin{aligned} I_2 =& \int_{\mathbb{H}^d\times\mathbb{R}^d} \chi_{\mathcal{O}_R}(x,v) \int_{T_1+x_1/|v_1|}^{T_2+x_1/|v_1|} g(t,(0,\overline{x}),v)^2\, dt\, dx\, dv \\ \le& \int_{\mathbb{H}^d\times\mathbb{R}^d} \chi_{\mathcal{O}_R}(x,v) \int_{T_1}^{T_2+R} g(t,(0,\overline{x}),v)^2\, dt\, dx\, dv \\ =& \int_{T_1}^{T_2+R} \int_{\mathbb{R}^d} \int_{\mathbb{R}^{d-1}} \chi_{\mathcal{O}_R}(x,v) x_1 g(t,(0,\overline{x}),v)^2\, d\overline{x}\, dv\, dt, \end{aligned}$$

where we used Lemma 4.5.5 and the fact that the integrand is positive to obtain the second line.

Applying once again Lemma 4.5.5, we find

$$I_2 \le \frac{R}{4} \int_{T_1}^{T_2+R} \int_{\mathbb{R}^d} \int_{\mathbb{R}_-\times\mathbb{R}^{d-1}} |v_1| g(t,(0,\overline{x}),v)^2 d\overline{x}\, dv\, dt.$$

This concludes the proof. □

4.6 Controlling f on the isolated region: Lemma 4.4.3

We begin with a lemma that is simple to prove. It essentially says if a point (x,v) is distance ρ to the boundary, then the path from (x,v) to the boundary that simply follows transport (without any changes in velocity) has to take at least time ρ.

Lemma 4.6.1. *If* $(x, v) \in \mathcal{N}_R$, *then*

$$x_1 \geq \frac{R}{10}|v_1|.$$

Proof. This is trivially true if $v_1 = 0$, so we assume that $v_1 \neq 0$. Recall from (4.28), that

$$\operatorname{dist}(\mathbb{R} \times \partial\mathbb{H}^d \times \mathbb{R}^d, (0, x, v)) \geq \sqrt{R/10}.$$

In view of (4.25), it follows that

$$\sqrt{\tfrac{R}{10}} \leq |||\zeta||| \qquad \text{where } \zeta = (\tau, \xi, \omega) = \left(\frac{x_1}{v_1}, (-x_1 - \tau v_1, 0), 0\right).$$

Unpacking the definition of $|||\cdot|||$ with the choice $w = 0$, we deduce that

$$\left(\frac{R}{10}\right)^{1/2} \leq |\tau|^{1/2},$$

which is precisely the claim. □

The proof of Lemma 4.4.3 is fairly straightforward, if tedious.

Proof of Lemma 4.4.3. To simplify the notation in this proof, let us set

$$\overline{R} = \frac{R}{10}.$$

For ψ and E to be chosen, we define

$$\widetilde{\mu}_R(x, v) = E\left(-\frac{\overline{R}v_1}{x_1}\right)\psi\left(-\frac{v_1^3}{x_1}\right)\psi\left(-\frac{2v_1}{\sqrt{\overline{R}}} - 1\right). \tag{4.61}$$

Below, it will be helpful to suppress the arguments, while keeping track of the three individual functions that make up $\widetilde{\mu}_R$. To this end, we write

$$\widetilde{\mu}_R(x, v) = E\psi\hat{\psi},$$

where ψ is shorthand for $\psi(-v_1^3/x_1)$ and $\hat{\psi}$ is shorthand for $\psi(-2v_1/\sqrt{\overline{R}} - 1)$.

To aid the reader, let us stress that, in all nontrivial cases in this proof, $v_1 < 0$. Thus, $-v_1$, $-v_1^3$, etc. are *positive* quantities.

Take E to be a decreasing function

$$E(\rho) = 1 \quad \text{if } \rho \leq 1, \qquad \text{and} \qquad E(\rho) \approx e^{-\rho} \quad \text{if } \rho \geq 0$$

such that

$$E'' + E' \leq 0 \qquad \text{for all } \rho \geq 0. \tag{4.62}$$

Moreover, we may take E such that

$$|E''|, |E'| \lesssim E.$$

Roughly, E is a mollification of $\min\{1, e^{1-\rho}\}$. This is somewhat simple to construct, so we omit its proof. Additionally, we let ψ be any increasing function such that

$$\psi(\rho) = \begin{cases} 1 & \text{if } \rho \geq 1, \\ 0 & \text{if } \rho \leq 1/2. \end{cases}$$

Let us first check that, with these choices, $\widetilde{\mu}_R$ satisfies (4.30). This is straightforward, except for the first case when $(x, v) \in \mathcal{N}_R \cap \{x_1 \leq -v_1^3\}$. Clearly ψ satisfies the correct bound. From Lemma 4.6.1, we have that $x_1 \leq \overline{R}|v_1|$. This implies that

$$E\left(-v_1\overline{R}/x_1\right) \geq E(1) \approx 1.$$

Finally, we notice that

$$\overline{R} \leq \frac{x_1}{|v_1|} \leq \frac{|v_1|^3}{|v_1|} = |v_1|^2.$$

It follows that $\hat{\psi} = 1$.

We now need to show the main estimate (4.31). This proceeds by considering each of the subdomains on which μ_R is nonzero one at a time. Given its definition (4.61), there are eight cases to check: there are three functions, each having one region where it takes the constant value one and one region where it varies.

 # Case one:

$$\frac{-\overline{R}v_1}{x_1} \geq 1, \quad \frac{-v_1^3}{x_1} \geq 1, \quad \text{and} \quad \frac{2v_1}{\sqrt{\overline{R}}} - 1 \geq 1.$$

Let us note that the last inequality yields

$$v_1^2 \geq \overline{R}. \tag{4.63}$$

In this case, letting $\rho = \frac{-\overline{R}v_1}{x_1}$,

$$\begin{aligned}\Delta_v \chi + v \cdot \nabla_x \widetilde{\mu}_R &= \frac{\overline{R}^2}{x_1^2} E'' + \frac{\overline{R}v_1^2}{x_1^2} E' = \frac{z^2}{v_1^2}\left(E'' + \frac{v_1^2}{\overline{R}} E'\right) \\ &\leq \frac{z^2}{v_1^2}\left(E'' + E'\right) \leq 0.\end{aligned}$$

The second-to-last inequality follows by (4.63) and the fact that $E' \leq 0$. The last inequality follows by (4.62). This is clearly (4.31) in this case.

 # Case two:

$$\frac{-\overline{R}v_1}{x_1} \geq 1, \quad \frac{-v_1^3}{x_1} \geq 1, \quad \text{and} \quad \frac{-2v_1}{\sqrt{\overline{R}}} - 1 \in (1/2, 1). \tag{4.64}$$

In this case,

$$\begin{aligned}(\Delta_v + v \cdot \nabla_x)\,\widetilde{\mu}_R &= \left(\frac{\overline{R}^2}{x_1^2} E''\hat{\psi} + \frac{4\sqrt{\overline{R}}}{x_1} E'\hat{\psi}' + \frac{4}{\overline{R}} E\hat{\psi}''\right) + \left(\frac{v_1^2 \overline{R}}{x_1^2} E'\hat{\psi}\right) \\ &\leq \frac{\overline{R}^2}{x_1^2} E''\hat{\psi} + \frac{4}{\overline{R}} E\hat{\psi}''.\end{aligned} \tag{4.65}$$

In the inequality, we used that E is decreasing, while $\hat{\psi}$ is increasing.

Notice that for any $\varepsilon > 0$ and $z \geq 0$,

$$1 \lesssim \frac{x_1}{|v_1|^3} e^{-\frac{\varepsilon v_1^3}{x_1}} \quad \text{and} \quad |E'(\rho)|, |E''(\rho)|, E(\rho) \lesssim e^{-\rho}.$$

We look only at the first term in (4.65); however, the second term is handled similarly. Then,

$$\begin{aligned}\frac{R^2}{x_1^2} E''\hat{\psi} &\lesssim \frac{\overline{R}^2}{x_1^2} e^{\frac{v_1 \overline{R}}{x_1}} \lesssim \frac{\overline{R}^2}{x_1^2} e^{\frac{v_1 \overline{R}}{x_1}} \left(\frac{x_1^3}{|v_1|^9} e^{-\frac{\varepsilon v_1^3}{x_1}}\right) \\ &= \frac{\overline{R}^{13/4}}{\overline{R}^{5/4}} \frac{x_1}{|v_1|^9} e^{\frac{v_1 \overline{R}}{x_1}} e^{-\frac{\varepsilon v_1^3}{x_1}} \leq \frac{\overline{R}^{13/4}}{\overline{R}^{5/4}} \frac{x_1}{|v_1|^9} e^{\frac{v_1^3}{x_1}} e^{-\frac{\varepsilon v_1^3}{x_1}} \\ &\approx \frac{\overline{R}^{13/4}}{\overline{R}^{5/4}} \frac{x_1}{\overline{R}^{13/4} |v_1|^{5/2}} e^{\frac{v_1^3}{x_1}} e^{-\frac{\varepsilon v_1^3}{x_1}} \lesssim \frac{1}{\overline{R}^{5/4}} \widetilde{\varphi}(x, v).\end{aligned}$$

In the inequality on the second line and in "≈" on the last line, we used that, by the third item in (4.64),

$$\frac{3\sqrt{\overline{R}}}{4} \leq -v_1 \leq \sqrt{\overline{R}}.$$

This is clearly (4.31) in this case.

 # Case three:

$$-\frac{v_1 \overline{R}}{x_1} \geq 1, \quad -\frac{v_1^3}{x_1} \in (1/2, 1) \quad \text{and} \quad -\frac{2v_1}{\sqrt{\overline{R}}} - 1 \geq 1.$$

We claim that this case cannot happen. Indeed, the first and third inequalities above yield

$$x_1 \leq -v_1 \overline{R} \leq -v_1^3,$$

while the second implies that

$$x_1 > -v_1^3.$$

This is a contradiction. Hence, case three cannot occur.

 # Case four:

$$-\frac{v_1\overline{R}}{x_1} \geq 1, \quad -\frac{v_1^3}{x_1} \in (1/2, 1), \quad \text{and} \quad -\frac{2v_1}{\sqrt{\overline{R}}} - 1 \in (1/2, 1).$$

This case involves the most derivatives since all three cutoff functions are varying. That said, it is fundamentally the same as case two, while being slightly easier because

$$x_1 \approx -v_1^3 \approx \overline{R}^{3/2} \quad \text{and} \quad \widetilde{\varphi}(x,v) \approx \sqrt{-v_1}.$$

As such, we omit its proof.

 # Case five:

$$-\frac{v_1\overline{R}}{x_1} < 1, \quad -\frac{v_1^3}{x_1} \geq 1, \quad \text{and} \quad -\frac{2v_1}{\sqrt{\overline{R}}} - 1 \geq 1.$$

This case is precisely when $\widetilde{\mu}_R \equiv 1$. Hence

$$(\Delta_v + v \cdot \nabla_x)\widetilde{\mu}_R = 0,$$

which clearly yields (4.31) in this case.

 # Case six:

$$-\frac{v_1\overline{R}}{x_1} < 1, \quad -\frac{v_1^3}{x_1} \geq 1 \quad \text{and} \quad -\frac{2v_1}{\sqrt{\overline{R}}} - 1 \in (1/2, 1). \tag{4.66}$$

This case cannot occur. Indeed, the first and second inequalities in (4.66) imply that

$$-v_1\overline{R} < x_1 \leq -v_1^3,$$

which implies that $v_1^2 > R$. On the other hand, the last inequality in (4.66) implies that

$$-v_1 < \sqrt{\overline{R}}.$$

These are in contradiction (recall that $-v_1 \geq 0$).

 # Case seven:

$$-\frac{v_1\overline{R}}{x_1} < 1, \quad -\frac{v_1^3}{x_1} \in (1/2, 1) \quad \text{and} \quad -\frac{2v_1}{\sqrt{\overline{R}}} - 1 \geq 1.$$

From the first and second inequalities, we see that

$$x_1 \approx -v_1^3 \gtrsim \overline{R}^{3/2}.$$

Hence,

$$\widetilde{\varphi}(x,v) \approx \sqrt{-v_1}.$$

We also notice that ψ is the only term in $\widetilde{\mu}_R$ not equal to 1 on this domain. Hence,

$$\begin{aligned}(\Delta_v + v\cdot\nabla_x)\,\widetilde{\mu}_R &= \frac{9v^4}{x^2}\psi'' - \frac{6v}{x}\psi' + \frac{v^4}{x^2}\psi' \\ &\lesssim \frac{1}{v_1^2} \lesssim \frac{\sqrt{-v_1}}{\overline{R}^{5/4}} \approx \frac{\widetilde{\varphi}(x,v)}{\overline{R}^{5/4}}.\end{aligned}$$

Thus, we have established (4.31).

 # Case eight:

$$-\frac{v_1\overline{R}}{x_1} < 1, \quad -\frac{v_1^3}{x_1} \in (1/2, 1), \quad \text{and} \quad \frac{2v_1}{\sqrt{\overline{R}}} - 1 \in (1/2, 1).$$

In this case, we have

$$x_1 \approx -v_1^3 \approx \overline{R}^{3/2}.$$

From this point, the proof is essentially the same as in the previous case, and, thus, is skipped. This completes the proof of Lemma 4.4.3. □

4.7 Other technical lemmas

4.7.1 Understanding (x,v) in $\mathcal{P}_R$

Proof of Lemma 4.5.4. Let $z = (0,x,v)$ for ease. It is easy to see that the infimum in (4.25) is attained, up to including the boundary $(\{0\}\times\mathbb{R}^{d-1})^2$, so we fix $\zeta = (\tau,\xi,\omega)$ such that

$$\begin{aligned}&z\circ\zeta \in \mathbb{R}\times\overline{\gamma}_- \quad \text{and} \\ &\|\!|\zeta|\!\| = \operatorname{dist}(\mathbb{R}\times\gamma_-, z) \le \sqrt{R}.\end{aligned} \tag{4.67}$$

Since $z\circ\zeta = (\tau, x+\xi+\tau v, v+\omega)$, then we have the constraints

$$0 = x_1 + \xi_1 + \tau v_1 \quad \text{and} \quad v_1 + \omega_1 > 0. \tag{4.68}$$

It is clear from the second inequality in (4.67), as well as the definition (4.23) of $|||\cdot|||$, that

$$|\omega_1| \leq 2\sqrt{R},$$

This can be seen by noting that

$$\sqrt{R} \geq \min_w \max\{|\omega - w|, |w|\} \geq \frac{|\omega|}{2}. \tag{4.69}$$

Hence, the second line of (4.68) yields

$$-v_1 < \omega_1 \leq 2\sqrt{R},$$

as desired.

Next, notice that the definition (4.23) of $|||\cdot|||$ and the second line of (4.67) implies that $|\tau| \leq R$. Hence, if

$$|\xi_1| \leq 3R^{3/2} \tag{4.70}$$

then we immediately have

$$x_1 = -\xi_1 - \tau v_1 \lesssim \max\{3R^{3/2}, Rv_1\},$$

as desired.

To see (4.70), let w be the minimizer in the definition (4.23) of $|||\cdot|||$. Then, by arguing exactly as in (4.69), we find $|w| \leq 2\sqrt{R}$. We deduce that

$$R^{3/2} \geq |\xi - \tau w| \geq |\xi| - 2R^{3/2},$$

from which (4.70) follows. This concludes the proof. □

4.7.2 Understanding (x, v) in $\mathcal{O}_R$

Proof of Lemma 4.5.5. By the symmetry of γ_- and γ_+, we immediately see that

$$x_1 \leq \frac{R}{10} \max\{-v_1, 3\sqrt{R/10}\} \tag{4.71}$$

since $\operatorname{dist}(\mathbb{R} \times \gamma_-, (0, x, v)) \leq \sqrt{R/10}$; see Lemma 4.5.4. This is useful in the sequel.

If $-v_1 \geq \sqrt{R/2}$, then we find

$$\frac{x_1}{|v_1|} \leq \begin{cases} \frac{3R}{10^{3/2}\sqrt{1/2}} & \text{if } -v_1 \leq 3\sqrt{R/10}, \\ \frac{R}{10} & \text{if } -v_1 \geq 3\sqrt{R/10}. \end{cases}$$

Hence,

$$\frac{x_1}{|v_1|} \leq \frac{R}{5},$$

and the proof is finished in this case.

Next consider when

$$|v_1| = -v_1 < \sqrt{R/2}. \tag{4.72}$$

In view of (4.71), this yields

$$x_1 \leq \frac{3R^{3/2}}{10^{3/2}} < \frac{R^{3/2}}{2^{3/2}}. \tag{4.73}$$

By definition, we find

$$\operatorname{dist}(\mathbb{R} \times \gamma_-, (0, x, v)) \geq \sqrt{R/2}.$$

Letting

$$\zeta = \left(\tfrac{R}{4}, \left(-x_1 - \tfrac{R}{4}v_1, 0\right), (-v_1, 0)\right)$$

we have that $z \circ \zeta \in \overline{\mathbb{R} \times \gamma_-}$ and, hence,

$$|||\zeta||| \geq \sqrt{R/2}.$$

Taking $w = 0$ in the definition (4.23) of $|||\cdot|||$, we find

$$\max\left\{|\tfrac{R}{4}|^{1/2}, \left|x_1 + \tfrac{Rv_1}{4}\right|^{1/3}, |0|, |v_1 - 0|\right\} \geq \sqrt{R/2}.$$

By (4.72), it follows that

$$\left(\frac{R}{2}\right)^{3/2} \leq \left|x_1 + \frac{Rv_1}{4}\right|.$$

If the term in the absolute value is negative, we find

$$x_1 < -\frac{Rv_1}{4} = \frac{R|v_1|}{4}$$

from which the conclusion follows. If the term in absolute value is nonnegative, we find

$$\left(\frac{R}{2}\right)^{3/2} + \frac{R|v_1|}{4} \leq x_1.$$

This contradicts (4.73). The proof is concluded. □

4.7.3 Understanding (x, v) in $\mathcal{N}_R$

Proof of Lemma 4.4.2. Fix $\varepsilon > 0$ to be chosen. Let us first consider the case where $|v_1| \geq \varepsilon\sqrt{R}$. Applying Lemma 4.6.1, we see that

$$x_1 \geq \frac{R|v_1|}{10} \geq \frac{\varepsilon R^{3/2}}{10}. \tag{4.74}$$

If $x_1 \geq |v_1|^3$ then, by (4.2),

$$\widetilde{\varphi}(x, v) \approx x_1^{1/6} \gtrsim R^{1/4}.$$

We used (4.74) in the last inequality.

If $x_1 < |v_1|^3$, then, applying the assumption $v_1 > 0$ and, by (4.2),

$$\widetilde{\varphi}(x, v) \approx \sqrt{v_1} \gtrsim R^{1/4}.$$

This finishes the proof in this case.

Now we consider the case $|v_1| \leq \varepsilon\sqrt{R}$. Recall from (4.28), that

$$\operatorname{dist}(\mathbb{R}_+ \times \partial\mathbb{H}^d \times \mathbb{R}^d, (0, x, v)) \geq \sqrt{\tfrac{R}{10}}.$$

Let

$$\zeta = (\tau, \xi, 0) = (R/20, (-x_1 - \tau v_1, 0), 0),$$

and notice that $0 = x_1 + \xi_1 + \tau v_1$. Hence,

$$\sqrt{\tfrac{R}{10}} \leq \operatorname{dist}(\mathbb{R}_+ \times \partial\mathbb{H}^d \times \mathbb{R}^d, (0, x, v)) \leq |||\zeta|||.$$

Taking $w = 0$ in the definition (4.23) of $|||\cdot|||$, we find

$$\sqrt{\tfrac{R}{10}} \leq \max\left\{|R/20|^{1/2}, |x_1 + \tau v_1|^{1/3}, |0|, |0|\right\}.$$

It follows that

$$\sqrt{\tfrac{R}{10}} \leq |x_1 + \tau v_1|^{1/3}.$$

Rearranging this, we find

$$\frac{R^3}{10^3} - \frac{R}{20}|v_1| \leq x_1.$$

Using that $|v_1| \leq \varepsilon\sqrt{R}$ and possibly decreasing ε, we deduce that

$$\frac{R^3}{10^4} \leq x_1.$$

Further decreasing ε, if necessary, we see that $|v_1|^2 \leq x_1$, whence

$$\widetilde{\varphi}(x, v) \approx x_1^{1/6} \gtrsim R^{1/4}.$$

Here, we once again used the asymptotics of $\widetilde{\varphi}$ given in (4.2). This concludes the proof. $\square$

4.8 The whole space case: Corollary 4.1.3

The proof of Theorem 4.1.2 follows exactly the outline of the proof of Proposition 4.5.2. The only modification to be made is to take the cutoff function ψ in (4.47)-(4.48) to be supported on B. As such, we omit the proof.

In this section, we provide a brief outline of the proof of the whole-space time decay. The work here is similar to, but much simpler than, the proof of Theorem 4.1.1.

Proof of Corollary 4.1.3. Let us note that

$$\frac{d}{dt}\int f\,dx\,dv = \int (\nabla \cdot (a\nabla_v f) - v\cdot\nabla_x f)\,dx\,dv = 0.$$

Hence, the L^1-norm is conserved:

$$\|f(t,\cdot,\cdot)\|_{L^1(\mathbb{R}^{2d})} = \int f_{\text{in}}\,dx\,dv. \tag{4.75}$$

The main step is obtaining an $L^1 \to L^2$ bound on the solution operator $S_t f_{\text{in}} = f(t)$. We do this by combining the Nash inequality with the energy equality.

Let us begin with the energy equality. Multiplying (4.9) by f, integrating, and then integrating by parts, we obtain, for any $0 \le t' \le t$,

$$E(t) + \int_{t'}^{t} D(s)\,ds \lesssim E(t) + \int_{t'}^{t}\int \nabla_v f a \nabla_v f\,dz \le E(t'), \tag{4.76}$$

where we borrow the notation for E and D from the proof of Theorem 4.1.1 (see (4.33)). Let us point out that, as a result of (4.76), E is decreasing in time.

As in the proof of Theorem 4.1.1, we note that

$$\int_{t'}^{t} D(s)\,ds \approx [\![f]\!]^2_{H^1_{\text{kin}}([t',t]\times\mathbb{R}^{2d})}$$

(recall (4.19)).

Now we introduce the Nash inequality to control the quantities in (4.76). Indeed, applying Theorem 4.1.2 with the choices $s_0 = \sqrt{t/4}$, $s = \sqrt{\varepsilon t}$ for $\varepsilon \in (0, 1/4)$ to be chosen,

$$\Omega_1 = [t/2, t] \times \mathbb{R}^{2d}, \quad \Omega_2 = [t/4, t]\times\mathbb{R}^{2d}, \quad \text{and} \quad B = \{z \in (0,1]\times\mathbb{R}^{2d} : \text{dist}(z,0) \le 1\},$$

we deduce that

$$\int_{t/2}^{t} E(s)\,ds \lesssim \frac{\varepsilon t}{\delta}\int_{t/4}^{t} D(s)\,ds + \delta\int_{t/4}^{t} E(s)\,ds + \frac{t^2}{(\varepsilon t)^{2d+1}}\Big(\int f_{\text{in}}\,dx\,dv\Big)^2, \tag{4.77}$$

where $\delta > 0$ is a parameter to be chosen and where we applied (4.75).

Fix $\overline{\alpha} > 0$ to be chosen and let

$$\alpha = \overline{\alpha}\Big(\int f_{\text{in}}\,dx\,dv\Big)^2.$$

Define

$$t_0 = \sup\Big\{t : E(s) \le \frac{\alpha}{s^{2d}} \quad \text{for all } s \in (0,t]\Big\}.$$

Up to approximation, we may assume that f_{in} is smooth and compactly supported, whence $t_0 > 0$. Our goal is to show that $t_0 = \infty$. Hence, we argue by contradiction assuming that t_0 is finite.

Clearly $E(t_0) = \alpha t_0^{-2d}$ and $E(s) \le \alpha s^{-2d}$ for all $s \le t_0$. Using this in (4.77) and recalling that E is decreasing in time yields

$$\frac{\alpha}{t_0^{2d-1}} = t_0 E(t_0) \lesssim \frac{\varepsilon t_0}{\delta}\int_{t_0/4}^{t_0} D(s)\,ds + \frac{\delta\alpha}{t_0^{2d-1}} + \frac{\alpha}{\bar{\alpha}\varepsilon^{2d+1}(\varepsilon t_0)^{2d-1}}. \tag{4.78}$$

Using (4.76) and the definition of t_0, we have

$$\int_{t_0/4}^{t_0} D(s)\,ds \le E({}^{t_0}/_4) \le \frac{4^{2d}\alpha}{t_0^{2d}}.$$

Including this in (4.78), we find

$$\frac{\alpha}{t_0^{2d-1}} \lesssim \frac{\varepsilon}{\delta}\frac{\alpha}{t_0^{2d-1}} + \frac{\delta\alpha}{t_0^{2d-1}} + \frac{\alpha}{\bar{\alpha}\varepsilon^{2d+1}(\varepsilon t_0)^{2d-1}}.$$

This is clearly a contradiction after choosing, in order, δ and ϵ small and $\overline{\alpha}$ large. It follows that $t_0 = \infty$.

The rest of the proof is simple functional analysis. From the fact that $t_0 = \infty$, we deduce that

$$\int_{\mathbb{R}^{2d}} f(t,x,v)^2\,dx\,dv \lesssim \frac{1}{t^{2d}}\int f_{\text{in}}\,dx\,dv.$$

This can be rephrased as

$$\|S_t\|_{L^1\to L^2} \lesssim \frac{1}{t^d}.$$

Of course, the same inequality follows for the solution operator $\widetilde{S}_t$ adjoint equation

$$(\partial_t - v\cdot\nabla_x)f = \nabla_v\cdot(a\nabla_v f) \qquad \text{in } \mathbb{R}_+\times\mathbb{R}^{2d}.$$

Hence, the operator $\widetilde{S}_t^* : L^2 \to L^\infty$, which is again a solution operator of (4.9), satisfies

$$\|\widetilde{S}_t^*\|_{L^2\to L^\infty} \lesssim \frac{1}{t^d}.$$

Writing $f(t) = \widetilde{S}_{t/2}^* \widetilde{S}_{t/2}^* f_{\text{in}}$, we deduce

$$\|f(t)\|_{L^\infty(\mathbb{R}^{2d})} = \|\widetilde{S}_{t/2}^* S_{t/2} f_{\text{in}}\|_{L^\infty(\mathbb{R}^{2d})} \lesssim \frac{1}{t^d}\|S_{t/2} f_{\text{in}}\|_{L^2(\mathbb{R}^{2d})} \lesssim \frac{1}{t^d}\frac{1}{t^d}\|f_{\text{in}}\|_{L^1(\mathbb{R}^{2d})},$$

which concludes the proof. □

4.A A kinetic version of Young's convolution inequality

Let us note that the main inequality in the following lemma is well-known. Indeed, it is known as Young's convolution inequality for integrals with respect to the bi-invariant Haar measure associated to a locally compact group. In our case, the group is $(\mathbb{R}^{2d+1}, \circ)$. We include it here for completeness and, importantly, because the change in domain of integration is crucial to our results above.

Lemma 4.A.1. *Fix any measurable sets $A, B \subset \mathbb{R}^{2d+1}$. Let g and ψ be measurable functions, and let the indices $r, p, q \in [1, \infty]$ satisfy*

$$\frac{1}{r} + 1 = \frac{1}{p} + \frac{1}{q}.$$

If $\operatorname{supp}\psi \subset B$, *then*

$$\|g * \psi\|_{L^r(A)} \leq \|g\|_{L^p(A\circ B^{-1})}\|\psi\|_{L^q(B)}.$$

We recall the definition (4.22) *of $A \circ B^{-1}$ and the definition* (4.26) *of the kinetic convolution.*

Proof. Fix any $z \in A$. Let us note that

$$\frac{1}{r} + \frac{r-p}{rp} + \frac{r-q}{rq} = \frac{1}{p} + \frac{1}{q} - \frac{1}{r} = 1.$$

Hence, applying the generalized Hölder inequality to the suitably re-written convolution, we have

$$\begin{aligned}
|(g*\psi)(z)| &\le \int_B |g(z\circ\widetilde{z}^{-1})||\psi(\widetilde{z})|\,d\widetilde{z}\\
&= \int_B |g(z\circ\widetilde{z}^{-1})|^{1+\frac{p}{r}-\frac{p}{r}}|\psi(\widetilde{z})|^{1+\frac{q}{r}-\frac{q}{r}}\,d\widetilde{z}\\
&= \int_B |g(z\circ\widetilde{z}^{-1})|^{\frac{p}{r}}|\psi(\widetilde{z})|^{\frac{q}{r}}|g(z\circ\widetilde{z}^{-1})|^{1-\frac{p}{r}}|\psi(\widetilde{z})|^{1-\frac{q}{r}}\,d\widetilde{z}\\
&= \int_B \left(|g(z\circ\widetilde{z}^{-1})|^{p}|\psi(\widetilde{z})|^{q}\right)^{\frac{1}{r}}|g(z\circ\widetilde{z}^{-1})|^{\frac{r-p}{r}}|\psi(\widetilde{z})|^{\frac{r-q}{r}}\,d\widetilde{z}\\
&\le \left\|\left(|g(z\circ\cdot^{-1})|^{p}|\psi|^{q}\right)^{\frac{1}{r}}\right\|_{L^r(B)}\left\||g(z\circ\cdot^{-1})|^{\frac{r-p}{r}}\right\|_{L^{\frac{pr}{r-p}}(B)}\left\||\psi|^{\frac{r-q}{r}}\right\|_{L^{\frac{qr}{r-q}}(B)}.
\end{aligned}$$

We simplify two of the norms above. First, we clearly have

$$\left\||\psi|^{\frac{r-q}{r}}\right\|_{L^{\frac{qr}{r-q}}(B)} = \|\psi\|_{L^q(B)}^{\frac{r-q}{r}}.$$

Additionally, by using that $z\circ B^{-1}\subset A\circ B^{-1}$, we see that

$$\left\||g(z\circ\cdot^{-1})|^{\frac{r-p}{r}}\right\|_{L^{\frac{pr}{r-p}}(B)} = \left\|g(z\circ\cdot^{-1})\right\|_{L^p(B)}^{\frac{r-p}{r}} = \|g(z\circ\cdot)\|_{L^p(B^{-1})}^{\frac{r-p}{r}} \le \|g\|_{L^p(A\circ B^{-1})}^{\frac{r-p}{r}}. \tag{4.A.1}$$

Here, we used that the Jacobian associated to $\widetilde{z}\mapsto\widetilde{z}^{-1}$ is one.

In summary, we have arrived at, for any fixed $z\in A$,

$$|(g*\psi)(z)| \le \left(\int_B |g(z\circ\widetilde{z}^{-1})|^p|\psi(\widetilde{z})|^q\,d\widetilde{z}\right)^{1/r}\|g\|_{L^p(A\circ B^{-1})}^{\frac{r-p}{r}}\|\psi\|_{L^q(B)}^{\frac{r-q}{r}}.$$

We now integrate over all $z\in A$, use the Fubini-Tonelli theorem, and enlarge the domains as we did in (4.A.1) to find:

$$\begin{aligned}
\|g*\psi\|_{L^r(A)}^r &\le \|g\|_{L^p(A\circ B^{-1})}^{r-p}\|\psi\|_{L^q(B)}^{r-q}\int_A\left(\int_B |g(z\circ\widetilde{z}^{-1})|^p|\psi(\widetilde{z})|^q\,d\widetilde{z}\right)dz\\
&\le \|g\|_{L^p(A\circ B^{-1})}^{r-p}\|\psi\|_{L^q(B)}^{r-q}\int_B |\psi(\widetilde{z})|^q\left(\int_A |g(z\circ\widetilde{z}^{-1})|^p\,dz\right)d\widetilde{z}\\
&\le \|g\|_{L^p(A\circ B^{-1})}^{r-p}\|\psi\|_{L^q(B)}^{r-q}\int_B |\psi(\widetilde{z})|^q\left(\int_{A\circ B^{-1}} |g(\zeta)|^p\,d\zeta\right)d\widetilde{z}\\
&= \|g\|_{L^p(A\circ B^{-1})}^{r}\|\psi\|_{L^q(B)}^{r}.
\end{aligned}$$

The proof is complete after taking each side to the $1/r$ power.

□

Bibliography

[1] Abedin, F., and Tralli, G. Harnack inequality for a class of Kolmogorov-Fokker-Planck equations in non-divergence form. *Arch. Ration. Mech. Anal. 233*, 2 (2019), 867–900.

[2] Albritton, D., Armstrong, S., Mourrat, J.-C., and Novack, M. Variational methods for the kinetic Fokker–Planck equation. *Anal. PDE 17*, 6 (2024), 1953–2010.

[3] Anceschi, F., Dietert, H., Guerand, J., Loher, A., Mouhot, C., and Rebucci, A. Poincaré inequality and quantitative De Giorgi method for hypoelliptic operators. *Preprint, arXiv:2401.12194* (2024).

[4] Anceschi, F., and Polidoro, S. A survey on the classical theory for Kolmogorov equation. *Matematiche (Catania) 75*, 1 (2020), 221–258.

[5] Anceschi, F., and Rebucci, A. A note on the weak regularity theory for degenerate Kolmogorov equations. *J. Differential Equations 341* (2022), 538–588.

[6] Anceschi, F., and Rebucci, A. On the fundamental solution for degenerate Kolmogorov equations with rough coefficients. *Journal of Elliptic and Parabolic Equations 9*, 1 (2023), 63–92.

[7] Barucci, E., Polidoro, S., and Vespri, V. Some results on partial differential equations and Asian options. *Math. Models Methods Appl. Sci. 11*, 3 (2001), 475–497.

[8] Biagi, S., and Bramanti, M. Schauder estimates for Kolmogorov-Fokker-Planck operators with coefficients measurable in time and Hölder continuous in space. *J. Math. Anal. Appl. 533*, 1 (2024), Paper No. 127996, 65.

[9] Bouin, E., Dolbeault, J., and Lafleche, L. Fractional hypocoercivity. *Comm. Math. Phys. 390*, 3 (2022), 1369–1411.

[10] Bouin, E., Dolbeault, J., Mischler, S., Mouhot, C., and Schmeiser, C. Hypocoercivity without confinement. *Pure Appl. Anal. 2*, 2 (2020), 203–232.

[11] Bramanti, M., and Brandolini, L. Schauder estimates for parabolic nondivergence operators of Hörmander type. *Journal of Differential Equations 234*, 1 (2007), 177 – 245.

[12] Bramanti, M., and Polidoro, S. Fundamental solutions for Kolmogorov-Fokker-Planck operators with time-depending measurable coefficients. *Math. Eng. 2*, 4 (2020), 734–771.

[13] Carrapatoso, K., and Mischler, S. The kinetic Fokker-Planck equation in a domain: Ultracontractivity, hypocoercivity and long-time asymptotic behavior. *Preprint, arXiv:2406.10112* (2024).

[14] Chandresekhar, S. Stochastic problems in physics and astronomy. *Rev. Modern Phys. 15* (1943), 1–89.

[15] Chapman, S., and Cowling, T. G. *The mathematical theory of nonuniform gases*, third ed. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1990. An account of the kinetic theory of viscosity, thermal conduction and diffusion in gases, In co-operation with D. Burnett, With a foreword by Carlo Cercignani.

[16] Chaudru de Raynal, P. E. Strong well posedness of McKean-Vlasov stochastic differential equations with Hölder drift. *Stochastic Process. Appl. 130*, 1 (2020), 79–107.

[17] Chaudru de Raynal, P.-E., Honoré, I., and Menozzi, S. Sharp Schauder estimates for some degenerate Kolmogorov equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 22*, 3 (2021), 989–1089.

[18] Chaudru de Raynal, P.-E., Honoré, I., and Menozzi, S. Strong regularization by Brownian noise propagating through a weak Hörmander structure. *Probab. Theory Related Fields 184*, 1-2 (2022), 1–83.

[19] Cushman, J. H., Park, M., Kleinfelter, N., and Moroni, M. Super-diffusion via Lévy lagrangian velocity processes. *Geophys. Res. Lett. 32*, 19 (2005).

[20] Da Prato, G., and Lunardi, A. On the Ornstein-Uhlenbeck operator in spaces of continuous functions. *J. Funct. Anal. 131*, 1 (1995), 94–114.

[21] Delarue, F., and Menozzi, S. Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal. 259*, 6 (2010), 1577–1630.

[22] Di Francesco, M., and Pascucci, A. On the complete model with stochastic volatility by Hobson and Rogers. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 460*, 2051 (2004), 3327–3338.

[23] Di Francesco, M., and Pascucci, A. On a class of degenerate parabolic equations of Kolmogorov type. *AMRX Appl. Math. Res. Express*, 3 (2005), 77–116.

[24] Di Francesco, M., and Polidoro, S. Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form. *Adv. Differential Equations 11*, 11 (2006), 1261–1320.

[25] Dong, H., and Yastrzhembskiy, T. Global L_p estimates for kinetic Kolmogorov-Fokker-Planck equations in nondivergence form. *Arch. Ration. Mech. Anal. 245*, 1 (2022), 501–564.

[26] Dong, H., and Yastrzhembskiy, T. Global Schauder estimates for kinetic Kolmogorov-Fokker-Planck equations. *Preprint, arXiv:2209.00769* (2022).

[27] Dong, H., and Yastrzhembskiy, T. Global L_p estimates for kinetic Kolmogorov-Fokker-Planck equations in divergence form. *SIAM J. Math. Anal. 56*, 1 (2024), 1223–1263.

[28] Duderstadt, J. J., and Martin, W. R. *Transport theory.* A Wiley-Interscience Publication. John Wiley & Sons, New York-Chichester-Brisbane, 1979.

[29] Eckmann, J.-P., Pillet, C.-A., and Rey-Bellet, L. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. *Comm. Math. Phys. 201*, 3 (1999), 657–697.

[30] Fabes, E. B., and Stroock, D. W. A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. *Arch. Rational Mech. Anal. 96*, 4 (1986), 327–338.

[31] Flandoli, F. *Random perturbation of PDEs and fluid dynamic models*, vol. 2015 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School].

[32] Folland, G. B., and Stein, E. M. *Hardy spaces on homogeneous groups*, vol. 28 of *Mathematical Notes*. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1982.

[33] Friedman, A. *Partial differential equations of parabolic type*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.

[34] Friedman, A. *Stochastic differential equations and applications*. Dover Publications, Inc., Mineola, NY, 2006. Two volumes bound as one, Reprint of the 1975 and 1976 original published in two volumes.

[35] Garofalo, N., and Tralli, G. Hardy-Littlewood-Sobolev inequalities for a class of non-symmetric and non-doubling hypoelliptic semigroups. *Math. Ann. 383*, 1-2 (2022), 1–38.

[36] Garofalo, N., and Tralli, G. A new proof of the geometric Sobolev embedding for generalised Kolmogorov operators. In *INdAM Meeting: Kolmogorov Operators and their Applications Workshop* (2022), Springer, pp. 117–134.

[37] Golse, F., Imbert, C., Mouhot, C., and Vasseur, A. Harnack inequality for kinetic Fokker-Planck equations with rough coefficients and application to the Landau equation. *Annali della Scuola Normale Superiore di Pisa XIX*, 1 (2019), 253–295.

[38] Groeneboom, P., Jongbloed, G., and Wellner, J. A. Integrated Brownian motion, conditioned to be positive. *Ann. Probab. 27*, 3 (1999), 1283–1303.

[39] Guerand, J., and Mouhot, C. Quantitative De Giorgi methods in kinetic theory. *J. Éc. polytech. Math. 9* (2022), 1159–1181.

[40] Hao, Z., Jabir, J.-F., Menozzi, S., Röckner, M., and Zhang, X. Propagation of chaos for moderately interacting particle systems related to singular kinetic mckean-vlasov sdes, 2024.

[41] Hao, Z., Wu, M., and Zhang, X. Schauder estimates for nonlocal kinetic equations and applications. *J. Math. Pures Appl. (9) 140* (2020), 139–184.

[42] Henderson, C., Lucertini, G., and Wang, W. A kinetic nash inequality and precise boundary behavior of the kinetic fokker-planck equation. *Preprint, arXiv:2407.08785* (2024).

[43] Henderson, C., and Snelson, S. C^∞ smoothing for weak solutions of the inhomogeneous Landau equation. *Arch. Ration. Mech. Anal. 236*, 1 (2020), 113–143.

[44] Henderson, C., and Wang, W. Kinetic Schauder estimates with time-irregular coefficients and uniqueness for the Landau equation. *Discrete Contin. Dyn. Syst. 44*, 4 (2024), 1026–1072.

[45] Hobson, D. G., and Rogers, L. C. G. Complete models with stochastic volatility. *Math. Finance 8*, 1 (1998), 27–48.

[46] Hörmander, L. Hypoelliptic second order differential equations. *Acta Math. 119* (1967), 147–171.

[47] Hou, M. Boundedness of weak solutions to degenerate Kolmogorov equations of hypoelliptic type in bounded domains. *Preprint, arXiv:2407.00800* (2024).

[48] Hwang, H. J., Jang, J., and Jung, J. The Fokker-Planck equation with absorbing boundary conditions in bounded domains. *SIAM J. Math. Anal. 50*, 2 (2018), 2194–2232.

[49] Hwang, H. J., Jang, J., and Velázquez, J. J. L. The Fokker-Planck equation with absorbing boundary conditions. *Arch. Ration. Mech. Anal. 214*, 1 (2014), 183–233.

[50] Hwang, H. J., and Kim, J. The Vlasov-Poisson-Fokker-Planck equation in an interval with kinetic absorbing boundary conditions. *Stochastic Process. Appl. 129*, 1 (2019), 240–282.

[51] Il'in, A. M. On a class of ultraparabolic equations. *Dokl. Akad. Nauk SSSR 159* (1964), 1214–1217.

[52] Imbert, C., and Mouhot, C. The Schauder estimate in kinetic theory with application to a toy nonlinear model. *Ann. H. Lebesgue 4* (2021), 369–405.

[53] Imbert, C., and Silvestre, L. The Schauder estimate for kinetic integral equations. *Anal. PDE 14*, 1 (2021), 171–204.

[54] Imbert, C., and Silvestre, L. Global regularity estimates for the Boltzmann equation without cut-off. *J. Amer. Math. Soc. 35*, 3 (2022), 625–703.

[55] Issoglio, E., Pagliarani, S., Russo, F., and Trevisani, D. Degenerate mckean-vlasov equations with drift in anisotropic negative besov spaces, 2024.

[56] Kolmogoroff, A. Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math. (2) 35*, 1 (1934), 116–117.

[57] Lanconelli, A., Pagliarani, S., and Pascucci, A. Local densities for a class of degenerate diffusions. *Ann. Inst. Henri Poincaré Probab. Stat. 56*, 2 (2020), 1440–1464.

[58] Lanconelli, A., and Pascucci, A. Nash estimates and upper bounds for non-homogeneous Kolmogorov equations. *Potential Anal. 47*, 4 (2017), 461–483.

[59] Lanconelli, E., Pascucci, A., and Polidoro, S. Linear and nonlinear ultraparabolic equations of Kolmogorov type arising in diffusion theory and in finance. In *Nonlinear problems in mathematical physics and related topics, II*, vol. 2 of *Int. Math. Ser. (N. Y.)*. Kluwer/Plenum, New York, 2002, pp. 243–265.

[60] Lanconelli, E., and Polidoro, S. On a class of hypoelliptic evolution operators. *Rend. Sem. Mat. Univ. Politec. Torino 52*, 1 (1994), 29–63. Partial differential equations, II (Turin, 1993).

[61] Lieberman, G. M. Intermediate Schauder theory for second order parabolic equations. IV. Time irregularity and regularity. *Differential Integral Equations 5*, 6 (1992), 1219–1236.

[62] Lipton, A. Kelvin Waves, Klein-Kramers and Kolmogorov Equations, Path-Dependent Financial Instruments: Survey and New results. *Preprint, arXiv:2309.04547* (2023).

[63] Loher, A. Quantitative Schauder estimates for hypoelliptic equations. *Preprint, arXiv:2305.00463* (2023).

[64] Lorenzi, L. Schauder estimates for degenerate elliptic and parabolic problems with unbounded coefficients in $\mathbb{R}^N$. *Differential Integral Equations 18*, 5 (2005), 531–566.

[65] Lucertini, G., Pagliarani, S., and Pascucci, A. Optimal schauder estimates for kinetic kolmogorov equations with time measurable coefficients. *Preprint, arXiv:2304.13392* (2023).

[66] Lunardi, A. Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in $\mathbf{R}^n$. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 24*, 1 (1997), 133–164.

[67] Manfredini, M. The Dirichlet problem for a class of ultraparabolic equations. *Adv. Differential Equations 2*, 5 (1997), 831–866.

[68] McKean, Jr., H. P., and Singer, I. M. Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry 1*, 1 (1967), 43–69.

[69] Menozzi, S. Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electron. Commun. Probab. 16* (2011), 234–250.

[70] Menozzi, S. Martingale problems for some degenerate Kolmogorov equations. *Stochastic Process. Appl. 128*, 3 (2018), 756–802.

[71] Pagliarani, S., Lucertini, G., and Pascucci, A. Optimal regularity for degenerate Kolmogorov equations in non-divergence form with rough-in-time coefficients. *J. Evol. Equ. 23*, 4 (2023), Paper No. 69, 37.

[72] Pagliarani, S., Pascucci, A., and Pignotti, M. Intrinsic Taylor formula for Kolmogorov-type homogeneous groups. *J. Math. Anal. Appl. 435*, 2 (2016), 1054–1087.

[73] Pagliarani, S., Pascucci, A., and Pignotti, M. Intrinsic expansions for averaged diffusion processes. *Stochastic Process. Appl. 127*, 8 (2017), 2560–2585.

[74] Pagliarani, S., and Polidoro, S. A Yosida's parametrix approach to Varadhan's estimates for a degenerate diffusion under the weak Hörmander condition. *J. Math. Anal. Appl. 517*, 1 (2023), Paper No. 126538, 42.

[75] Pascucci, A. *PDE and martingale methods in option pricing*, vol. 2 of *Bocconi & Springer Series*. Springer, Milan; Bocconi University Press, Milan, 2011.

[76] Pascucci, A., and Pesce, A. The parametrix method for parabolic SPDEs. *Stochastic Process. Appl. 130*, 10 (2020), 6226–6245.

[77] Pascucci, A., and Pesce, A. On stochastic Langevin and Fokker-Planck equations: the two-dimensional case. *J. Differential Equations 310* (2022), 443–483.

[78] Pascucci, A., and Pesce, A. Sobolev embeddings for kinetic Fokker-Planck equations. *J. Funct. Anal. 286*, 7 (2024), Paper No. 110344, 40.

[79] Polidoro, S. On a class of ultraparabolic operators of Kolmogorov-Fokker-Planck type. *Matematiche (Catania) 49*, 1 (1994), 53–105 (1995).

[80] Polidoro, S., Rebucci, A., and Stroffolini, B. Schauder type estimates for degenerate Kolmogorov equations with Dini continuous coefficients. *Commun. Pure Appl. Anal. 21*, 4 (2022), 1385–1416.

[81] Priola, E. Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math. 194*, 2 (2009), 117–153.

[82] Silvestre, L. Hölder estimates for kinetic Fokker-Planck equations up to the boundary. *Ars Inven. Anal.* (2022), Paper No. 6, 29.

[83] Soize, C. *The Fokker-Planck equation for stochastic dynamical systems and its explicit steady state solutions*, vol. 17 of *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co., Inc., River Edge, NJ, 1994.

[84] Sonin, I. M. A class of degenerate diffusion processes. *Teor. Verojatnost. i Primenen 12* (1967), 540–547.

[85] VILLANI, C. A review of mathematical topics in collisional kinetic theory. In *Handbook of mathematical fluid dynamics, Vol. I.* North-Holland, Amsterdam, 2002, pp. 71–305.

[86] VILLANI, C. Hypocoercivity. *Mem. Amer. Math. Soc. 202*, 950 (2009), iv+141.

[87] WEBER, M. The fundamental solution of a degenerate partial differential equation of parabolic type. *Trans. Amer. Math. Soc. 71* (1951), 24–37.

[88] ZHANG, X., AND ZHANG, X. Cauchy problem of stochastic kinetic equations. *Ann. Appl. Probab. 34*, 1A (2024), 148–202.

[89] ZHU, Y. Regularity of kinetic Fokker-Planck equations in bounded domains. *Preprint, arXiv:2206.04536* (2022).