Weak Solutions to a Fractional Fokker-Planck Equation via Splitting and Wasserstein Gradient Flow

by

Malcolm Bowles
B.Sc., University of Victoria, 2012

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics & Statistics

Weak Solutions to a Fractional Fokker-Planck Equation via Splitting and Wasserstein Gradient Flow

by

Malcolm Bowles
B.Sc., University of Victoria, 2012

Supervisory Committee

Dr. Martial Agueh, Supervisor
(Department of Mathematics & Statistics)

Dr. Reinhard Illner, Departmental Member
(Department of Mathematics & Statistics)

Supervisory Committee

Dr. Martial Agueh, Supervisor
(Department of Mathematics & Statistics)

Dr. Reinhard Illner, Departmental Member
(Department of Mathematics & Statistics)

ABSTRACT

In this thesis, we study a linear fractional Fokker-Planck equation that models non-local ('fractional') diffusion in the presence of a potential field. The non-locality is due to the appearance of the 'fractional Laplacian' in the corresponding PDE, in place of the classical Laplacian which distinguishes the case of regular (Gaussian) diffusion. We introduce the fractional Laplacian via the Fourier transform, and show equivalence of the Fourier definition with a singular integral formulation which explicitly characterizes the non-local effects.

Motivated by the observation that, in contrast to the classical Fokker-Planck equation (describing regular diffusion in the presence of a potential field), there is no natural gradient flow formulation for its fractional counterpart, we prove existence of weak solutions to this fractional Fokker-Planck equation by combining a splitting technique together with a Wasserstein gradient flow formulation. An explicit iterative construction is given, which we prove weakly converges to a weak solution of this PDE.

Contents

Supervisory Committee ii

Abstract iii

Table of Contents iv

Acknowledgements vi

1 Introduction 1

1.1 Notation 9

1.2 Assumptions on Initial Data and Potential 10

1.3 Statement of Main Result 11

2 The Fractional Laplacian 12

2.1 The Fractional Laplacian through the Fourier Transform 12

2.2 The Fractional Laplacian as a Singular Integral 14

2.2.1 Equality of Fourier and Singular Integral Representation on Non-Schwartz Functions 23

2.2.2 Integration by Parts 24

3 The Fractional Heat Equation 25

3.1 Properties of Solutions to the Fractional Heat Equation 25

4 The Transport Equation as a Gradient Flow 30

4.1 Gradient Flow in Metric Spaces 30

4.2 Optimal Transportation & the 2-Wasserstein Distance 33

4.3 Transport as Steepest Descent of the Potential Energy 35

4.4 The Characteristic Equation 42

5 Operator Splitting on the Fractional Fokker-Planck Equation 44

5.1 Construction 44
5.2 Time-Dependent Approximation 47
5.3 Convergence to a Weak Solution 56
5.3.1 Proof of the Main Result 61

6 Conclusion **63**
6.1 Concluding Remarks 63
6.2 Open Questions 64
6.2.1 Regularity and Uniqueness 64
6.2.2 Extension of the Method 66

Bibliography **67**

ACKNOWLEDGEMENTS

I would like to thank:

my supervisor, Dr. Martial Agueh, for his endless patience, encouragement, and advice. One could not ask for a greater supervisor.

NSERC, and the University of Victoria, whose funding support is gratefully acknowledged, and

my family, and colleagues in the Math Department, for their encouragement, support, and helpful discussions.

Chapter 1

Introduction

The diffusion, or heat equation, $\partial_t \rho = \Delta \rho$, is a classical and intensively studied PDE which has been very successful in describing a wide range of physical phenomena [16]. In the study of continuous-time stochastic processes, it is closely connected to the theory of Brownian motion (or Wiener processes); in particular, if $X = \{X_t : 0 \leq t < \infty\}$ is a Brownian motion that admits, at each time t, a probability density $\rho(t)$, then in fact ρ solves the heat equation [20]. On a more intuitive level, it is well known that a Brownian motion can be constructed from a suitable limit of a discrete random walk with finite variance, and it is not hard to check that the probability distribution of this random walk satisfies a discrete version of the heat equation [20]. It is this random walk we imagine when we think of the physical process of diffusion.

An alternative viewpoint of diffusion is that of an irreversible process from thermodynamics. Irreversible processes are, in particular, characterized by the fact that their entropy (given by $S = -\int \rho \log \rho$ in the continuous case, where ρ is a probability distribution over the continuous state space) always increases. In particular, as thermodynamic equilibrium of a system is achieved for a state of maximum entropy by the Second Law of Thermodynamics, we imagine entropy as 'driving' the evolution, i.e. diffusion is a result of a system 'seeking' to maximize its entropy at any given instant in time.

In their seminal paper [19], Jordan, Kinderlehrer, and Otto, were (as a special case) able to make a connection between the time evolution of a solution to $\partial_t \rho = \Delta \rho$

and its corresponding entropy $-\int \rho \log \rho$. They proved that

$$\begin{cases} \rho_t = \Delta\rho + \operatorname{div}(\rho\nabla\Psi(x)) & \text{in } \mathbb{R}^d \times (0,\infty) \\ \rho = \rho^0 \quad \text{on } \mathbb{R}^d \times \{t=0\} \end{cases} \tag{1.1}$$

which models a diffusing particle moving in a potential field Ψ, is a gradient flow, or steepest descent, of the free energy functional $F(\rho) := \int_{\mathbb{R}^d} \rho \log \rho + \int_{\mathbb{R}^d} \rho\Psi$ with respect to the metric W_2, called the 2-Wasserstein metric, on the space of probability measures (see Chapter 4)[19] . That is, at each instant in time, solutions of (1.1) follow the direction of steepest descent of $F(\rho)$ w.r.t. the 2-Wasserstein distance. In particular, $\Psi \equiv 0$ gives a precise meaning to the idea that dynamics of the heat equation occur because the system seeks to maximize its entropy at every instant in time [19].

Let us return to the random walk interpretation of the heat equation. For review purposes, we sketch out the connection. Consider a particle, starting at the origin, that at each time step τ has an equal probability to jump to one of the lattice points $\pm he_i$ of $h\mathbb{Z}^d$, where $e_i = (0,\dots,0,1,0\dots,0)$ is a unit vector in the ith direction, and $h > 0$ is a given step size. The probability $p_0(x, t+\tau)$ that the particle is at $x \in h\mathbb{Z}^d$ at time $t+\tau \in \tau\mathbb{N}$, given that it started at the origin, satisfies the following relation

$$p_0(x, t+\tau) = \frac{1}{2d}\sum_{i=1}^{d} p_0(x+he_i, t) + p_0(x-he_i, t),$$

or equivalently,

$$\frac{p_0(x,t+\tau) - p_0(x,t)}{\tau} = \frac{h^2}{2d\tau}\sum_{i=1}^{d} \frac{p_0(x+he_i,t) - 2p_0(x,t) + p_0(x-he_i,t)}{h^2}.$$

We imagine h and τ to correspond to the mean distance and time between collisions. In the above display, the right-hand side has the form of a discretization of the Laplacian. Assuming $h^2 \propto 2d\tau$, i.e. h scales according to the square root of τ, as h, $\tau \to 0$, we obtain a continuous probability distribution ρ satisfying the heat equation

$$\begin{cases} \partial_t\rho(x,t) = \Delta\rho(x,t), & x \in \mathbb{R}^d, t > 0, \\ \rho(x,0) = \delta(x) \quad x \in \mathbb{R}^d, \end{cases}$$

where δ is the Dirac measure (i.e. with probability 1, the particle started at the origin). The solution for $t > 0$ is given by $\rho(x,t) = \Phi(x,t)$, where $\Phi(x,t) = \frac{1}{(4\pi t)^{d/2}} e^{-|x|^2/4t}$ is a Gaussian distribution for each fixed $t > 0$ [16]. (If instead, we have an initial distribution ρ^0 for the particle rather than a precise starting location, then the convolution $\rho(x,t) = \Phi(t) * \rho^0(x)$ furnishes the probability distribution of the particle at time $t > 0$.)

The second moment of a solution to the heat equation, $\int_{\mathbb{R}^d} x^2 \rho(x,t)\,\mathrm{d}x$, is characterized by the fact that it increases in proportion to t (we omit the computation here). Thus in an experiment measuring the *mean square displacement* of a particle (which is equivalent to the second moment, if we choose the particle to be initially at the origin), we expect a linear dependence with time if the process is well described by the classical heat equation, see e.g. the famous work by Perrin [23]. However, certain experiments involving diffusion (see e.g. [25], or [9] and references therein) have shown that the mean-square displacement is not proportional to t, but instead to $t^\alpha, \alpha \neq 1$. This suggests that Gaussian diffusion, and in particular, on a discrete level, the classical random walk, is no longer a good model for the observed physical process. Instead, we introduce another random walk from [27], and formally investigate its limit. We remark that such a random walk cannot have finite variance (since this will lead to Brownian motion).

Therefore, suppose now that at any given point in the lattice, there is a non-zero probability to jump to any of the other lattice points in $h\mathbb{Z}^d$, that is, long-range effects are present. Specifically [27], let $K : \mathbb{R}^d \to [0,\infty)$ be a function satisfying $K(-x) = K(x)$ with normalization $\sum_{i\in\mathbb{Z}^d} K(i) = 1$, specifying the distribution of these jump sizes. Then with the same notation as above

$$p_0(x, t+\tau) - p_0(x,t) = \sum_{i\in\mathbb{Z}^d} K(i)\left[p_0(x+ih,t) - p_0(x,t)\right],$$

or

$$\frac{p_0(x, t+\tau) - p_0(x,t)}{\tau} = \sum_{i\in\mathbb{Z}^d} \frac{K(i)}{\tau}\left[p_0(x+ih,t) - p_0(x,t)\right].$$

The classical case is recovered when $K(i) = 1/2d$ for $i \in \mathbb{Z}^d$ satisfying $|i| = 1$, and $K(i) = 0$ otherwise. For convenience, we rewrite the above using the symmetry in K as

$$\frac{p_0(x, t+\tau) - p_0(x,t)}{\tau} = \frac{1}{2}\sum_{i\in\mathbb{Z}^d} \frac{K(i)}{\tau}\left[p_0(x+ih,t) + p(x-ih,t) - 2p_0(x,t)\right].$$

Without any motivation here, let us choose K to be a homogeneous 'heavy-tailed' distribution, depending on a parameter $s \in (0,1)$, $K_s(x) := \frac{C}{|x|^{d+2s}}$, $|x| > 0$, $K_s(0) = 0$, with an appropriate normalizing constant C.

Our first observation [27] is that for such a choice, the second moment,

$$\sum_{i \in \mathbb{Z}^d \setminus \{0\}} |i|^2 K(i) = C \sum_{i \in \mathbb{Z}^d \setminus \{0\}} |i|^{2-d-2s} = +\infty.$$

In particular, this random walk has an infinite variance for every $s \in (0,1)$.

Now we wish to formally investigate the limit $\tau, h \to 0$. To this end, suppose τ scales according to h^{2s}, $\tau \propto h^{2s}$. Then (up to constants), $\frac{K_s(i)}{\tau} = h^d K_s(ih)$, so

$$\frac{p_0(x,t+\tau) - p_0(x,t)}{\tau} = \frac{h^d}{2} \sum_{i \in \mathbb{Z}^d} K_s(ih) \left[p_0(x+ih,t) + p(x-ih,t) - 2p_0(x,t) \right].$$

Formally, the right-hand side of the above display is a Riemann sum, while the left-hand side is a discretization of a derivative in t. Therefore if $\tau, h \to 0$ with $\tau \propto h^{2s}$, we anticipate (up to constants) the equation

$$\begin{cases} \partial_t \rho(x,t) = \displaystyle\int_{\mathbb{R}^d} \frac{\rho(x+y,t) + \rho(x-y,t) - 2\rho(x,t)}{|y|^{d+2s}} \, \mathrm{d}y, & x \in \mathbb{R}^d, t > 0 \\ \rho(x,0) = \delta(x) \quad x \in \mathbb{R}^d. \end{cases}$$

This singular integral on the right-hand side is, up to a constant (which depends on s), a non-local linear operator called the *fractional* Laplacian, and denoted by $(-\Delta)^s$ (see Chapter 2 for more details). The corresponding PDE is known as the fractional heat equation,

$$\partial_t \rho = -(-\Delta)^s \rho. \tag{1.2}$$

Although the variance of a solution ρ to (1.2) is infinite (see Chapter 2) which is non-physical, one can still define a 'pseudo-variance', $\left[\int_{\mathbb{R}^d} x^\beta \rho(x,t) \, \mathrm{d}x\right]^{2/\beta}$ where $\beta < 2s$. It can be shown that this pseudo-variance satisfies $\left[\int_{\mathbb{R}^d} x^\beta \rho(x,t) \, \mathrm{d}x\right]^{2/\beta} \propto t^{1/s}$. Thus, the fractional heat equation can be considered as a model for situations where there is non-Gaussian diffusion.

The continuous-time stochastic process corresponding to the limit of this random walk is not a Brownian motion as in the classical random walk case, but instead belongs to a more general class of stochastic processes called Lévy processes, to which

Brownian motion belongs [4].

Formally speaking, a Lévy process X is a stochastic process which has stationary and independent increments [4]; in particular, a Brownian motion is a Lévy process for which the independent increments have a Gaussian distribution. If X is a *symmetric pure jump* $2s$*-stable Lévy process* that admits a density $\rho(t)$ at each time t, then $\partial_t \rho = -(-\Delta)^s \rho$. This terminology comes from the celebrated Lévy-Itô decomposition [4] which says, roughly speaking, that every Lévy process is the sum of a deterministic drift, a Brownian motion, and a jump process (related to a compound Poisson process - a Poisson process with random jump sizes). A pure jump process is a Lévy process which contains no drift or Brownian motion. More precisely, a Lévy process can be classified by its *characteristic function*, which determines the probability distribution of the process. The Lévy-Khintchine formula [4] gives a canonical representation for the characteristic function, which is given by a *Lévy triple* (b, A, ν), where $b \in \mathbb{R}^d$ is related to a deterministic drift, $A \in \mathbb{R}^{d \times d}$ is related to a Brownian motion, and ν is a (Lévy) measure on $\mathbb{R}^d \backslash \{0\}$ related to a jump process. A pure jump Lévy process has Lévy triple $(0, 0, \nu)$. In particular, the pure jump process that corresponds to the fractional Laplacian has (up to constants) $\mathrm{d}\nu(y) = |y|^{-d-2s}\,\mathrm{d}y$. Since $\nu(-A) = \nu(A)$ it is a *symmetric* pure jump process. Finally, the terminology *stable* means that there exist real-valued sequences $\{c_n\}$ and $\{d_n\}$ such that $X_1 + \ldots + X_n$ is equal in distribution to $c_n X + d_n$ for each n, where X_i is an independent copy of the Lévy process X. It can be shown (see references in [4]) that c_n can take only the form $c_n = \sigma n^{1/2s}$, $0 < s \leq 1$, and thus $2s$ is said to be the index of stability.

The above discussion has been rather brief and formal, but it is not our aim to fully develop the theory of Lévy processes here; for the interested reader we refer to [4]. Rather, we wish simply to draw a connection between the fractional heat equation, the 'heavy-tailed' random walk, and the corresponding Lévy process, in the same way as that of the heat equation, standard random walk, and the corresponding Brownian motion.

We consider the fractional heat equation as characterizing a non-Gaussian diffusion, and refer to this as 'fractional diffusion'. In particular, the solution to (1.2) in $\mathbb{R}^d$ with initial distribution ρ^0 is given by $\rho(t) = \Phi_s(t) * \rho^0$, where now Φ_s is a non-Gaussian kernel (see Chapter 3).

One may wonder if there is a similar gradient flow interpretation of the fractional heat equation involving the entropy $-\int_{\mathbb{R}^d} \rho \log \rho$ as there was for the heat equation. Indeed, Mathias Erbar [15] rigorously proved that the fractional heat equation is the gradient flow of the entropy, not with respect to the 2-Wasserstein distance, but with respect to a new 'modified Wasserstein' distance built from the Lévy measure and based on the Benamou-Brenier variant of the 2-Wasserstein distance [28]; see [15] for details. However, there appears to be no such extension to the 'fractional' Fokker-Planck equation corresponding to (1.1),

$$\begin{cases} \rho_t = -(-\Delta)^s \rho + \operatorname{div}\left(\rho \nabla \Psi(x)\right) \quad \text{in } \mathbb{R}^d \times (0,\infty), \quad s \in (0,1) \\ \rho = \rho^0 \quad \text{on } \mathbb{R}^d \times \{t = 0\}\,. \end{cases} \tag{1.3}$$

It is unknown if it is even possible to regard (1.3) as a gradient flow of an energy functional in some metric space. Indeed, there does not seem to be any obvious extension of the work by Erbar to (1.3), since the distance there was seemingly designed with precisely the entropy $-\int \rho \log \rho$ in mind.

Instead, we think of (1.3) as really consisting of the two separate processes of fractional diffusion, and transport in the field of the potential Ψ. Moreover, we think it is natural to consider transport dynamics as arising from the tendency of a particle to minimize its potential energy in this field, that is, as a gradient flow of the potential energy (with respect to the 2-Wasserstein distance; see Chapter 4).

It is therefore our interest to see if solutions to (1.3) can in fact be obtained by separating, or splitting, (1.3) into these two processes, and solving each separately, on a vanishingly small interval of time. That is, within some small time interval of duration τ, we imagine that dynamics of (1.3) correspond to evolving a given initial distribution according to the fractional heat equation $\partial_t \rho = -(-\Delta)^s \rho$, and then running a gradient flow of the potential energy in the 2-Wasserstein distance. When $\tau \to 0$, we hope to recover a solution of (1.3). More precisely, we recursively iterate the following two connected subproblems for $n = 0, 1, \ldots, N-1$, given some finite time horizon $T < \infty$ and time-step $\tau = T/N$:

1. *(The fractional heat equation)*

$$\begin{aligned} &\partial_t u(x,t) = -(-\Delta)^s u(x,t), \quad (x,t) \in \mathbb{R}^d \times (0,\infty) \\ &u(x,0) = \rho_\tau^n(x) \end{aligned}$$

   Set $\tilde{\rho}_\tau^{n+1}(x) := u(x,\tau)$.

2. *(Gradient Flow of the Potential Energy)*

Minimize

$$\rho \mapsto \frac{1}{2\tau} W_2(\tilde{\rho}_\tau^{n+1}, \rho)^2 + \int_{\mathbb{R}^d} \rho \Psi \, \mathrm{d}x \tag{1.4}$$

Set $\rho_\tau^{n+1}(x)$ *as the minimizer.*

We will explain (1.4) in Chapter 4.

The idea of splitting is well-known from numerical analysis. It has been applied to other 'fractional PDE's' [2, 13], such as the so-called fractional conservation law, $\partial_t u(x,t) + \operatorname{div}(f(u)) + (-\Delta)^s u(x,t) = 0$, as well as on other PDE's, to obtain existence of a solution; see e.g. [18] and references therein.

To see why splitting is a plausible approximation scheme, we run it on the simple ODE

$$\begin{cases} u'(t) = (A+B)u(t) \\ u(t=0) = u^0 \in \mathbb{R}^d \end{cases}$$

where $A, B \in \mathbb{R}^{d\times d}$ are $d \times d$ matrices with real-valued entries. The solution at time $t > 0$ is formally given by $u(t) = e^{t(A+B)}u^0$. If now, given some time-step $\tau > 0$, we solve the ODE's

$$\begin{cases} v'(t) = Av(t), \quad \text{with } v(t=0) = u^0 \\ w'(t) = Bw(t), \quad \text{with } w(t=0) = v(\tau) \end{cases}$$

then $w(\tau) = e^{\tau B} e^{\tau A} u^0$ is an approximation of $u(\tau)$. This is easily seen by the Taylor expansions

$$\begin{aligned} u(\tau) &= u^0 + \tau(A+B)u^0 + \frac{1}{2}\tau^2(A+B)^2 u^0 + o(\tau^2) \\ w(\tau) &= u^0 + \tau(A+B)u^0 + \frac{1}{2}\tau^2(A^2 + 2BA + B^2)u^0 + o(\tau^2), \end{aligned}$$

so that

$$|u(\tau) - w(\tau)| \le \tau^2 |(AB - BA)u^0| + o(\tau^2),$$

and so at some time $t = n\tau$,

$$|u(t) - w(t)| \le C\tau + o(\tau).$$

Returning now to (1.3), we remark that previous research [5, 17, 26] specifically on (1.3) has focused only on the long-time behaviour in the specific case $\Psi(x) = |x|^2/2$, where they prove exponential convergence of solutions to equilibrium. More precisely, they obtain ‘entropy’ inequalities [17] of type

$$\mathrm{Ent}^{\gamma}_{u_\infty}\left(\frac{u(t)}{u_\infty}\right) \le e^{-\frac{t}{C}}\mathrm{Ent}^{\gamma}_{u_\infty}\left(\frac{u_0}{u_\infty}\right),$$

where u is assumed to solve (1.3) with $\Psi = |x|^2/2$, u_∞ is the equilibrium solution, (the solution of $(-\Delta)^s u = \mathrm{div}\,(xu)$), and $\mathrm{Ent}^{\gamma}_{u_\infty}$ is defined for nonnegative functions f by

$$\mathrm{Ent}^{\gamma}_{u_\infty}(f) := \int_{\mathbb{R}^d} \gamma(f) u_\infty \,\mathrm{d}x - \gamma\left(\int_{\mathbb{R}^d} f u_\infty \,\mathrm{d}x\right),$$

where $\gamma : \mathbb{R}^+ \to \mathbb{R}$ is a smooth convex function. Since we are interested in proving existence of solutions via splitting, we do not find occasion to make use of these results in the sequel, and encourage interested readers to consult the above references for further details.

To the best of our knowledge, existence of solutions to (1.3) has not been proven via a splitting in this fashion before. We suspect, however, that existence by some other means may have already been established, but were unable to find any exact references in the literature. Indeed, it can be checked that the Duhamel-type formula

$$\rho(x,t) = \Phi_s(t) * \rho^0(x) + \int_0^t \Phi_s(t-t') * \mathrm{div}\,(\rho(t')\nabla\Psi)\,(x)\,\mathrm{d}t'$$

formally solves (1.3) (where Φ_s is the fractional heat kernel). Placing the spatial derivative on Φ_s instead of $\rho(t')\nabla\Psi$ in the above (‘integration by parts’) gives the notion of a *mild solution*, i.e. a ρ satisfying

$$\rho(x,t) = \Phi_s(t) * \rho^0(x) - \int_0^t \nabla\Phi_s(t-t') * [\rho(t')\nabla\Psi]\,(x)\,\mathrm{d}t'.$$

Provided the right-hand side of the above display makes sense, it may be possible to prove the existence of a mild solution by running a fixed point argument in, e.g., the Banach space $C((0,T);L^1(\mathbb{R}^d))$ [13]. If we were to continue in this direction, it seems that one should impose $\nabla\Psi \in L^\infty(\mathbb{R}^d)$, since we anticipate $\rho(t') \in L^1(\mathbb{R}^d)$, in order for the right-hand side to be well-defined. Such an assumption is not needed however in the following. Moreover, we apply splitting to (1.3) with the aim to see if a similar technique can be applied to other PDE’s which cannot be fully realized as a Wasserstein gradient flow.

This thesis is organized as follows. The rest of this chapter is concerned with setting some notation, giving the assumptions which will be used in the sequel, and a statement of the main result. In Chapter 2, we establish rigorous definitions and examine properties of the fractional Laplacian. This is followed by the brief exposition of Chapter 3 which will establish properties of solutions to the fractional heat equation. Chapter 4 discusses the gradient flow formulation of the transport equation. Finally, Chapter 5 is where the construction and convergence of the splitting is established.

1.1 Notation

In this section we set the notation we shall use. Other notation which is used locally is defined in each relevant section.

1. C is a constant that might vary from line to line.

2. We denote x for the spatial coordinate(s), and t for the 'time' coordinate.

3. We will usually suppress spatial dependence for functions, in particular when integrating. This means that if $f = f(x) : \mathbb{R}^d \to \mathbb{R}$ and $\varphi = \varphi(x,t) : \mathbb{R}^d \times (0,\infty) \to \mathbb{R}$, then
$$\int_{\mathbb{R}^d} f\varphi(t)\,\mathrm{d}x := \int_{\mathbb{R}^d} f(x)\varphi(x,t)\,\mathrm{d}x$$
We will always indicate dependence on t.

4. L^p spaces will be denoted as usual by
$$L^p(\mathbb{R}^d) := \left\{ f : \mathbb{R}^d \to \mathbb{R} \,:\, \|f\|^p_{L^p(\mathbb{R}^d)} := \int_{\mathbb{R}^d} |f|^p\,\mathrm{d}x < \infty \right\}, \quad 1 \le p < \infty,$$
$$L^\infty(\mathbb{R}^d) := \left\{ f : \mathbb{R}^d \to \mathbb{R} \,:\, \|f\|_{L^\infty(\mathbb{R}^d)} := \mathrm{essup}_{x\in\mathbb{R}^d}|f| < \infty \right\}.$$

5. If $\alpha = (\alpha_1, \dots, \alpha_d)$ is a d-tuple of non-negative integers, and $|\alpha| = \sum_{i=1}^d \alpha_i$, then for $f : \mathbb{R}^d \to \mathbb{R}$,
$$D^\alpha f(x) := \partial^{\alpha_1}_{x_1} \dots \partial^{\alpha_d}_{x_d} f(x).$$

6. If $f : \mathbb{R}^d \to \mathbb{R}$, then
$$\left\|D^2 f\right\|_{L^\infty(\mathbb{R}^d)} := \|g\|_{L^\infty(\mathbb{R}^d)}, \quad \text{where } g := \quad |D^2 f| = \left(\sum_{|\alpha|=2} |D^\alpha f|^2 \right)^{1/2}.$$

7. C^k functions,

$$C^0(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{R}, f \text{ is continuous}\}$$
$$C^k(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{R}, f \text{ is } k \text{ times continuously differentiable}\}$$

8. Let $0 < \alpha \leq 1$. The Hölder spaces

$$C^{0,\alpha}(\mathbb{R}^d) := \left\{ f : \mathbb{R}^d \to \mathbb{R} \,:\, f \in C^0(\mathbb{R}^d), \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x-y|^\alpha} < \infty \right\}$$
$$C^{k,\alpha}(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{R} \,:\, f \in C^k(\mathbb{R}^d), D^\beta f \in C^{0,\alpha}(\mathbb{R}^d) \text{ for all } \beta \text{ with } |\beta| = k\}$$

9. $\mathcal{P}_a^2(\mathbb{R}^d)$ is the set of absolutely continuous (w.r.t. Lebesgue) probability measures on $\mathbb{R}^d$ that have finite second moments, which we will identify with their densities,

$$\mathcal{P}_a^2(\mathbb{R}^d) := \left\{ \rho : \mathbb{R}^d \to \mathbb{R} \,:\, \rho \geq 0 \text{ a.e. }, \int_{\mathbb{R}^d} \rho \, \mathrm{d}x = 1, \int_{\mathbb{R}^d} |x|^2 \rho \, \mathrm{d}x < \infty \right\}.$$

   We will not make a distinction between a measure and its density, but the usage will be clear from the context.

10. B_R and $B_R(x)$ denote the open ball of radius R centred at the origin and at x, respectively; $1_{B_R}(x) := 1$ if $x \in B_R$, and 0 otherwise, denotes the indicator function.

1.2 Assumptions on Initial Data and Potential

In the sequel, we impose the following assumptions on ρ^0 and Ψ in (1.3).

(A1) $\rho^0 \in \mathcal{P}_a^2(\mathbb{R}^d) \cap L^p(\mathbb{R}^d)$ for some $1 < p \leq \infty$, $\int_{\mathbb{R}^d} \rho^0 \Psi \, \mathrm{d}x < \infty$.

(A2) $\Psi \in C^{1,1} \cap C^{2,1}(\mathbb{R}^d)$, $\Psi \geq 0$.

Remark 1.2.1. *We remark on the assumptions. We require $\Psi \in C^{1,1} \cap C^2(\mathbb{R}^d)$ so that $D^2\Psi$ is bounded. This allows us to have an estimate for the potential energy of a solution to the fractional heat equation, in terms of the potential energy of the initial data- see (5.5). Together with the assumption $\rho^0 \in L^p(\mathbb{R}^d)$ for $p > 1$, it allows us to prove a uniform L^p bound on the time-dependent approximate solution to (1.3)* *obtained from the splitting - see (5.43), crucial for obtaining (weak) compactness in* L^p.

We additionally impose $\Psi \in C^{2,1}(\mathbb{R}^d)$ *so that* $\nabla\Psi \cdot \xi \in C_c^{1,1}(\mathbb{R}^d)$ *for every* $\xi \in C_c^\infty(\mathbb{R}^d)$, *and consequently we have* $(-\Delta)^s [\nabla\Psi \cdot \xi] \in L^\infty(\mathbb{R}^d)$ *by Proposition 2.2.5. The nonnegativity of* Ψ *is a convenience so that* $\int_{\mathbb{R}^d} \rho\Psi \,\mathrm{d}x \geq 0$ *for all* $\rho \in \mathcal{P}_a^2(\mathbb{R}^d)$.

A typical example of a potential satisfying these properties is the quadratic function $\Psi(x) = |x|^2/2$.

1.3 Statement of Main Result

Our main result is as follows (see Theorem 5.3.5).

Theorem 1.3.1. *Let* $T < \infty$ *and* $\tau = T/N$ *for some* $N \in \mathbb{N}$, *and assume* ρ^0 *and* Ψ *satisfy the above given assumptions. Then there exists a sequence of functions* $\rho_\tau : \mathbb{R}^d \times (0,T) \to \mathbb{R}$ *(which is constructed from the splitting scheme outlined above) and a* $\rho \in L^1 \cap L^p(\mathbb{R}^d \times (0,T))$ *(where* $p > 1$*) such that*

1. ρ_τ *weakly converges to* ρ *in* $L^p(\mathbb{R}^d \times (0,T))$ *as* $\tau \to 0$,
2. $\int_{\mathbb{R}^d} \rho(x,t)\,\mathrm{d}x = \int_{\mathbb{R}^d} \rho^0(x)\,\mathrm{d}x$ *for a.e.* $t \in (0,T)$,
3. $\rho(x,t) \geq 0$ *for a.e.* $(x,t) \in \mathbb{R}^d \times (0,T)$, *and*
4. $\int_0^T \int_{\mathbb{R}^d} \rho(t)\left[\partial_t\varphi(t) - (-\Delta)^s\varphi(t) - \nabla\Psi \cdot \nabla\varphi(t)\right] \mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^d} \rho^0\varphi(0)\,\mathrm{d}x = 0$ *for all* $\varphi \in C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ *with time support in* $[-T,T]$.

Chapter 2

The Fractional Laplacian

In this chapter we establish some basic properties of the fractional Laplacian. Some questions which motivated the following exposition include, *For what functions does the fractional Laplacian exist (in the classical pointwise sense)? How does the fractional Laplacian act with regards to regularity and integrability? Can we integrate by parts for the fractional Laplacian?* We give answers to these questions, but do not attempt to recover results in full generality.

We first begin by detailing equivalent definitions of $(-\Delta)^s$ on $\mathbb{R}^d$, the first through the Fourier transform, and the second as a singular integral.

2.1 The Fractional Laplacian through the Fourier Transform

The simplest approach to defining the fractional Laplacian operator is through the Fourier transform on the space of smooth, rapidly decaying (Schwartz) functions on $\mathbb{R}^d$, which we denote by $\mathcal{S}(\mathbb{R}^d)$. Formally, we recall a function belongs $\mathcal{S}(\mathbb{R}^d)$ if the function, and all its derivatives, vanish as $|x| \to \infty$ faster than any function with polynomial growth.

We first recall the definition of the Fourier transform. Let $f \in L^1(\mathbb{R}^d)$. The Fourier transform of f, denoted by $\mathcal{F}[f]$, is defined by

$$\mathcal{F}[f](\xi) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i\langle x,\xi\rangle} f(x)\, \mathrm{d}x, \quad (\xi \in \mathbb{R}^d)$$

with the inverse Fourier transform

$$\mathcal{F}^{-1}(g)(x) := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle} g(\xi)\, \mathrm{d}\xi, \quad (x \in \mathbb{R}^d),$$

where $\langle x, y\rangle := \sum_{i=1}^{d} x_i y_i$ denotes the standard scalar product of $x, y \in \mathbb{R}^d$. We remark that occasionally we will use $\hat{f}$ instead of $\mathcal{F}[f]$ for clarity.

Proposition 2.1.1. *(Useful properties of the Fourier transform) The following properties hold for $f, g \in L^1(\mathbb{R}^d)$ (see [16]):*

1. $\mathcal{F}^{-1}(\mathcal{F}[f])(x) = f(x)$,
2. $\mathcal{F}[f * g] = (2\pi)^{d/2}\mathcal{F}[f]\,\mathcal{F}[g]$,
3. $\mathcal{F}[D^\alpha f] = (i\xi)^\alpha \mathcal{F}[f]$ *for each multiindex* α *and* $D^\alpha f \in L^1(\mathbb{R}^d)$,

Suppose $f \in \mathcal{S}(\mathbb{R}^d)$. By the properties above, the Fourier transform of $-\Delta f$, where $\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}$ is the classical Laplacian, is given by

$$-\Delta f(x) = \mathcal{F}^{-1}\left(|\cdot|^2 \mathcal{F}[f]\right)(x).$$

It is then a small step to formally change $|\xi|^2$ to $|\xi|^{2s}$ for $s \in (0,1)$, which gives the following definition for the fractional Laplacian, on $\mathcal{S}(\mathbb{R}^d)$.

Definition 2.1.2. *(The fractional Laplacian) For any $f \in \mathcal{S}(\mathbb{R}^d)$, the fractional Laplacian of f (of order s), denoted by $(-\Delta)^s f$, is defined by*

$$(-\Delta)^s f(x) := \mathcal{F}^{-1}\left(|\cdot|^{2s} \mathcal{F}[f]\right)(x), \quad s \in (0,1).$$

Remark 2.1.3. *Although in principle the above definition holds for $s > 1$, we will see from the integral representation below that only when $s \in (0,1)$ are we assured of a 'maximum principle' for the fractional heat equation (3.1). This is one of the reasons why previous literature on the fractional Laplacian has only been concerned with s in this range.*

From the definition, we can see for $s \uparrow 1$ and $s \downarrow 0$, we recover $-\Delta f$ and f as expected.

Remark 2.1.4. *The change $|\xi|^2 \to |\xi|^{2s}$ introduces a decrease in regularity of the function $\xi \mapsto |\xi|^{2s}\mathcal{F}[f](\xi)$ at $\xi = 0$. By properties of the Fourier transform, this* *corresponds in the real variable x to a slow decay at infinity, and thus $(-\Delta)^s f$ is not a Schwartz function since it is no longer rapidly decreasing.*

2.2 The Fractional Laplacian as a Singular Integral

An equivalent way [12, 14] of defining the fractional Laplacian on the space of Schwartz functions $\mathcal{S}(\mathbb{R}^d)$ is given by the following proposition. This singular integral formulation will allow us to extend the class of functions for which the fractional Laplacian is well-defined.

Proposition 2.2.1. *(The fractional Laplacian as a singular integral) For all $f \in \mathcal{S}(\mathbb{R}^d)$,*

$$(-\Delta)^s f(x) = -C_{d,s}\left[\int_{B_r} \frac{f(x+y) - f(x) - \nabla f(x)\cdot y}{|y|^{d+2s}}\,\mathrm{d}y + \int_{\mathbb{R}^d\setminus B_r} \frac{f(x+y)-f(x)}{|y|^{d+2s}}\,\mathrm{d}y\right] \tag{2.1}$$

for every $r > 0$, where $C_{d,s} = \frac{s2^{2s}\Gamma\left(\frac{d+2s}{2}\right)}{\pi^{d/2}\Gamma(1-s)}$ and $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\,\mathrm{d}x$.

It is also equivalent to write

$$(-\Delta)^s f(x) = C_{d,s}\lim_{\epsilon\to 0}\int_{\mathbb{R}^d\setminus B_\epsilon(x)} \frac{f(x)-f(y)}{|x-y|^{d+2s}}\,\mathrm{d}y := C_{d,s}P.V.\int_{\mathbb{R}^d}\frac{f(x)-f(y)}{|x-y|^{d+2s}}\,\mathrm{d}y, \tag{2.2}$$

or

$$(-\Delta)^s f(x) = -\frac{1}{2}C_{d,s}\int_{\mathbb{R}^d}\frac{f(x+y)+f(x-y)-2f(x)}{|y|^{d+2s}}\,\mathrm{d}y. \tag{2.3}$$

Remark 2.2.2. *Following from Remark (2.1.3), we will use representation (2.3) to formally show that when $s \in (0,1)$ we are assured of a 'maximum principle' for the fractional heat equation (3.1).*

Assume u is a smooth solution of (3.1), and the fractional Laplacian of u can be written in the form (2.3). If at some time $t > 0$, u has a global maximum at $x_0 \in \mathbb{R}^d$, then it is easy to see that $(-\Delta)^s u(x_0,t) \geq 0$, and hence $\partial_t u(x_0,t) = -(-\Delta)^s u(x_0,t) \leq 0$. Thus $u(x_0,t') \leq u(x_0,t)$ for all $t' > t$.

If $s > 1$ (assume for simplicity $s = 1+\sigma$ where $0 < \sigma < 1$), then using the Fourier definition we can see,

$$(-\Delta)^s u = (-\Delta)^\sigma\left[-\Delta u\right],$$

and it is not guaranteed that $(-\Delta)^\sigma[-\Delta u](x_0,t) \geq 0$ *if* u *has a global maximum at* x_0 *at time* t.

Proof. The following proof is taken from [14], see also [12]. We first consider the case $s \in (0,1)$ with $d \geq 2$, however the following argument also holds when $d = 1$ if $s > 1/2$. Let $f \in \mathcal{S}(\mathbb{R}^d)$. Then we can write

$$(-\Delta)^s f(x) = -\mathcal{F}^{-1}\left(|\cdot|^{2s-2}\mathcal{F}[\Delta f]\right)(x). \tag{2.4}$$

The function $\xi \mapsto |\xi|^{2s-2}$ is locally integrable for any $s \in (0,1)$, provided $d \geq 2$, since

$$\int_{B_R} |\xi|^{2s-2}\,\mathrm{d}\xi \leq C\int_0^R r^{d+2s-3}\,\mathrm{d}r = CR^{d+2s-2} < \infty$$

for any $R > 0$. It therefore defines a tempered distribution $T_s \in \mathcal{S}'(\mathbb{R}^d)$ defined through its action on elements $\varphi \in \mathcal{S}(\mathbb{R}^d)$ by

$$\langle T_s, \varphi\rangle := \int_{\mathbb{R}^d} |x|^{2s-2}\varphi(x)\,\mathrm{d}x.$$

Therefore we can consider $\mathcal{F}^{-1}(|\cdot|^{2s-2})$ in the sense of distributions, i.e.

$$\left\langle \mathcal{F}^{-1}(T_s), \varphi\right\rangle := \left\langle T_s, \mathcal{F}^{-1}(\varphi)\right\rangle.$$

Let us now show that $\mathcal{F}^{-1}(|\cdot|^{2s-2}) = \overline{C}_{d,s}|\cdot|^{-d-(2s-2)}$ for some constant $\overline{C}_{d,s}$ to be determined. First we recall that a distribution $T \in \mathcal{S}'(\mathbb{R}^d)$ is *homogeneous of degree* a if for all $t > 0$,

$$t^{-d}\langle T, \varphi(\cdot/t)\rangle = t^a \langle T, \varphi\rangle,$$

and *radial* if for all orthogonal transformations A on $\mathbb{R}^d$

$$\langle T, \varphi \circ A\rangle = \langle T, \varphi\rangle.$$

It is easy to see that T_s is homogeneous of degree $2s-2$. Now we show $\mathcal{F}^{-1}(|\cdot|^{2s-2})$

is homogeneous of degree $-d-(2s-2)$. By direct computation

$$\begin{aligned}
t^{-d}\left\langle T_s, \mathcal{F}^{-1}\left(\varphi(\cdot/t)\right)\right\rangle &= t^{-d}\int_{\mathbb{R}^d}|x|^{2s-2}(2\pi)^{-d/2}\left[\int_{\mathbb{R}^d} e^{i\langle \xi,x\rangle}\varphi(\xi/t)\,\mathrm{d}\xi\right]\mathrm{d}x\\
&= \int_{\mathbb{R}^d}|x|^{2s-2}(2\pi)^{-d/2}\left[\int_{\mathbb{R}^d} e^{i\langle \gamma,tx\rangle}\varphi(\gamma)\,\mathrm{d}\gamma\right]\mathrm{d}x\\
&= t^{-d-(2s-2)}\int_{\mathbb{R}^d}|y|^{2s-2}(2\pi)^{-d/2}\left[\int_{\mathbb{R}^d} e^{i\langle \gamma,y\rangle}\varphi(\gamma)\,\mathrm{d}\gamma\right]\mathrm{d}y\\
&= t^{-d-(2s-2)}\left\langle T_s, \mathcal{F}^{-1}\left(\varphi\right)\right\rangle
\end{aligned}$$

where $\gamma=\xi/t$ and $y=tx$ shows that $\mathcal{F}^{-1}\left(|\cdot|^{2s-2}\right)$ is homogeneous of degree $-d-(2s-2)$. It is easily checked that $\mathcal{F}^{-1}\left(|\cdot|^{2s-2}\right)$ is radial. Clearly $T_1(x):=|x|^{-d-2s}$ satisfies these two properties. If T_2 is any other distribution satisfying the same properties, then $\frac{T_2}{T_1}$ is radial and homogeneous of degree 0, i.e. $\frac{T_2}{T_1}$ is a constant. Thus

$$\mathcal{F}^{-1}\left(|\cdot|^{2s-2}\right)=\overline{C}_{d,s}|\cdot|^{-d-(2s-2)} \tag{2.5}$$

for some constant $\overline{C}_{d,s}$, where again equality is in the sense of distributions, i.e.

$$\int_{\mathbb{R}^d}|x|^{2s-2}\mathcal{F}^{-1}\left(\varphi\right)(x)\,\mathrm{d}x=\overline{C}_{d,s}\int_{\mathbb{R}^d}|x|^{-d-(2s-2)}\varphi(x)\,\mathrm{d}x,\quad \forall\varphi\in\mathcal{S}(\mathbb{R}^d).$$

In particular by selecting the test function $e^{-|x|^2/2}$ which is invariant under the Fourier transform, we can find the constant $\overline{C}_{d,s}$.

$$\begin{aligned}
&\int_{\mathbb{R}^d}|x|^{2s-2}e^{-|x|^2/2}\,\mathrm{d}x=\overline{C}_{d,s}\int_{\mathbb{R}^d}|x|^{-d-(2s-2)}e^{-|x|^2/2}\,\mathrm{d}x;\quad \text{setting } r=|x|,\\
&\int_0^\infty r^{d+2s-3}e^{-r^2/2}\,\mathrm{d}r=\overline{C}_{d,s}\int_0^\infty r^{1-2s}e^{-r^2/2}\,\mathrm{d}r;\quad \text{setting } R=r^2/2,\\
&2^{d/2+s-2}\underbrace{\int_0^\infty R^{\frac{d+2s-4}{2}}e^{-R}\,\mathrm{d}R}_{\Gamma\left(\frac{d+2s}{2}-1\right)}=\overline{C}_{d,s}2^{-s}\underbrace{\int_0^\infty R^{-s}e^{-R}\,\mathrm{d}R}_{\Gamma(1-s)}.
\end{aligned}$$

Thus $\overline{C}_{d,s}=2^{2s}\cdot 2^{d/2-2}\frac{\Gamma\left(\frac{d+2s}{2}-1\right)}{\Gamma(1-s)}$. Referring back to (2.4) with (2.5) in hand, we can write

$$(-\Delta)^s f(x) = -(2\pi)^{-d/2}\mathcal{F}^{-1}\left(|\cdot|^{2s-2}\right) * [\Delta f](x)$$
$$= -\frac{2^{2s}\Gamma\left(\frac{d+2s}{2}-1\right)}{4\pi^{d/2}\Gamma(1-s)}|\cdot|^{-d-(2s-2)} * [\Delta f](x).$$

which is well-defined since $|\cdot|^{-d-(2s-2)}$ is locally integrable (for all $s \in (0,1)$) and Δf is a Schwartz function. Therefore

$$(-\Delta)^s f(x) = -\frac{2^{2s}\Gamma\left(\frac{d+2s}{2}-1\right)}{4\pi^{d/2}\Gamma(1-s)}\int_{\mathbb{R}^d}|z|^{-d-(2s-2)}\Delta f(x+z)\,\mathrm{d}z.$$

The idea now is to integrate by parts, but we need to be careful about integrability near 0. For example, formally integrating by parts twice in the above display gives $\int_{\mathbb{R}^d}|z|^{-d-2s}f(x+z)\,\mathrm{d}z$, and it is not clear if this is well-defined.

To this end, let $r > 0, x \in \mathbb{R}^d$ be given and let $\theta \in C_c^\infty(\mathbb{R}^d)$ be an even function with $\theta \equiv 1$ on B_r. Defining the function

$$\phi_x(z) := f(x+z) - f(x) - \nabla f(x)\cdot z\theta(z) \tag{2.6}$$

(which can be seen is of order $|z|^2$ near the origin and bounded at infinity, and thus $z \mapsto |z|^{-d-2s}\phi_x(z)$ is integrable in a neighbourhood of the origin), we have

$$\Delta\phi_x(z) = \Delta f(x+z) + \nabla f(x)\cdot\Delta\left(z\theta(z)\right)$$

and (ignoring the constant)

$$(-\Delta)^s f(x) = -\int_{\mathbb{R}^d}|z|^{-d-(2s-2)}\Delta f(x+z)\,\mathrm{d}z$$
$$= -\int_{\mathbb{R}^d}|z|^{-d-(2s-2)}\Delta\phi_x(z)\,\mathrm{d}z$$
$$- \nabla f(x)\cdot\int_{\mathbb{R}^d}|z|^{-d-(2s-2)}\Delta(z\theta(z))\,\mathrm{d}z,$$

both integrals being well-defined (finite) because $\Delta\phi_x(z)$ and $\Delta(z\theta(z))$ are both Schwartz functions. Since $z \mapsto \Delta(z\theta(z))$ is odd, the second integral vanishes, and we are left with

$$(-\Delta)^s f(x) = -\int_{\mathbb{R}^d}|z|^{-d-(2s-2)}\Delta\phi_x(z)\,\mathrm{d}z.$$

Now we rigorously justify an integration by parts for the above integral. Let $\epsilon > 0$ and define $C_\epsilon := \{z : \epsilon \leq |z| \leq 1/\epsilon\}$ to be the annulus between ϵ and $1/\epsilon$. Then an application of Green's formula gives

$$\begin{aligned} &\int_{C_\epsilon} |z|^{-d-(2s-2)} \Delta\phi_x(z)\,\mathrm{d}z \\ &\quad = \int_{C_\epsilon} \Delta\left(|z|^{-d-(2s-2)}\right)\phi_x(z)\,\mathrm{d}z \\ &\qquad + \int_{\partial C_\epsilon} \left[\phi_x(z)\nabla\left(|z|^{-d-(2s-2)}\right)\cdot n(z) - |z|^{-d-(2s-2)}\nabla\phi_x(z)\cdot n(z)\right]\,\mathrm{d}\sigma_\epsilon(z) \end{aligned} \tag{2.7}$$

where $n(z)$ is the unit outer normal to z, $\partial C_\epsilon = \{|z| = \epsilon\} \cup \{|z| = 1/\epsilon\}$ is the boundary of C_ϵ, and σ_ϵ is the surface measure on ∂C_ϵ. Let us show that the integral over the boundary vanishes as $\epsilon \to 0$.

By a finite Taylor expansion, it is easy to see that in any neighbourhood of the origin (small enough so that $\theta(z) \equiv 1$ there),

$$|\phi_x(z)| \leq C|z|^2, \quad |\nabla\phi_x(z)| \leq C|z|, \quad \text{and} \quad \left|\nabla\left(|z|^{-d-(2s-2)}\right)\right| \leq C|z|^{-d+1-2s}.$$

Thus

$$\begin{aligned} &\left|\int_{\{|z|=\epsilon\}} \phi_x(z)\nabla\left(|z|^{-d-(2s-2)}\right)\cdot n(z) - |z|^{-d-(2s-2)}\nabla\phi_x(z)\cdot n(z)\,\mathrm{d}\sigma_\epsilon(z)\right| \\ &\quad \leq C\epsilon^{-d+3-2s}\int_{\{|x|=\epsilon\}} \mathrm{d}\sigma_\epsilon(z) \\ &\quad \leq C\epsilon^{2-2s} \to 0. \end{aligned}$$

Similarly, since $\nabla\phi_x(z) = \nabla f(x+z)$ for large $|z|$,

$$\begin{aligned} &\left|\int_{\{|z|=1/\epsilon\}} \phi_x(z)\nabla\left(|z|^{-d-(2s-2)}\right)\cdot n(z) - |z|^{-d-(2s-2)}\nabla\phi_x(z)\cdot n(z)\,\mathrm{d}\sigma_\epsilon(z)\right| \\ &\quad \leq C\left(\epsilon^{2s} + \epsilon^{2s-1}\sup_{\{|z|=1/\epsilon\}}|\nabla f(x+z)|\right) \to 0. \end{aligned}$$

In the above argument, the justification that $\epsilon^{2s-1}\sup_{\{|z|=1/\epsilon\}}|\nabla f(x+z)| \to 0$ for $s < 1/2$ $(2s-1 < 0)$ is because f is a Schwartz function. In particular, defining the Schwartz function $g(z) := \nabla f(x+z)$ for the fixed x, and letting $R := \epsilon^{-1}$, we see $\lim_{\epsilon\downarrow 0}\epsilon^{2s-1}\sup_{\{|z|=1/\epsilon\}}|g(z)| = \lim_{R\uparrow\infty} R^{1-2s}\sup_{\{|z|=R\}}|g(z)| = 0$.

Therefore, returning back to (2.7), we know then

$$\begin{aligned}\int_{C_\epsilon} |z|^{-d-(2s-2)} \Delta\phi_x(z)\,\mathrm{d}z &= \int_{C_\epsilon} \Delta\left(|z|^{-d-(2s-2)}\right)\phi_x(z)\,\mathrm{d}z + \mathcal{O}(\epsilon^\alpha)\\ &= 2s(d+2s-2)\int_{C_\epsilon} |z|^{-d-2s}\phi_x(z)\,\mathrm{d}z + \mathcal{O}(\epsilon^\alpha)\end{aligned}$$

for some $\alpha > 0$. Since $|\phi_x(z)| \le C|z|^2$ near the origin, then $\int_{C_\epsilon} |z|^{-d-2s}\phi_x(z)\,\mathrm{d}z$ is integrable for all $\epsilon > 0$. Therefore we can now let $\epsilon \to 0$ to obtain the equality we were looking for,

$$\int_{\mathbb{R}^d} |z|^{-d-(2s-2)}\Delta\phi_x(z)\,\mathrm{d}z = 2s(d+2s-2)\int_{\mathbb{R}^d} |z|^{-d-2s}\phi_x(z)\,\mathrm{d}z.$$

Putting back the constant, we see that

$$\begin{aligned}(-\Delta)^s f(x) &= -\frac{s2^{2s}\left(\frac{d+2s}{2}-1\right)\Gamma\left(\frac{d+2s}{2}-1\right)}{\pi^{d/2}\Gamma(1-s)}\int_{\mathbb{R}^d}|z|^{-d-2s}\phi_x(z)\,\mathrm{d}z\\ &= -\frac{s2^{2s}\Gamma\left(\frac{d+2s}{2}\right)}{\pi^{d/2}\Gamma(1-s)}\int_{\mathbb{R}^d}|z|^{-d-2s}\phi_x(z)\,\mathrm{d}z, \qquad (2.8)\end{aligned}$$

where we have used the property that $(t-1)\Gamma(t-1) = \Gamma(t)$. All that remains is to write $\int_{\mathbb{R}^d}|z|^{-d-2s}\phi_x(z)\,\mathrm{d}z$ in a final form. By definition of ϕ_x and θ, we have

$$\begin{aligned}\int_{\mathbb{R}^d}|z|^{-d-2s}\phi_x(z)\,\mathrm{d}z &= \int_{B_r}\frac{f(x+z)-f(x)-\nabla f(x)\cdot z}{|z|^{-d-2s}}\,\mathrm{d}z\\ &\quad + \int_{\mathbb{R}^d\setminus B_r}\frac{f(x+z)-f(x)-\nabla f(x)\cdot z\theta(z)}{|z|^{-d-2s}}\,\mathrm{d}z.\end{aligned}$$

Since both $\frac{f(x+z)-f(x)}{|z|^{d+2s}}$ and $\frac{\nabla f(x)\cdot z\theta(z)}{|z|^{d+2s}}$ are integrable on $\mathbb{R}^d\setminus B_r$, and $z \mapsto \frac{\nabla f(x)\cdot z\theta(z)}{|z|^{d+2s}}$ is odd,

$$\int_{\mathbb{R}^d\setminus B_r}\frac{\nabla f(x)\cdot z\theta(z)}{|z|^{d+2s}}\,\mathrm{d}z = 0.$$

Hence we obtain (2.1), where, by (2.8), $C_{d,s} = \frac{s2^{2s}\Gamma\left(\frac{d+2s}{2}\right)}{\pi^{d/2}\Gamma(1-s)}$.

The case when $s \in (0, 1/2]$ and $d = 1$ is obtained by an analytic extension argument [14] which we do not give here.

To obtain the other equivalent expressions, we note that

$$\int_{B_r} \frac{f(x+z) - f(x) - \nabla f(x) \cdot z}{|z|^{d+2s}} \, \mathrm{d}z$$

is integrable for all $s \in (0,1)$, and thus in the limit $r \to 0$ it vanishes, leaving

$$\begin{aligned}(-\Delta)^s f(x) &= -C_{d,s} \lim_{r \to 0} \int_{\mathbb{R}^d \setminus B_r} \frac{f(x+z) - f(x)}{|z|^{d+2s}} \, \mathrm{d}z \\ &= C_{d,s} \lim_{r \to 0} \int_{\mathbb{R}^d \setminus B_r(x)} \frac{f(x) - f(y)}{|x-y|^{d+2s}} \, \mathrm{d}y,\end{aligned}$$

which by definition is (2.2). Finally, (2.3) follows by the change of variable $z \mapsto -z$

$$\int_{\mathbb{R}^d \setminus B_r} \frac{f(x+z) - f(x)}{|z|^{d+2s}} \, \mathrm{d}z = \int_{\mathbb{R}^d \setminus B_r} \frac{f(x-z) - f(x)}{|z|^{d+2s}} \, \mathrm{d}z$$

and

$$\int_{B_r} \frac{f(x+z) - f(x) - \nabla f(x) \cdot z}{|z|^{d+2s}} \, \mathrm{d}z = \int_{B_r} \frac{f(x-z) - f(x) + \nabla f(x) \cdot z}{|z|^{d+2s}} \, \mathrm{d}z,$$

from which

$$\begin{aligned}\int_{\mathbb{R}^d \setminus B_r} \frac{f(x+z) - f(x)}{|z|^{d+2s}} \, \mathrm{d}z &= \frac{1}{2} \int_{\mathbb{R}^d \setminus B_r} \frac{f(x+z) + f(x-z) - 2f(x)}{|z|^{d+2s}} \, \mathrm{d}z \\ \int_{B_r} \frac{f(x+z) - f(x) - \nabla f(x) \cdot z}{|z|^{d+2s}} \, \mathrm{d}z &= \frac{1}{2} \int_{B_r} \frac{f(x+z) + f(x-z) - 2f(x)}{|z|^{d+2s}} \, \mathrm{d}z,\end{aligned}$$

giving (2.3). $\square$

The integral representation allows us to extend the pointwise fractional Laplacian definition to functions which do not have as nice smoothness and integrability properties as Schwartz functions [24]. We will be content with showing the integral representation makes sense for functions belonging to certain Hölder spaces. Indeed we have the following from [24].

Proposition 2.2.3. *[24] Let $f \in C^{0,\alpha}(\mathbb{R}^d)$ for some $2s < \alpha \leq 1$. Then $(-\Delta)^s f \in C^{0,\alpha-2s}$. If, in addition, f is bounded, then $(-\Delta)^s f \in L^\infty(\mathbb{R}^d)$.*

Remark 2.2.4. *As $s \uparrow 1$, we see that there exists no α satisfying $2s < \alpha \leq 1$. Indeed, this is the case for $s \geq 1/2$, and therefore we cannot expect $(-\Delta)^s$ to be well-*

defined for $C^{0,\alpha}$ functions for s in this range. We might anticipate this if we think of $-(-\Delta)^{1/2} \approx \nabla$ and $-(-\Delta)^1 = \Delta$, since, in general, $C^{0,\alpha}$ functions do not possess any smoothness properties. Thus when s passes above $1/2$ we 'require' at least one derivative, and when $s = 1$ we need two (see Proposition 2.2.5 below) for $(-\Delta)^s$ to be well-defined.

Proof. Fix $x_1, x_2 \in \mathbb{R}^d$, and let $R := |x_1 - x_2|$. Then for $i = 1, 2$,

$$\int_{B_R} \frac{|f(x_i + z) + f(x_i - z) - 2f(x_i)|}{|z|^{d+2s}} \, \mathrm{d}z \leq C|x_1 - x_2|^{\alpha - 2s}.$$

Outside B_R, we have

$$\frac{|f(x_1 + z) + f(x_1 - z) - 2f(x_1) - f(x_2 + z) - f(x_2 - z) + 2f(x_2)|}{|z|^{d+2s}} \leq C\frac{|x_1 - x_2|^\alpha}{|z|^{d+2s}},$$

and

$$\int_{\mathbb{R}^d \setminus B_R} \frac{|x_1 - x_2|^\alpha}{|z|^{d+2s}} \, \mathrm{d}z \leq C|x_1 - x_2|^\alpha R^{-2s} = C|x_1 - x_2|^{\alpha - 2s}.$$

Thus it follows $|(-\Delta)^s f(x_1) - (-\Delta)^s f(x_2)| \leq C|x_1 - x_2|^{\alpha - 2s}$.

If, in addition, f is bounded, it is easy to see $(-\Delta)^s f \in L^\infty(\mathbb{R}^d)$, since for any $R > 0$,

$$\int_{B_R} \frac{|f(x + z) + f(x - z) - 2f(x)|}{|z|^{d+2s}} \, \mathrm{d}z \leq CR^{\alpha - 2s}$$
$$\int_{\mathbb{R}^d \setminus B_R} \frac{|f(x + z) + f(x - z) - 2f(x)|}{|z|^{d+2s}} \, \mathrm{d}z \leq CR^{-2s}.$$

□

Similar ideas used in the above can be used to prove the following.

Proposition 2.2.5. *[24] Let $f \in C^{1,\alpha}(\mathbb{R}^d)$ for some $0 < \alpha \leq 1$.*

1. *If $\alpha > 2s$, then $(-\Delta)^s f \in C^{1,\alpha - 2s}(\mathbb{R}^d)$.*
2. *If $\alpha < 2s$, then $(-\Delta)^s f \in C^{0,\alpha - 2s + 1}(\mathbb{R}^d)$.*

Additionally, if $1 + \alpha > 2s$ and f is bounded, then $(-\Delta)^s f \in L^\infty(\mathbb{R}^d)$.

Proof. We only show $(-\Delta)^s f \in L^\infty(\mathbb{R}^d)$ if f is also bounded and $1 + \alpha > 2s$, since we will need this estimate later. We refer to [24] for proof of the other claims. For fixed $R > 0$, it is easy to estimate using representation (2.1)

$$|f(x+z) - f(x) - \nabla f(x) \cdot z| \leq C|z| \, |\nabla f(x + \lambda z) - \nabla f(x)| \leq C|z|^{1+\alpha},$$

where $|\lambda| < 1$ from a first-order Taylor expansion with Lagrange remainder. Thus

$$\int_{B_R} \frac{|f(x+z) - f(x) - \nabla f(x) \cdot z|}{|z|^{d+2s}} \, \mathrm{d}z \leq C \int_0^R r^{\alpha - 2s} \, \mathrm{d}r \leq C R^{1+\alpha-2s}.$$

The other integral over $\mathbb{R}^d \backslash B_R$ is easily seen to be uniformly bounded in x because f is bounded. □

Finally it comes as no surprise that $(-\Delta)^s f$ still retains nice regularity and integrability properties when $f \in C_c^\infty(\mathbb{R}^d)$.

Lemma 2.2.6. *Let $f \in C_c^\infty(\mathbb{R}^d)$. Then $(-\Delta)^s f \in L^p \cap C^\infty(\mathbb{R}^d)$ for every $1 \leq p \leq \infty$.*

Proof. The boundedness of $(-\Delta)^s f$ follows as in the above, and smoothness is by differentiation under the integral. We show $(-\Delta)^s f \in L^1(\mathbb{R}^d)$. Fix $R, R' > 0$ such that $\operatorname{spt}(f) \subset B_{R'}$, and let $g_{R,s}(x) := \int_{B_R} \frac{|f(x+z) - f(x) - \nabla f(x) \cdot z|}{|z|^{d+2s}} \, \mathrm{d}z$. It is easy to see that $\operatorname{spt}(g_{R,s}) \subset B_{R+R'}$ and $g_{R,s} \in L^\infty(\mathbb{R}^d)$. Then

$$\int_{\mathbb{R}^d} |(-\Delta)^s f(x)| \, \mathrm{d}x \leq \|g_{R,s}\|_{L^\infty(\mathbb{R}^d)} |B_{R+R'}| + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d \backslash B_R} \frac{|f(x+z) - f(x)|}{|z|^{d+2s}} \, \mathrm{d}z \, \mathrm{d}x.$$

where $|B_{R+R'}|$ denotes the Lebesgue measure of $B_{R+R'}$. To estimate the last integral, we can write

$$\begin{aligned}
\int_{\mathbb{R}^d \backslash B_R} |z|^{-d-2s} \int_{\mathbb{R}^d} &|f(x+z) - f(x)| \, \mathrm{d}x \, \mathrm{d}z \\
&\leq \int_{\mathbb{R}^d \backslash B_R} |z|^{-d-2s} \int_{\mathbb{R}^d} \left(|f(x+z)| + |f(x)|\right) \, \mathrm{d}x \, \mathrm{d}z \\
&\leq 2 \, \|f\|_{L^1(\mathbb{R}^d)} \int_{\mathbb{R}^d \backslash B_R} |z|^{-d-2s} \, \mathrm{d}z < \infty.
\end{aligned}$$

Since $(-\Delta)^s f \in L^1 \cap L^\infty(\mathbb{R}^d)$, it is therefore in every L^p space, $1 \leq p \leq \infty$, by interpolation. □

2.2.1 Equality of Fourier and Singular Integral Representation on Non-Schwartz Functions

Now we turn to the following question: Suppose that f is not a Schwartz function, but $\mathcal{F}^{-1}\left(|\cdot|^{2s}\mathcal{F}\left[f\right]\right)(x)$ is well-defined, and also $\int_{\mathbb{R}^d}\frac{f(x+z)+f(x-z)-2f(x)}{|z|^{d+2s}}\,\mathrm{d}z$ is well defined. Do they agree? That is, $(-\Delta)^s f$ should not depend on which representation we use, if both exist. To begin, we have the following.

Lemma 2.2.7. *Let $f\in L^1(\mathbb{R}^d)$. Denote*

$$\begin{aligned}
A_s(f)(x) &:= -\frac{1}{2}C_{d,s}\int_{\mathbb{R}^d}\frac{f(x+z)+f(x-z)-2f(x)}{|z|^{d+2s}}\,\mathrm{d}z, \quad \textit{and}\\
B_s(f)(x) &:= \mathcal{F}^{-1}\left(|\cdot|^{2s}\mathcal{F}\left[f\right]\right)(x).
\end{aligned}$$

If $A_s(f)\in L^\infty(\mathbb{R}^d)$, and $|\cdot|^{2s}\hat{f}\in L^1(\mathbb{R}^d)$, then the respective equalities

$$\int_{\mathbb{R}^d}A_s(f)\eta\,\mathrm{d}x=\int_{\mathbb{R}^d}fA_s(\eta)\,\mathrm{d}x \quad \textit{and} \quad \int_{\mathbb{R}^d}B_s(f)\eta\,\mathrm{d}x=\int_{\mathbb{R}^d}fB_s(\eta)\,\mathrm{d}x,$$

hold for every $\eta\in C_c^\infty(\mathbb{R}^d)$.

Proof. Suppose $A_s(f)$ is in $L^\infty(\mathbb{R}^d)$. Since

$$\begin{aligned}
\left|\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{f(x+z)+f(x-z)-2f(x)}{|z|^{d+2s}}\eta(x)\,\mathrm{d}z\,\mathrm{d}x\right| &= \left|\int_{\mathbb{R}^d}A_s(f)\eta\,\mathrm{d}x\right|\\
&\le \left\|A_s(f)\right\|_{L^\infty(\mathbb{R}^d)}\left\|\eta\right\|_{L^1(\mathbb{R}^d)},
\end{aligned}$$

then we can apply the Fubini-Tonelli theorem to interchange the integrals in x and z,

$$\begin{aligned}
&\int_{\mathbb{R}^d}\eta(x)\int_{\mathbb{R}^d}\frac{f(x+z)+f(x-z)-2f(x)}{|z|^{d+2s}}\,\mathrm{d}z\,\mathrm{d}x\\
&\quad=\int_{\mathbb{R}^d}|z|^{-d-2s}\left[\int_{\mathbb{R}^d}\eta(x)f(x+z)\,\mathrm{d}x+\int_{\mathbb{R}^d}\eta(x)f(x-z)\,\mathrm{d}x-2\int_{\mathbb{R}^d}f(x)\eta(x)\,\mathrm{d}x\right]\mathrm{d}z\\
&\quad=\int_{\mathbb{R}^d}|z|^{-d-2s}\left[\int_{\mathbb{R}^d}\eta(x-z)f(x)\,\mathrm{d}x+\int_{\mathbb{R}^d}\eta(x+z)f(x)\,\mathrm{d}x-2\int_{\mathbb{R}^d}f(x)\eta(x)\,\mathrm{d}x\right]\mathrm{d}z\\
&\quad=\int_{\mathbb{R}^d}|z|^{-d-2s}\left[\int_{\mathbb{R}^d}f(x)\left(\eta(x+z)+\eta(x-z)-2\eta(x)\right)\,\mathrm{d}x\right]\mathrm{d}z,
\end{aligned}$$

and we can interchange the last integral since $f\in L^1(\mathbb{R}^d)$ and $(-\Delta)^s\eta\in L^\infty(\mathbb{R}^d)$ to conclude the result.

The condition $|\cdot|^{2s}\hat{f} \in L^1(\mathbb{R}^d)$ implies that $B_s(f) \in L^\infty(\mathbb{R}^d)$. We write

$$\begin{aligned}\int_{\mathbb{R}^d} B_s(f)\eta \,\mathrm{d}x &= \mathcal{F}\left[B_s(f)\eta\right](\xi = 0) \\ &= \mathcal{F}\left[B_s(f)\right] * \mathcal{F}\left[\eta\right](\xi = 0) \\ &= \left[|\cdot|^{2s}\hat{f}\right] * \hat{\eta}(\xi = 0) \\ &= \int_{\gamma\in\mathbb{R}^d} |0-\gamma|^{2s}\hat{f}(0-\gamma)\hat{\eta}(\gamma)\,\mathrm{d}\gamma\end{aligned}$$

and the conclusion follows by noting that we can change $\gamma \to -\gamma$, and reverse the steps to get

$$\int_{\mathbb{R}^d} f B_s(\eta)\,\mathrm{d}x.$$

□

Lemma 2.2.8. *Let A_s, B_s be defined as in Lemma 2.2.7, and suppose $f \in L^1(\mathbb{R}^d)$, $|\cdot|^{2s}\hat{f} \in L^1(\mathbb{R}^d)$, and $A_s(f) \in L^\infty(\mathbb{R}^d)$. Then $A_s(f) = B_s(f)$ a.e. $x \in \mathbb{R}^d$.*

Proof. By the previous lemma, and equality of A_s and B_s on the space of Schwartz functions,

$$\int_{\mathbb{R}^d} A_s(f)\eta\,\mathrm{d}x = \int_{\mathbb{R}^d} f A_s(\eta)\,\mathrm{d}x = \int_{\mathbb{R}^d} f B_s(\eta)\,\mathrm{d}x = \int_{\mathbb{R}^d} B_s(f)\eta\,\mathrm{d}x,$$

for all $\eta \in C_c^\infty(\mathbb{R}^d)$. Hence $A_s(f) = B_s(f)$ a.e., and we may use $(-\Delta)^s f$ without ambiguity. □

2.2.2 Integration by Parts

For convenience we extend Lemma 2.2.7.

Lemma 2.2.9. *(Integration by Parts) Let $f, g \in L^1 \cap L^\infty(\mathbb{R}^d)$, with $(-\Delta)^s f, (-\Delta)^s g \in L^\infty(\mathbb{R}^d)$. Then*

$$\int_{\mathbb{R}^d} \left[(-\Delta)^s f\right] g\,\mathrm{d}x = \int_{\mathbb{R}^d} f\left[(-\Delta)^s g\right]\,\mathrm{d}x.$$

Proof. See Lemma 2.2.7. We impose $f, g \in L^\infty(\mathbb{R}^d)$ as well as in $L^1(\mathbb{R}^d)$ so that the integrals, e.g. $\int_{\mathbb{R}^d} g(x)f(x+z)\,\mathrm{d}x$, are finite. □

Chapter 3

The Fractional Heat Equation

In this section, we are interested in studying solutions to the *fractional heat equation*,

$$\begin{cases} \partial_t u = -(-\Delta)^s u \quad \text{in } \mathbb{R}^d \times (0,\infty), \quad s \in (0,1) \\ \quad u = u^0 \quad \text{ on } \mathbb{R}^d \times \{t=0\}, \end{cases} \tag{3.1}$$

where u^0 is a probability density on $\mathbb{R}^d$.

3.1 Properties of Solutions to the Fractional Heat Equation

Recall that solutions to the classical heat equation on $\mathbb{R}^d$ are obtained by convolving the initial data with the Gaussian heat kernel,

$$\frac{1}{(4\pi t)^{d/2}} e^{-|x|^2/4t}.$$

Moreover, these solutions are smooth, except possibly at $t = 0$, and satisfy a maximum principle [16]. The solutions to (3.1) also turn out to have many of the same properties.

We give a formal discussion first. Suppose $u = u(x,t)$ solves (3.1). Then taking the Fourier transform of (3.1) gives

$$\begin{cases} \partial_t \hat{u}(\xi,t) = -|\xi|^{2s}\hat{u}(\xi,t), \quad \xi \in \mathbb{R}^d \\ \hat{u}(\xi,0) = \hat{u^0}(\xi). \end{cases}$$

This has solution

$$\hat{u}(\xi, t) = e^{-t|\xi|^{2s}}\hat{u^0}(\xi),$$

which upon inverting back to real space, and using the convolution property of the Fourier transform, yields

$$u(x,t) = \frac{1}{(2\pi)^{d/2}}\mathcal{F}^{-1}\left(e^{-t|\cdot|^{2s}}\right) * u^0(x).$$

Thus we can define the 'fractional heat kernel' Φ_s to be

$$\Phi_s(x,t) := \frac{1}{(2\pi)^{d/2}}\mathcal{F}^{-1}\left(e^{-t|\cdot|^{2s}}\right)(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle}e^{-t|\xi|^{2s}}\,\mathrm{d}\xi, \quad t > 0. \tag{3.2}$$

It is the solution to the fractional heat equation (3.1) when the initial distribution is a point source. For general s, Φ_s is not known explicitly; when $s = 1$, the Gaussian heat kernel is recovered. Thus, in some sense, the classical heat equation is just one member of a family of equations parametrized by s, where each kernel Φ_s is the generator of a contraction semigroup on L^1 [16], in the language of semigroup theory.

Some basic properties that we anticipate of the fractional heat kernel include the following. Since derivatives transform to powers of ξ under the Fourier transform, and $e^{-t|\xi|^{2s}}$ vanishes faster than any function with polynomial growth in ξ, we expect $\Phi_s \in C^\infty(\mathbb{R}^d \times (0,\infty))$. Moreover, since $s < 1$ we also formally see that unlike the classical Gaussian case, $\Phi_s(t)$ has an infinite second moment, since computing $\int_{\mathbb{R}^d} |x|^2\Phi_s(x,t)\,\mathrm{d}x$ is the same as computing the second derivative of the Fourier transform $\frac{\partial^2}{\partial\xi^2}e^{-t|\xi|^{2s}}$ at $\xi = 0$, which is singular, since $\lim_{|\xi|\to 0}|\xi|^{2s-2} = +\infty$. This means that the fractional heat kernel $\Phi_s(t)$ decays much more slowly than its Gaussian counterpart.

We now list some standard properties that Φ_s satisfies, which will be used in the sequel. Some of the following are taken from [13].

Proposition 3.1.1. *The fractional heat kernel Φ_s given by (3.2) satisfies the following properties. For every $t > 0$,*

1. $\partial_t\Phi_s(x,t) = -(-\Delta)^s\Phi_s(x,t)$, *for all* $x \in \mathbb{R}^d$.

2. *(A Scaling Property)* $\Phi_s(x,t) = t^{-d/2s}\Phi_s(t^{-1/2s}x, 1)$,

3. *(Regularization)* $\Phi_s \in C^\infty(\mathbb{R}^d \times (0,\infty))$,

4. *(Radial Symmetry)* $\Phi_s(x,t) = \Phi_s(|x|,t)$,

5. *(A two-sided estimate)*

$$C^{-1}\left(t^{-d/2s} \wedge \frac{t}{|x|^{d+2s}}\right) \leq \Phi_s(x,t) \leq C\left(t^{-d/2s} \wedge \frac{t}{|x|^{d+2s}}\right) \tag{3.3}$$

for all $x \in \mathbb{R}^d$, *where* $a \wedge b := \min\{a,b\}$ *for* $a, b \in \mathbb{R}$. *In particular,* $\Phi_s(t)$ *is nonnegative.*

6. *(Unit)* $\|\Phi_s(t)\|_{L^1(\mathbb{R}^d)} = 1$,

7. *(Infinite Second Moment)* $\int_{\mathbb{R}^d} |x|^2 \Phi_s(x,t)\,\mathrm{d}x = +\infty$ *for every* $s \in (0,1)$.

Remark 3.1.2. *The inequality (3.3) for* Φ_s *translates to*

$$\left.\begin{array}{l} C^{-1}t^{-d/2s} \\ C^{-1}\frac{t}{|x|^{d+2s}} \end{array}\right\} \leq \Phi_s(x,t) \leq \left\{\begin{array}{ll} Ct^{-d/2s}, & |x| \leq t^{1/2s} \\ C\frac{t}{|x|^{d+2s}}, & |x| > t^{1/2s}. \end{array}\right.$$

Proof. 1. This follows immediately from the definition of Φ_s.

2. By definition $\Phi_s(x,t) = (2\pi)^{-d/2}\mathcal{F}^{-1}\left(e^{-t|\cdot|^{2s}}\right)(x) = (2\pi)^{-d}\int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle}e^{-t|\xi|^{2s}}\,\mathrm{d}\xi$. By rescaling $\gamma = t^{1/2s}\xi$, we obtain the result.

3. For any multiindex α, it is easy to see that the function $\xi \mapsto |\xi|^{|\alpha|}e^{-t|\xi|^{2s}}$ is integrable over $\xi \in \mathbb{R}^d$ for $t > 0$. (Indeed, it is enough to show that $r^{k+d-1}e^{-tr^{2s}} \leq 1$ for all large enough r, where $r = |\xi|$ and $|\alpha| = k$.) Therefore

$$\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle}(i\xi)^\alpha e^{-t|\xi|^{2s}}\,\mathrm{d}\xi$$

exists, which by properties of the Fourier transform is exactly $D_x^\alpha \Phi_s(x,t)$. Moreover, since $e^{-t|\xi|^{2s}}$ is infinitely differentiable with respect to t, by differentiation under the integral, all t-derivatives of Φ_s also exist.

4. If $R : \mathbb{R}^d \to \mathbb{R}^d$ is a rotation operator, so that $|Rx| = |x|$, then

$$\Phi_s(Rx,t) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\langle Rx,\xi\rangle}e^{-t|\xi|^{2s}}\,\mathrm{d}\xi = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\left\langle x,R^{-1}\xi\right\rangle}e^{-t|\xi|^{2s}}\,\mathrm{d}\xi$$

and the result follows by a change of variable $\gamma = R^{-1}\xi$.

5. Let us first establish the (seemingly obvious) property that $\Phi_s(x,1)$ is strictly positive for all $x \in \mathbb{R}^d$. Let $y \in \mathbb{R}^d$ satisfy $|y| = 1/|x|$ for $x \neq 0$. Then since

$$\Phi_s(x,1) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\langle x,\xi\rangle} e^{-|\xi|^{2s}}\,\mathrm{d}\xi = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \cos\left(\langle x,\xi\rangle\right) e^{-|\xi|^{2s}}\,\mathrm{d}\xi$$
$$\geq -\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-|\xi|^{2s}}\,\mathrm{d}\xi,$$

it follows that

$$\Phi_s(x,1)\Phi_s(y,1) \geq \left[\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-|\xi|^{2s}}\,\mathrm{d}\xi\right]^2 > 0. \tag{3.4}$$

This implies that $\Phi_s(x,1) \neq 0$ for all $x \in \mathbb{R}^d \setminus \{0\}$. Moreover, since $\Phi_s(0,1) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-|\xi|^{2s}}\,\mathrm{d}\xi > 0$, we must also have $\Phi_s(x,1) > 0$ for all $x \in \mathbb{R}^d \setminus \{0\}$, for otherwise $\Phi_s(x,1) < 0$ implies, by continuity of $\Phi_s(\cdot,1)$, that there exists $z \in \mathbb{R}^d$, $0 < |z| < |x|$ satisfying $\Phi_s(z,1) = 0$, which is strictly forbidden.

By the scaling property we then conclude $\Phi_s(t) > 0$ for all $t > 0$.

Now we establish the estimates. By the scaling property above,

$$\Phi_s(x,t) = \frac{t^{-d/2s}}{(2\pi)^d}\int_{\mathbb{R}^d} e^{i\left\langle t^{-1/2s}x,\xi\right\rangle} e^{-|\xi|^{2s}}\,\mathrm{d}\xi \leq Ct^{-d/2s}\int_{\mathbb{R}^d} e^{-|\xi|^{2s}}\,\mathrm{d}\xi \leq Ct^{-d/2s}$$

for every $t > 0$ and $x \in \mathbb{R}^d$. This gives one of the estimates. For the other estimate, we extract from [7] the result

$$\lim_{|x|\to\infty} |x|^{d+2s}\Phi_s(x,1) = C.$$

Therefore using the scaling property again, we have

$$\Phi_s(x,t) \leq C\frac{t}{|x|^{d+2s}}, \quad \text{large } |x|, t > 0.$$

Since $\Phi_s(\cdot,t)$ is continuous, it is bounded in a ball centred at the origin, and since $C\frac{t}{|x|^{d+2s}} \to \infty$ as $|x| \to 0$, we can choose C large enough so that the above estimate holds for all $x \neq 0 \in \mathbb{R}^d$,

$$\Phi_s(x,t) \leq C\frac{t}{|x|^{d+2s}}, \quad t > 0, x \in \mathbb{R}^d\setminus\{0\}.$$

For the reverse inequality, we let $y \in \mathbb{R}^d$ satisfy $|y| = 1/|x|$ for $x \neq 0$. Then the above estimates give

$$\begin{aligned} C\frac{t}{|x|^{d+2s}} &\leq \frac{1}{\Phi_s(y,1/t)}, \\ Ct^{-d/2s} &\leq \frac{1}{\Phi_s(y,1/t)} \end{aligned}$$

for $t > 0$. Now we use (3.4) to have $C\frac{1}{\Phi_s(y,1/t)} \leq \Phi_s(x,t)$ and obtain the result.

6. Note that for every $t > 0$,

$$\int_{\mathbb{R}^d} \Phi_s(x,t)\,\mathrm{d}x = (2\pi)^{d/2}\mathcal{F}\left[\Phi_s(t)\right](\xi = 0) = e^{-t|0|^{2s}} = 1.$$

7. By (3.3), for any $t > 0$ and $R > t^{1/2s}$,

$$\int_{B_R} |x|^2\Phi_s(t,x)\,\mathrm{d}x \geq Ct\int_{t^{1/2s}}^{R} r^{1-2s}\,\mathrm{d}r \geq Ct\left(R^{2-2s} - t^{(1-s)/s}\right).$$

Thus $\int_{B_R} |x|^2\Phi_s(t,x)\,\mathrm{d}x \uparrow \infty$ as $R \uparrow \infty$.

$\square$

Corollary 3.1.3. *Define u by*

$$u(x,t) := \Phi_s(t) * u^0(x), \quad t > 0, \tag{3.5}$$

where u^0 is a probability density on $\mathbb{R}^d$. Then

1. $u \in C^\infty(\mathbb{R}^d \times (0,\infty))$,

2. $\partial_t u(x,t) = -(-\Delta)^s u(x,t)$ *for* $x \in \mathbb{R}^d$ *and* $t > 0$,

3. $\|u(t)\|_{L^1(\mathbb{R}^d)} = \|u^0\|_{L^1(\mathbb{R}^d)}$.

Chapter 4

The Transport Equation as a Gradient Flow

In this chapter we want to pursue the view that the linear transport equation

$$\begin{cases} \partial_t v = \operatorname{div}(v \nabla \Psi) \\ v(0) = v^0 \in \mathcal{P}_a^2(\mathbb{R}^d). \end{cases} \tag{4.1}$$

is a gradient flow of the potential energy $\int_{\mathbb{R}^d} \rho \Psi$ with respect to the 2-Wasserstein distance. In order to proceed with the splitting scheme in Chapter 5, such a development is not strictly necessary. Indeed, it is straightforward to obtain the existence of a weak solution to (4.1) by applying the method of characteristics [21]. However, we like to think that viewing (4.1) as a gradient flow of the potential energy is a more 'natural' viewpoint of the dynamics, and this is what we develop here.

We use a time-discrete variational scheme to prove the gradient flow assertion. The scheme will be a simplification of the one used in [19], which is introduced in Section 4.3. We first give a brief motivation for gradient flows in metric spaces.

4.1 Gradient Flow in Metric Spaces

A large amount of theory has been developed about the notion of gradient flows in metric spaces, especially in the now-classic book by Ambrosio, Gigli, and Savaré [3]. Here we attempt to explain somewhat formally one way to extend the usual notion of a gradient flow in $\mathbb{R}^d$ to metric spaces. This approach is sometimes called the Minimizing Movement Scheme [3].

The classical notion of a gradient flow in $\mathbb{R}^d$ is defined by a function $f \in C^1(\mathbb{R}^d)$, and the equation

$$\begin{cases} \dot{x}(t) = -\nabla f(x(t)), \quad t > 0, \\ x(0) = x^0 \in \mathbb{R}^d. \end{cases} \tag{4.2}$$

A C^1 solution $x : \mathbb{R} \to \mathbb{R}^d$ is the gradient flow of f if it satisfies (4.2) [11].

In a metric space, we may have no structure other than the metric itself. With this in mind, let us fix a time step $\tau > 0$ and apply an implicit Euler scheme to (4.2)

$$\frac{x_\tau^n - x_\tau^{n-1}}{\tau} = -\nabla f(x_\tau^n) \tag{4.3}$$

where x_τ^n approximates (4.2) at time $t_n := n\tau$. We note that x_τ^n solves (4.3) if and only if x_τ^n is the minimizer of

$$x \mapsto \frac{1}{2\tau}|x - x_\tau^{n-1}|^2 + f(x) \tag{4.4}$$

under suitable assumptions on f (e.g. f convex). In this fashion, we obtain a discrete-time sequence $\left\{x_\tau^k\right\}_{k=0,1,\ldots}$ for the given τ. To investigate the limit $\tau \downarrow 0$, we construct by interpolation a function $x_\tau = x_\tau(t)$ defined for all time, and attempt to obtain compactness of the sequence $\{x_\tau\}_{\tau\downarrow 0}$ in some suitable topology. The topology should be strong enough to deduce that the limit function $x = x(t)$ is a solution to (4.2).

In the above, for instance, if x_τ is a linear interpolation of the x_τ^k,

$$x_\tau(t) := \frac{t_n - t}{\tau} x_\tau^{n-1} + \frac{t - t_{n-1}}{\tau} x_\tau^n, \quad t \in [t_{n-1}, t_n]$$

then we have the following taken from [11]. Suppose for simplicity f is convex and ∇f is Lipschitz. To obtain compactness of $\{x_\tau\}_{\tau\downarrow 0}$, we have the estimate

$$|x_\tau'(t)| = \frac{|x_\tau^n - x_\tau^{n-1}|}{\tau} = |\nabla f(x_\tau^n)| \leq \left|\nabla f(x_\tau^{n-1})\right|, \quad t \in [t_{n-1}, t_n].$$

(This follows because $|[y + \tau\nabla f(y)] - [z + \tau\nabla f(z)]| \geq |y - z|$ for $y, z \in \mathbb{R}^d$ (by convexity of f), and $x_\tau^n - x_\tau^{n-1} = -\tau\nabla f(x_\tau^n)$. Then

$$\tau|\nabla f(x_\tau^n)| = |x_\tau^n - x_\tau^{n-1}| \leq \left|x_\tau^n + \tau\nabla f(x_\tau^n) - (x_\tau^{n-1} + \tau\nabla f(x_\tau^{n-1}))\right| \leq \tau|\nabla f(x_\tau^{n-1})|.)$$

Thus

$$|x'_\tau(t)| \leq \left|\nabla f(x^0_\tau)\right| = \left|\nabla f(x^0)\right|$$

is uniformly bounded above. Therefore $\{x_\tau\}_{\tau\downarrow 0}$ is compact w.r.t. the uniform norm on any finite time interval $[0,T]$ by the Ascoli-Arzelà theorem [10], and converges up to a subsequence to some x. To deduce that x solves (4.2) we introduce [11] the piecewise constant interpolant

$$\bar{x}_\tau(t) := x^n_\tau, \quad t \in (t_{n-1}, t_n],$$

and note that

$$|x_\tau(t) - \bar{x}_\tau(t)| \leq \left|x^n_\tau - x^{n-1}_\tau\right| = \tau\left|\nabla f(x^{n+1}_\tau)\right| \leq \tau\left|\nabla f(x^0)\right|$$

for $t \in (t_{n-1}, t_n]$. We also have that $x'_\tau(t) = -\nabla f(\bar{x}_\tau(t))$ a.e. t, from which we have the integrated form

$$x_\tau(t) - x^0 = -\int_0^t \nabla f(\bar{x}_\tau(s))\,\mathrm{d}s.$$

Then for $t \in [0,T]$,

$$\begin{aligned}\left|x(t) - x^0 + \int_0^t \nabla f(x(s))\,\mathrm{d}s\right| &\leq |x(t) - x_\tau(t)| + \int_0^t |\nabla f(x(s)) - \nabla f(\bar{x}_\tau(s))|\,\mathrm{d}s \\ &\leq |x(t) - x_\tau(t)| + \int_0^t |x(s) - \bar{x}_\tau(s)|\,\mathrm{d}s.\end{aligned}$$

Since $|x(s) - \bar{x}_\tau(s)| \leq |x(s) - x_\tau(s)| + \tau\,|\nabla f(x^0)|$, it follows that x solves (4.2).

Returning to the task of generalizing the notion of a gradient flow, since (4.4) involves only the Euclidean distance, the scheme makes sense for a general metric space (X,d),

$$\textit{Minimize} \quad x \mapsto \frac{1}{2\tau} d(x, x^{n-1}_\tau)^2 + \mathcal{F}(x) \quad \textit{over all } x \in X.$$

where $\mathcal{F} : X \to \mathbb{R}$ is a functional on X.

If X is a function space, existence of a minimizer can be established through the Direct Method in the Calculus of Variations. One important step in this is to establish compactness of a minimizing sequence in some topology. The topology can be weaker than the topology induced by the metric d, but the functional $x \mapsto \frac{1}{2\tau}d(x, x^{n-1}_\tau)^2 + \mathcal{F}(x)$ should be lower semi-continuous w.r.t. this topology.

As before, we obtain a discrete-time sequence $\{x_\tau^k\}_{k=0,1,\dots} \subset X$ for each $\tau > 0$ to interpolate with, giving $x_\tau = x_\tau(t)$. If we can then obtain compactness of the sequence $\{x_\tau\}_{\tau\downarrow 0}$ in a topology for which we can deduce that the limit x solves some given PDE (in, eg. the weak sense), then we say that this PDE is a gradient flow, or steepest descent, of the functional $\mathcal{F}$, with respect to the metric d on the space X.

In the following sections we are going to establish that the transport equation is a gradient flow of the potential energy in the 2-Wasserstein metric on the space $\mathcal{P}_a^2(\mathbb{R}^d)$ in the sense described above. But first, we need some definitions and results. We first establish the definition of a weak solution to (4.1).

Definition 4.1.1. *Given $T < \infty$, a function $v : \mathbb{R}^d \times (0,T) \to [0,\infty)$ is a weak solution of (4.1) if $\int_{\mathbb{R}^d} v(t)\,\mathrm{d}x = \int_{\mathbb{R}^d} v^0\,\mathrm{d}x$ for a.e. $t \in (0,T)$, and*

$$\int_0^T \int_{\mathbb{R}^d} v(t)\partial_t\varphi(t)\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^d} v^0\varphi(0)\,\mathrm{d}x = \int_0^T \int_{\mathbb{R}^d} v(t)\nabla\Psi\cdot\nabla\varphi(t)\,\mathrm{d}x\,\mathrm{d}t \tag{4.5}$$

for all $\varphi \in C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ with time support in $[-T,T]$.

4.2 Optimal Transportation & the 2-Wasserstein Distance

An important definition in this section is the push-forward.

Definition 4.2.1. *(Push forward) [28] Let μ, ν be two probability measures on $\mathbb{R}^d$. A map $T : \mathbb{R}^d \to \mathbb{R}^d$ is said to push μ forward to ν (or ν is the push-forward of μ by the map T) and we write $T\#\mu = \nu$ if for all ν-measurable $B \subset \mathbb{R}^d$,*

$$\nu\left[B\right] = \mu\left[T^{-1}(B)\right],$$

or, alternatively, for every $\xi \in L^1(\mathrm{d}\nu)$,

$$\int_{\mathbb{R}^d} \xi\,\mathrm{d}\nu = \int_{\mathbb{R}^d} \xi \circ T\,\mathrm{d}\mu.$$

The interpretation of the above condition is that the amount of mass in B is the same as the amount of mass that was transported to B under the transport map T. If μ and ν are absolutely continuous w.r.t. Lebesgue, with densities f and

g, respectively, and $T \in C^1(\mathbb{R}^d;\mathbb{R}^d)$ is injective, then using the change of variables $x = T(y)$, the equality

$$\int_{\mathbb{R}^d} \xi(T(y))f(y)\,\mathrm{d}y = \int_{\mathbb{R}^d} \xi\,(x)\,g(x)\,\mathrm{d}x$$

is equivalent to

$$f(y) = g\,(T(y))\,|\det \nabla T(y)|\,.$$

Let $\mathcal{P}^2(\mathbb{R}^d)$ be the collection of probability measures on $\mathbb{R}^d$ with finite second moments; i.e. if $\mu \in \mathcal{P}^2(\mathbb{R}^d)$, then $\mu[\mathbb{R}^d] = 1$ and $\int_{\mathbb{R}^d} |x|^2\,\mathrm{d}\mu < \infty$. We can define a metric on this space, the 2-Wasserstein metric. A proof of the following can be found in [28].

Proposition 4.2.2. *(2-Wasserstein metric) [28]. Let $\mu, \nu \in \mathcal{P}^2(\mathbb{R}^d)$. Then the function $W_2 : \mathcal{P}^2(\mathbb{R}^d) \times \mathcal{P}^2(\mathbb{R}^d) \to [0,\infty)$*

$$W_2(\mu,\nu) := \left[\inf_{\gamma}\left\{\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\,\mathrm{d}\gamma(x,y)\ :\ \gamma \in \Gamma(\mu,\nu)\right\}\right]^{1/2} \tag{4.6}$$

defines a metric on $\mathcal{P}^2(\mathbb{R}^d)$. Here $\Gamma(\mu,\nu)$ is the set of all probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν. This means that

$$\gamma \in \Gamma(\mu,\nu) \Longleftrightarrow \begin{cases} \gamma[A \times \mathbb{R}^d] = \mu[A] \\ \gamma[\mathbb{R}^d \times B] = \nu[B] \end{cases}$$

for all measurable $A, B \subset \mathbb{R}^d$. Equivalently, $\gamma \in \Gamma(\mu,\nu)$ if and only if

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} [\varphi(x) + \psi(y)]\ \mathrm{d}\gamma(x,y) = \int_{\mathbb{R}^d} \varphi(x)\,\mathrm{d}\mu + \int_{\mathbb{R}^d} \psi(y)\,\mathrm{d}\nu$$

for all $\varphi \in L^1(\,\mathrm{d}\mu)$ and $\psi \in L^1(\,\mathrm{d}\nu)$.

The 2-Wasserstein distance is closely connected to the theory of optimal transportation. The square of the 2-Wasserstein distance is the *Kantorovich optimal transportation problem* [28]

$$\textit{Minimize}\quad I[\gamma] := \int_{\mathbb{R}^d\times\mathbb{R}^d} c(x,y)\,\mathrm{d}\gamma(x,y) \quad \textit{for} \quad \gamma \in \Gamma(\mu,\nu).$$

when the cost function $c(x,y) = |x-y|^2$; an admissible γ is often referred to as a *transference plan*. It is a relaxed form of *Monge's optimal transport problem* [28]

$$\textit{Minimize} \quad I[T] := \int_{\mathbb{R}^d} c(x, T(x))\,\mathrm{d}\mu(x) \quad \textit{over all} \quad T\#\mu = \nu$$

where T is said to be an *optimal transport map*.

A great deal of theory, especially for the quadratic cost function, has been developed surrounding the question of when an optimal transference plan γ gives rise to a transport map T, i.e. when a minimizer for Kantorovich is actually a minimizer for Monge, $\gamma = (Id \times T)\#\mu$. From [28] we extract the following celebrated Brenier's theorem providing an answer for the quadratic case.

Theorem 4.2.3. *(Brenier's Theorem) [28]. Let $\mu, \nu \in \mathcal{P}^2(\mathbb{R}^d)$. If μ is absolutely continuous with respect to Lebesgue, then there is a unique optimal γ for $W_2(\mu,\nu)^2$, which is given by*

$$\mathrm{d}\gamma(x,y) = \mathrm{d}\mu(x)\delta\left[y = \nabla\varphi(x)\right],$$

where $\nabla\varphi$ is the unique gradient of a convex function which pushes μ onto ν, and δ is the Dirac measure.

In particular, if μ has density f, and $\nu \in \mathcal{P}^2(\mathbb{R}^d)$, then there exists $T = \nabla\varphi$ pushing μ to ν where φ is convex, and

$$W_2(\mu,\nu)^2 = \int_{\mathbb{R}^d} |x - \nabla\varphi(x)|^2 f(x)\,\mathrm{d}x.$$

4.3 Transport as Steepest Descent of the Potential Energy

In [19], Jordan, Kinderlehrer, and Otto identified the Fokker-Planck equation $\partial_t\rho = \Delta\rho + \operatorname{div}(\rho\nabla\Psi)$ as a gradient flow of the free energy $F(\rho) = \int_{\mathbb{R}^d} \rho\log\rho + \rho\Psi\,\mathrm{d}x$ in the 2-Wasserstein distance. More precisely, they proved that the time discrete scheme

Given $\rho_\tau^{n-1} \in \mathcal{P}_a^2(\mathbb{R}^d)$ with $F(\rho_\tau^{n-1}) < \infty$, find the minimizer ρ_τ^n of the functional

$$\rho \mapsto \frac{1}{2\tau} W_2(\rho_\tau^{n-1}, \rho)^2 + F(\rho) \tag{4.7}$$

over all $\rho \in \mathcal{P}_a^2(\mathbb{R}^d)$

converges for each $t \in (0, \infty)$ in the weak L^1 topology on $\mathbb{R}^d$ (after the time interpolation $\rho_\tau(t) = \rho_\tau^n$, $t \in [n\tau, (n+1)\tau)$), as the time step $\tau \downarrow 0$, to a solution ρ of the Fokker-Planck equation.

We plan to run the same argument for the transport equation. The above variational problem should therefore be simplified to

Given $\rho_\tau^{n-1} \in \mathcal{P}_a^2(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \rho_\tau^{n-1} \Psi \, \mathrm{d}x < \infty$, find the minimizer ρ_τ^n of the functional

$$\rho \mapsto I_{\rho_\tau^{n-1}}[\rho] := \frac{1}{2\tau} W_2(\rho_\tau^{n-1}, \rho)^2 + \int_{\mathbb{R}^d} \rho \Psi \, \mathrm{d}x \tag{4.8}$$

over all $\rho \in \mathcal{P}_a^2(\mathbb{R}^d)$.

A first step is to establish the existence of a minimizer to (4.8). Although the above functional is quite simple, we cannot deduce the existence of a minimizer to (4.8) in the same way as [19] did for (4.7), because while $\rho \mapsto \rho \log \rho + \rho\Psi$ is superlinear, $\rho \mapsto \rho\Psi$ is not. In particular, [19] obtains (relative) compactness of a minimizing sequence $\{\rho_\nu\}$ in the weak L^1 topology on $\mathbb{R}^d$ by proving that $\int_{\mathbb{R}^d} F(\rho_\nu) \, \mathrm{d}x \leq C$ and $\int_{\mathbb{R}^d} |x|^2 \rho_\nu \, \mathrm{d}x \leq C$, where $F(x) = x \log x$ is a superlinear function. This is enough to conclude tightness and uniform integrability of the sequence (see [8, 22]).

We do not have any 'superlinear bound' here. We only have a second moment bound, which is enough to ensure tightness of the minimizing sequence, and apply Prokhorov's theorem to establish that there exists an optimal measure. From there, a little more work will show that the measure admits a Lebesgue density. This general technique has been applied in, e.g. [1], from which we adapt to our situation. For an alternative method of establishing existence of a minimizer, we refer to [21]. We first review the relevant concepts.

Definition 4.3.1. *(Tightness) Let $\{\mu_n\}$ be a collection of probability measures on $\mathbb{R}^d$. Then $\{\mu_n\}$ is tight if, for all $\epsilon > 0$, there exists a compact $K_\epsilon \subset \mathbb{R}^d$ such that*

$$\mu_n \left(\mathbb{R}^d \backslash K_\epsilon\right) < \epsilon, \quad \textit{for all } n,$$

(equivalently, $\mu_n (K_\epsilon) > 1 - \epsilon$). That is, 'no mass escapes to infinity'.

Lemma 4.3.2. *(Second Moment Bound Implies Tightness) Suppose $\{\mu_n\}$ is a collection of probability measures on $\mathbb{R}^d$ satisfying*

$$\int_{\mathbb{R}^d} |x|^2 d\mu_n(x) \leq C, \quad \textit{for all } n.$$

Then $\{\mu_n\}$ is tight.

Proof. Let $\epsilon > 0$, and set $K_\epsilon := \{x \in \mathbb{R}^d : |x|^2 \leq 1/\epsilon\}$. Then

$$\int_{\mathbb{R}^d \setminus K_\epsilon} d\mu_n(x) = \int_{\{|x|^2 > 1/\epsilon\}} d\mu_n(x) \leq \int_{\{|x|^2 > 1/\epsilon\}} \epsilon |x|^2 \,\mathrm{d}\mu_n(x) \leq C\epsilon.$$

□

Definition 4.3.3. *(Weak Convergence of Probability Measures) Let $\{\mu_n\}$ be a collection of probability measures on $\mathbb{R}^d$. Then $\{\mu_n\}$ weakly converges to a probability measure μ on $\mathbb{R}^d$ if*

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f \,\mathrm{d}\mu_n = \int_{\mathbb{R}^d} f \,\mathrm{d}\mu$$

for all real-valued continuous bounded functions f on $\mathbb{R}^d$.

Proposition 4.3.4. *(Portmanteau) [6] $\{\mu_n\}$ weakly converges to a probability measure μ on $\mathbb{R}^d$ if and only if*

$$\int_{\mathbb{R}^d} f \,\mathrm{d}\mu \leq \liminf \int_{\mathbb{R}^d} f \,\mathrm{d}\mu_n$$

for every real-valued lower semi-continuous function f on $\mathbb{R}^d$ bounded from below.

Theorem 4.3.5. *(Prokhorov's theorem) [6] Let $\{\mu_n\}$ be a collection of probability measures on $\mathbb{R}^d$. Then $\{\mu_n\}$ is tight if and only if there exists a subsequence of $\{\mu_n\}$ which weakly converges in the space of probability measures on $\mathbb{R}^d$.*

With the above results in hand, we can now turn to the problem (4.8). We establish the result when $n = 1$ in (4.8).

Proposition 4.3.6. *The variational problem (4.8) admits a unique minimizer $\rho^1 \in \mathcal{P}_a^2(\mathbb{R}^d)$ for τ sufficiently small. In addition, if $T\#\rho^0 = \rho^1$ is the optimal map for $W_2(\rho^0, \rho^1)^2$, then T satisfies the equation*

$$\frac{T(x) - x}{\tau} = -\nabla\Psi(T(x)), \quad x \in \mathbb{R}^d, \tag{4.9}$$

and its inverse $T^{-1}\#\rho^1 = \rho^0$ is explicitly given by

$$T^{-1}(y) = y + \tau\nabla\Psi(y). \tag{4.10}$$

In particular, ρ^1 is explicitly given by

$$\rho^1(x) = \rho^0\left(T^{-1}(x)\right)\det\nabla\left(T^{-1}\right)(x). \tag{4.11}$$

Moreover,

$$\left|\int_{\mathbb{R}^d}\frac{\rho^1-\rho^0}{\tau}\xi\,\mathrm{d}x + \int_{\mathbb{R}^d}\rho^1\nabla\Psi\cdot\nabla\xi\,\mathrm{d}x\right| \leq \frac{1}{2\tau}\left\|D^2\xi\right\|_{L^\infty(\mathbb{R}^d)} W_2(\rho^0,\rho^1)^2, \tag{4.12}$$

for every $\xi \in C_c^\infty(\mathbb{R}^d)$.

Proof. We first show that (4.8) admits a minimizer. The argument is well-known (see e.g. [19]) however we detail it here for convenience. Since $0 \leq I_{\rho^0}[\rho]$ for all admissible ρ and $I_{\rho^0}[\rho^0] = \int_{\mathbb{R}^d}\rho^0\Psi\,\mathrm{d}x < \infty$, then the infimum (4.8) is finite. Let $\{\rho_\nu\}$ be a minimizing sequence. Then

$$W_2(\rho^0,\rho_\nu)^2 \leq 2\tau I_{\rho^0}[\rho_\nu] \leq 2\tau\int_{\mathbb{R}^d}\rho^0\Psi\,\mathrm{d}x$$

is uniformly bounded in ν. Since $|x|^2 \leq 2|x-y|^2 + 2|y|^2$ for all $x, y \in \mathbb{R}^d$, we have

$$\int_{\mathbb{R}^d}|x|^2\rho_\nu\,\mathrm{d}x \leq 2W_2(\rho^0,\rho_\nu)^2 + 2\int_{\mathbb{R}^d}|y|^2\rho^0\,\mathrm{d}y \leq 4\tau\int_{\mathbb{R}^d}\rho^0\Psi\,\mathrm{d}x + 2\int_{\mathbb{R}^d}|y|^2\rho^0\,\mathrm{d}y.$$

Therefore $\{\rho_\nu\,\mathrm{d}x\}$ is tight, and hence there exists a probability measure μ_1 on $\mathbb{R}^d$ such that $\{\rho_\nu\,\mathrm{d}x\}$ converges weakly to μ_1. By Proposition 4.3.4, $\int_{\mathbb{R}^d}\Psi\,\mathrm{d}\mu_1 \leq \liminf_\nu\int_{\mathbb{R}^d}\Psi\rho_\nu\,\mathrm{d}x$. Moreover, (see [19]), $W_2(\rho^0,\mu_1)^2 \leq \liminf_\nu W_2(\rho^0,\rho_\nu)^2$ (in particular, this implies $\mu_1 \in \mathcal{P}^2(\mathbb{R}^d)$). Therefore μ_1 is a minimizer for (4.8).

For uniqueness, we have that $\mu \mapsto W_2(\rho^0,\mu)^2$ is strictly convex over the admissible set $\mu \in \mathcal{P}^2(\mathbb{R}^d)$ [19]. This is because if μ, β are admissible, and $\lambda \in (0,1)$, then (applying Brenier's theorem (Theorem 4.2.3) since $\rho^0 \in \mathcal{P}_a^2(\mathbb{R}^d)$) letting $\nabla\varphi_\mu$ and $\nabla\varphi_\beta$ be the optimal map for $W_2(\rho^0,\mu)^2$ and $W_2(\rho^0,\beta)^2$, respectively, we have $\lambda\nabla\varphi_\mu +$

$(1-\lambda)\nabla\varphi_\beta$ is optimal for $W_2(\rho^0, \lambda\mu + (1-\lambda)\beta)^2$, so by definition

$$\begin{aligned} W_2(\rho^0, \lambda\mu + (1-\lambda)\beta)^2 &= \int_{\mathbb{R}^d} |x - \lambda\nabla\varphi_\mu - (1-\lambda)\nabla\varphi_\beta|^2 \rho^0 \,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} |\lambda(x - \nabla\varphi_\mu) + (1-\lambda)(x - \nabla\varphi_\beta)|^2 \rho^0 \,\mathrm{d}x \\ &\leq \lambda \int_{\mathbb{R}^d} |x - \nabla\varphi_\mu|^2 \rho^0 \,\mathrm{d}x + (1-\lambda)\int_{\mathbb{R}^d} |x - \nabla\varphi_\beta|^2 \rho^0 \,\mathrm{d}x \\ &= \lambda W_2(\rho^0, \mu)^2 + (1-\lambda) W_2(\rho^0, \beta)^2, \end{aligned}$$

with equality if and only if $\lambda = 0, 1$, by strict convexity of $x \mapsto |x|^2$. Since additionally $\mu \mapsto \int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu$ is linear, $\mu \mapsto \frac{1}{2\tau} W_2(\rho^0, \mu)^2 + \int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu$ is strictly convex, and hence (4.8) admits at most one minimizer.

Let us now derive the Euler-Lagrange equation for μ_1. We follow the technique in [19] while also drawing from [1]. Fix some smooth vector field $\xi \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$, and for $\epsilon \in \mathbb{R}$ let $\epsilon \mapsto \alpha_\epsilon \in \mathbb{R}^d$ be the flow solving

$$\begin{cases} \partial_\epsilon \alpha_\epsilon = \xi(\alpha_\epsilon) \\ \alpha_0 = Id. \end{cases} \tag{4.13}$$

We fix a variation $\mu_\epsilon := \alpha_\epsilon \# \mu_1$. Then

$$\frac{1}{\epsilon}\left[\frac{1}{2\tau} W_2(\rho^0, \mu_\epsilon)^2 + \int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_\epsilon - \frac{1}{2\tau} W_2(\rho^0, \mu_1)^2 - \int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_1\right] \geq 0,$$

for all $\epsilon \in \mathbb{R}$. Hence

$$\begin{aligned} &\frac{1}{2\tau} \limsup_{\epsilon \to 0} \left[\frac{W_2(\rho^0, \mu_\epsilon)^2 - W_2(\rho^0, \mu_1)^2}{\epsilon}\right] + \limsup_{\epsilon \to 0} \left[\frac{\int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_\epsilon - \int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_1}{\epsilon}\right] \\ &\geq \limsup_{\epsilon \to 0} \frac{1}{\epsilon}\left[\frac{1}{2\tau} W_2(\rho^0, \mu_\epsilon)^2 + \int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_\epsilon - \frac{1}{2\tau} W_2(\rho^0, \mu_1)^2 - \int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_1\right] \geq 0, \end{aligned}$$

and we will investigate each limit separately.

Since

$$\frac{\int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_\epsilon - \int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_1}{\epsilon} = \int_{\mathbb{R}^d} \frac{\Psi(\alpha_\epsilon) - \Psi}{\epsilon} \,\mathrm{d}\mu_1,$$

and $\Psi \in C^1(\mathbb{R}^d), \xi \in C_c^\infty(\mathbb{R}^d)$, the estimate

$$\left|\frac{\Psi(\alpha_\epsilon) - \Psi}{\epsilon}\right| \leq \|\nabla\Psi \cdot \xi\|_{L^\infty(\mathbb{R}^d)},$$

holds by a first-order Taylor expansion. Since furthermore $\|\nabla\Psi \cdot \xi\|_{L^\infty(\mathbb{R}^d)}$ is integrable w.r.t. μ_1 (since $\int_{\mathbb{R}^d} \mathrm{d}\mu_1 = 1$), we can apply the dominated convergence theorem [16] to conclude

$$\limsup_{\epsilon\to 0} \frac{\int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_\epsilon - \int_{\mathbb{R}^d} \Psi \,\mathrm{d}\mu_1}{\epsilon} = \int_{\mathbb{R}^d} \nabla\Psi \cdot \xi \,\mathrm{d}\mu_1.$$

From [19] we retrieve the estimate

$$\limsup_{\epsilon\to 0} \frac{W_2(\rho^0,\mu_\epsilon)^2 - W_2(\rho^0,\mu_1)^2}{2\epsilon} \le \int_{\mathbb{R}^d\times\mathbb{R}^d} (y-x)\cdot\xi(y)\,\mathrm{d}\gamma_1(x,y),$$

where γ_1 is the optimal probability measure for $W_2(\rho^0,\mu_1)^2$ on $\mathbb{R}^d\times\mathbb{R}^d$. Therefore we have

$$\frac{1}{\tau}\int_{\mathbb{R}^d\times\mathbb{R}^d} (y-x)\cdot\xi(y)\,\mathrm{d}\gamma_1(x,y) + \int_{\mathbb{R}^d} \nabla\Psi(y)\cdot\xi(y)\,\mathrm{d}\mu_1(y) \ge 0,$$

and interchanging $\xi \to -\xi$, we get equality

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \left[\frac{y-x}{\tau} + \nabla\Psi(y)\right]\cdot\xi(y)\,\mathrm{d}\gamma_1(x,y) = 0.$$

In fact, we can say more about the optimal γ_1. By Brenier's Theorem, we have since $\rho^0 \in \mathcal{P}_a^2(\mathbb{R}^d)$,

$$\mathrm{d}\gamma_1(x,y) = \rho^0(x)\,\mathrm{d}x\,\delta\left[y = \nabla\varphi(x)\right],$$

for a unique $\nabla\varphi\#\rho^0 = \mu_1$ where φ is convex. Therefore

$$\int_{\mathbb{R}^d} \left[\frac{\nabla\varphi(x)-x}{\tau} + \nabla\Psi\left(\nabla\varphi(x)\right)\right]\cdot\xi(\nabla\varphi(x))\rho^0(x)\,\mathrm{d}x = 0. \tag{4.14}$$

Now, define $\varphi^* : \mathbb{R}^d \to \mathbb{R}$ by $\varphi^*(y) = \frac{|y|^2}{2} + \tau\Psi(y)$. Then for small enough $\tau > 0$, φ^* is superlinear and strictly convex, since $D^2\varphi^* = I + \tau D^2\Psi$, and $\Psi \in C^{1,1}(\mathbb{R}^d)$.

It therefore follows [10] that its Legendre transform, which we denote by φ^{**}, is finite everywhere, C^1, and is convex also, where

$$\varphi^{**}(x) := \sup_{y\in\mathbb{R}^d} \left\{x\cdot y - \varphi^*(y)\right\}.$$

Moreover, [28] $\nabla\varphi^{**} = \left(\nabla\varphi^*\right)^{-1}$. Since $\nabla\varphi^*(x) = x + \tau\nabla\Psi(x)$, it follows that $\nabla\varphi^{**}$

satisfies

$$y = \nabla\varphi^{**}(y) + \tau\nabla\Psi\left(\nabla\varphi^{**}(y)\right), \quad y \in \mathbb{R}^d,$$

i.e., $\nabla\varphi^{**}$ satisfies (4.14). By uniqueness, we must therefore have $\nabla\varphi = \nabla\varphi^{**}$, Lebesgue almost everywhere. This proves (4.9) and (4.10) where $T := \nabla\varphi$.

We still need to show μ_1 is absolutely continuous w.r.t. Lebesgue. To this end, we will use the facts that $\nabla\varphi\#\rho^0 = \mu_1$, and ρ^0 is absolutely continuous w.r.t. Lebesgue. Let $B \subset \mathbb{R}^d$ be a subset with Lebesgue measure 0. Since $\mu_1[B] = \rho^0[\nabla\varphi^*(B)]$ it suffices to show the Lebesgue measure of $\nabla\varphi^*(B)$ is zero.

$$\begin{aligned}\mathcal{L}(\nabla\varphi^*(B)) &= \int_{\mathbb{R}^d} \chi_{\nabla\varphi^*(B)}(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^d} \chi_B(\nabla\varphi(x))\,\mathrm{d}x \\ &= \int_B \left|\det D^2\varphi^*(x)\right|\,\mathrm{d}x\end{aligned}$$

where χ is the indicator function. This implies that $\mathcal{L}(\nabla\varphi^*(B)) \le C\mathcal{L}(B) = 0$, and hence μ_1 is absolutely continuous w.r.t. Lebesgue; we denote its density by ρ^1. Thus we obtain (4.11) from the fact that $T^{-1} \in C^1(\mathbb{R}^d;\mathbb{R}^d)$ and $T^{-1}\#\rho^1 = \rho^0$.

Finally we show that ρ^1 solves an approximate weak form of the transport equation (4.12). The following can be found in [21], however, we provide the details here for convenience. From (4.14), we choose $\xi = \nabla\zeta$ for $\zeta \in C_c^\infty(\mathbb{R}^d)$, and use $T\#\rho^0 = \rho^1$, $T^{-1}\#\rho^1 = \rho^0$, to write

$$\int_{\mathbb{R}^d} \rho^1 \frac{T^{-1}(x) - x}{\tau} \cdot \nabla\zeta\,(x)\;\mathrm{d}x - \int_{\mathbb{R}^d} \rho^1 \nabla\Psi\,(x) \cdot \nabla\zeta(x)\,\mathrm{d}x = 0. \tag{4.15}$$

With the Taylor expansion

$$\zeta(T^{-1}(x)) - \zeta(x) = \nabla\zeta(x)\cdot(T^{-1}(x) - x) + \frac{1}{2}(T^{-1}(x) - x)^t D^2\zeta(\lambda_{x,\tau})(T^{-1}(x) - x)$$

where t is the transpose, and $\lambda_{x,\tau}$ is an intermediate point between x and $T(x)$, we substitute into (4.15), and obtain

$$\begin{aligned}\left|\int_{\mathbb{R}^d} \frac{\rho^1 - \rho^0}{\tau}\zeta\,\mathrm{d}x + \int_{\mathbb{R}^d} \rho^1\nabla\Psi\cdot\nabla\zeta\,\mathrm{d}x\right| &\le \frac{1}{2\tau}\left\|D^2\zeta\right\|_{L^\infty(\mathbb{R}^d)} \int_{\mathbb{R}^d} |T^{-1}(x) - x|^2 \rho^1(x)\,\mathrm{d}x \\ &\le \frac{1}{2\tau}\left\|D^2\zeta\right\|_{L^\infty(\mathbb{R}^d)} W_2(\rho^0,\rho^1)^2.\end{aligned}$$

□

To conclude the identification of (4.1) as a gradient flow of the potential energy with respect to the 2-Wasserstein distance on $\mathcal{P}_a^2(\mathbb{R}^d)$, we still need to define a function ρ_τ defined for all $t \in (0, T)$, obtain compactness of the sequence $\{\rho_\tau\}_{\tau\downarrow 0}$ in some topology to deduce the existence of a candidate solution ρ, and then show this ρ is indeed a weak solution of the transport equation. We omit the details here as we will be following similar steps in the sequel when combining both fractional heat and transport together. We refer the reader to [21] for further discussion regarding these steps for the transport equation.

4.4 The Characteristic Equation

We conclude discussion on the transport equation with the following. Consider

$$\begin{cases} \dfrac{\partial T}{\partial t}(x,t) = -\nabla \Psi \left(T(x,t)\right), \\ T(x,0) = x. \end{cases} \tag{4.16}$$

The existence of a unique C^1 map T to (4.16) is assured through a fixed-point argument, because $\Psi \in C^{1,1} \cap C^{2,1}(\mathbb{R}^d)$ [16]. We then have the following from [21].

Proposition 4.4.1. *Let $T(t,\cdot)$ be the unique C^1 map solving (4.16). Then $v(t) := T(t,\cdot)\#v^0$ is a weak solution of (4.1).*

Proof. To avoid the notational difficulty here of denoting the time as T and the map T solving (4.16), we let $\tilde{T} < \infty$ denote the time. Let $\varphi \in C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ with time support in $[-\tilde{T}, \tilde{T}]$. Then

$$\int_{\mathbb{R}^d} \int_0^{\tilde{T}} \varphi(x,t) v(x,t) = \int_{\mathbb{R}^d} \int_0^{\tilde{T}} \varphi\left(T(x,t), t\right) v^0(x) \, \mathrm{d}x$$

so that

$$\begin{aligned}
\int_{\mathbb{R}^d}\int_0^{\tilde{T}} \partial_t\varphi(x,t)v(x,t)\,\mathrm{d}t\,\mathrm{d}x &= \int_{\mathbb{R}^d}\int_0^{\tilde{T}} \left(\partial_t\varphi\right)\left(T(x,t),t\right) v^0(x)\,\mathrm{d}t\,\mathrm{d}x \\
&= \int_{\mathbb{R}^d}\int_0^{\tilde{T}} \frac{\mathrm{d}}{\mathrm{d}t}\left[\varphi(T(x,t),t)\right] v^0(x)\,\mathrm{d}t\,\mathrm{d}x \\
&\quad + \int_{\mathbb{R}^d}\int_0^{\tilde{T}} \nabla\varphi\left(T(x,t),t\right)\cdot\nabla\Psi\left(T(x,t)\right) v^0(x)\,\mathrm{d}t\,\mathrm{d}x \\
&= \int_{\mathbb{R}^d}\int_0^{\tilde{T}} \nabla\varphi(x)\cdot\nabla\Psi(x)v(x,t)\,\mathrm{d}t\,\mathrm{d}x \\
&\quad - \int_{\mathbb{R}^d} \varphi(x,0)v^0(x)\,\mathrm{d}x.
\end{aligned}$$

Mass preservation follows by definition of the push-forward. $\square$

Applying an implicit Euler scheme to (4.16) yields a T_τ satisfying $\frac{T_\tau(x)-x}{\tau} = -\nabla\Psi(T_\tau(x))$, which is exactly the equation satisfied by the optimal transport map earlier. Thus, in this case, the gradient flow interpretation is equivalent to an implicit Euler scheme for the characteristic equation [21].

Chapter 5

Operator Splitting on the Fractional Fokker-Planck Equation

5.1 Construction

We return now to the central question of operator splitting. Recall that in the splitting scheme for obtaining a discrete time approximation to (1.3) up to some finite time horizon $T < \infty$, we fix a time step $\tau > 0$ so that $N\tau = T$ for some $N \in \mathbb{N}$, set $\rho_\tau^0 = \rho^0$, and then recursively iterate the following two connected subproblems for $n = 0, 1, \ldots, N-1$:

1. *(The fractional heat equation) Solve*

$$\begin{aligned} &\partial_t u(x,t) = -(-\Delta)^s u(x,t), \quad (x,t) \in \mathbb{R}^d \times (0,\infty) \\ &u(x,0) = \rho_\tau^n(x) \end{aligned}$$

   and set $\tilde{\rho}_\tau^{n+1}(x) := u(x,\tau)$.

2. *(The transport equation as a gradient flow)*
   Minimize

$$\rho \mapsto I_{\tilde{\rho}_\tau^{n+1}}[\rho] := \frac{1}{2\tau} W_2(\tilde{\rho}_\tau^{n+1}, \rho)^2 + \int_{\mathbb{R}^d} \rho \Psi \, \mathrm{d}x \tag{5.1}$$

   over all $\rho \in \mathcal{P}_a^2(\mathbb{R}^d)$, *and set* $\rho_\tau^{n+1}(x)$ *as the minimizer.*

We already have a good understanding of the properties of the solution to each of these problems. However, as noted in Proposition 3.1.1, we have to deal with initial data $\tilde{\rho}_\tau^{n+1}$ in the variational problem (5.1) which may not possess a finite second moment. In particular, this implies that the potential energy of the initial data may be infinite (e.g. if $\Psi(x) = |x|^2/2$), which violates the starting estimate $I_{\tilde{\rho}_\tau^{n+1}}[\tilde{\rho}_\tau^{n+1}] = \int_{\mathbb{R}^d} \tilde{\rho}_\tau^{n+1} \Psi \, \mathrm{d}x < \infty$ to show existence of a minimizer.

Therefore, a possible workaround to place the problem back in the framework of Chapter 4 is to find some approximation of the initial data $\tilde{\rho}_\tau^{n+1}$. Ideally, we want the error in the approximation to be uniform with respect to n in, e.g. the L^1 norm (see below in (5.31)). To achieve this, we propose the following.

Introduce an additional parameter $R > 0$, and starting with $n = 0$, replace $\tilde{\rho}_\tau^1 = \Phi_s^\tau * \rho^0$ in (5.1) with the normalized 'approximation'

$$\bar{\rho}_{\tau,R}^1 := \frac{(\Phi_s^\tau 1_{B_R}) * \rho^0}{\|\Phi_s^\tau 1_{B_R}\|_{L^1(\mathbb{R}^d)}}, \tag{5.2}$$

where $\Phi_s^\tau(x) := \Phi_s(x, \tau)$. We continue this procedure for every n. Note that the original initial data $\tilde{\rho}_\tau^1$ in (5.1) satisfies $\tilde{\rho}_\tau^1 = \lim_{R\to\infty} \bar{\rho}_{\tau,R}^1$.

It is clear that now $\bar{\rho}_{\tau,R}^1 \in \mathcal{P}_a^2(\mathbb{R}^d)$ whenever $\rho^0 \in \mathcal{P}_a^2(\mathbb{R}^d)$, and we also see below that $\int_{\mathbb{R}^d} \bar{\rho}_{\tau,R}^1 \Psi < \infty$ whenever $\int_{\mathbb{R}^d} \rho^0 \Psi < \infty$. Then with the modified initial condition (5.2) we can replace (5.1) with its approximated version

Minimize

$$\rho \mapsto I_{\bar{\rho}_{\tau,R}^{n+1}}[\rho] := \frac{1}{2\tau} W_2(\bar{\rho}_{\tau,R}^{n+1}, \rho)^2 + \int_{\mathbb{R}^d} \rho \Psi \, \mathrm{d}x \tag{5.3}$$

over all $\rho \in \mathcal{P}_a^2(\mathbb{R}^d)$, and set $\rho_{\tau,R}^{n+1}(x)$ as the minimizer.

We then know (5.3) admits a unique minimizer $\rho_{\tau,R}^1$ which has the explicit expression

$$\rho_{\tau,R}^1(x) = \bar{\rho}_{\tau,R}^1 \circ T^{-1}(x) \det(\nabla T^{-1}(x)), \quad \text{where } T^{-1}(x) = x + \tau \nabla \Psi(x). \tag{5.4}$$

We make this precise in the following.

Lemma 5.1.1. *Let $\rho^0 \in \mathcal{P}_a^2(\mathbb{R}^d)$ and $\Psi \in C^{1,1} \cap C^2(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \rho^0 \Psi \, \mathrm{d}x < \infty$. Define $\bar{\rho}_{\tau,R}^1$ by (5.2). Then*

1. $\bar{\rho}_{\tau,R}^1 \in \mathcal{P}_a^2(\mathbb{R}^d)$;

2. $$\int_{\mathbb{R}^d} \bar{\rho}_{\tau,R}^1 \Psi \, \mathrm{d}x \le \int_{\mathbb{R}^d} \rho_{\tau,R}^0 \Psi \, \mathrm{d}x + \frac{\|D^2\Psi\|_{L^\infty(\mathbb{R}^d)}}{2} \frac{\int_{B_R} |x|^2 \Phi_s^\tau(x) \, \mathrm{d}x}{\int_{B_R} \Phi_s^\tau \, \mathrm{d}x} < \infty; \tag{5.5}$$

3. *there exists a unique minimizer* $\rho^1_{\tau,R}$ *of (5.3) which is explicitly given by (5.4).*

Proof. 1. It is immediate that $\bar\rho^1_{\tau,R} \geq 0$ and $\int_{\mathbb{R}^d} \bar\rho^1_{\tau,R} = 1$. The last property to check is $\int_{\mathbb{R}^d} |x|^2 \bar\rho^1_{\tau,R}(x)\,\mathrm{d}x \leq 2\int_{\mathbb{R}^d} |x|^2 \rho^0(x)\,\mathrm{d}x + 2\frac{\int_{B_R} |y|^2 \Phi^\tau_s(y)\,\mathrm{d}y}{\int_{B_R} \Phi^\tau_s(y)\,\mathrm{d}y} < \infty$.

2. By definition,

$$\int_{\mathbb{R}^d} \bar\rho^1_{\tau,R} \Psi \,\mathrm{d}x = \frac{\int_{\mathbb{R}^d} \Psi(x) \int_{B_R} \rho^0_{\tau,R}(x-y)\Phi^\tau_s(y)\,\mathrm{d}y\,\mathrm{d}x}{\int_{B_R} \Phi^\tau_s(y)\,\mathrm{d}y}. \tag{5.6}$$

Focusing on the numerator, we have by a change of variable $z = x - y$,

$$\int_{\mathbb{R}^d} \Psi(x) \int_{B_R} \rho^0_{\tau,R}(x-y)\Phi^\tau_s(y)\,\mathrm{d}y\,\mathrm{d}x = \int_{B_R} \Phi^\tau_s(y) \int_{\mathbb{R}^d} \Psi(y+z)\rho^0_{\tau,R}(z)\,\mathrm{d}z\,\mathrm{d}y.$$

Now a finite Taylor expansion for Ψ gives

$$\Psi(y+z) = \Psi(z) + y\cdot\nabla\Psi(z) + \frac{1}{2} y^t D^2\Psi(\xi_{y,z})y, \tag{5.7}$$

where $\xi_{y,z}$ is some intermediate point on the line joining y and z. Since the Hessian of Ψ is bounded we can then estimate

$$\int_{B_R} \Phi^\tau_s(y) \int_{\mathbb{R}^d} \Psi(y+z)\rho^0_{\tau,R}(z)\,\mathrm{d}z\,\mathrm{d}y \tag{5.8}$$
$$\leq \int_{B_R} \Phi^\tau_s(y)\,\mathrm{d}y \int_{\mathbb{R}^d} \Psi(z)\rho^0_{\tau,R}(z)\,\mathrm{d}z \tag{5.9}$$
$$+ \left| \int_{B_R} \Phi^\tau_s(y) \int_{\mathbb{R}^d} y\cdot\nabla\Psi(z)\rho^0_{\tau,R}(z)\,\mathrm{d}z\,\mathrm{d}y \right| \tag{5.10}$$
$$+ \frac{\|D^2\Psi\|_{L^\infty(\mathbb{R}^d)}}{2} \int_{B_R} |y|^2 \Phi^\tau_s(y) \underbrace{\int_{\mathbb{R}^d} \rho^0_{\tau,R}(z)\,\mathrm{d}z}_{1}\,\mathrm{d}y \tag{5.11}$$

and use the fact that

$$\int_{B_R} y\Phi^\tau_s(y)\,\mathrm{d}y = 0$$

to conclude (5.10) vanishes. Thus by (5.6), (5.9), and (5.11), we obtain the desired conclusion.

3. We have just established that $\bar\rho^1_{\tau,R}$ is admissible, and $\int_{\mathbb{R}^d} \bar\rho^1_{\tau,R}\Psi\,\mathrm{d}x < \infty$. Then we can appeal to Proposition 4.3.6 to conclude the desired result.

□

Remark 5.1.2. *The additional parameter R will later be set to a function of τ.*

The last few paragraphs have now motivated the 'splitting' scheme we will use, of which is now extremely simple since everything is known explicitly.

The Splitting Scheme

1. *Set*
$$\rho^0_{\tau,R} := \rho^0.$$
For $n = 0, \ldots, N-1$,

2. *Define*
$$\bar{\rho}^{n+1}_{\tau,R} := \frac{(\Phi^\tau_s 1_{B_R}) * \rho^n_{\tau,R}}{\|\Phi^\tau_s 1_{B_R}\|_{L^1(\mathbb{R}^d)}}. \tag{5.12}$$

3. *Set*

$$\rho^{n+1}_{\tau,R} = \bar{\rho}^{n+1}_{\tau,R} \circ T^{-1} \det(\nabla T^{-1}), \quad \textit{where } T^{-1}(x) = x + \tau\nabla\Psi(x). \tag{5.13}$$

5.2 Time-Dependent Approximation

Through the scheme outlined above, we obtain a discrete-time sequence $\{\rho^k_{\tau,R}\}_{0\le k\le N}$. The goal now is to choose some suitable time-interpolation of the sequence $\{\rho^k_{\tau,R}\}$, and prove that this time-interpolation converges (to be defined in a suitable sense) to a weak solution of the original PDE (1.3). We first define our notion of a weak solution.

Definition 5.2.1. *(Weak Solution) Let $T < \infty$. We say that $\rho = \rho(x,t) : \mathbb{R}^d \times [0,T) \to [0,\infty)$ is a weak solution to (1.3) if,*

1. *For every $\varphi \in C^\infty_c(\mathbb{R}^d \times \mathbb{R})$ with time support in $[-T,T]$,*

$$\int_0^T \int_{\mathbb{R}^d} \rho(x,t)\left[\partial_t\varphi(x,t) - (-\Delta)^s\varphi(x,t) - \nabla\Psi(x)\cdot\nabla\varphi(x,t)\right] \mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^d} \rho^0(x)\varphi(x,0)\,\mathrm{d}x = 0,$$

2. *$\rho(x,t) \ge 0$ for a.e. $(x,t) \in \mathbb{R}^d \times (0,T)$,*

3. $\int_{\mathbb{R}^d} \rho(x,t)\,\mathrm{d}x = \int_{\mathbb{R}^d} \rho^0\,\mathrm{d}x = 1$ *for a.e.* $t \in (0,T)$.

The non-negativity condition should be expected since ρ^0 is a probability density, while the mass-preserving condition is natural, since formally, if ρ is a solution, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} \rho(t)\,\mathrm{d}x = -\int_{\mathbb{R}^d} (-\Delta)^s \rho(t)\,\mathrm{d}x + \int_{\mathbb{R}^d} \operatorname{div}\left(\rho\nabla\Psi(x)\right)\,\mathrm{d}x = 0$$

because $\int_{\mathbb{R}^d} \operatorname{div}\left(\rho\nabla\Psi(x)\right)\,\mathrm{d}x = 0$, while

$$\int_{\mathbb{R}^d} (-\Delta)^s \rho(t)\,\mathrm{d}x = \mathcal{F}\left[(-\Delta)^s\rho(t)\right](\xi = 0) = \left(|\cdot|^{2s}\hat{\rho}(\cdot,t)\right)(\xi = 0) = 0.$$

Time Interpolation

To obtain a function defined for all $t \in [0,T)$, we set

$$\rho_{\tau,R}(t) := \Phi_s(t - t_n) * \rho^n_{\tau,R}, \quad t \in [t_n, t_{n+1}). \tag{5.14}$$

The reason for our choice of interpolation (5.14) is because on each interval $[t_n, t_{n+1})$, $\rho_{\tau,R}$ is a solution of the fractional heat equation with initial condition $\rho^n_{\tau,R}$.

The following lemma tells us how 'close' the interpolation $\rho_{\tau,R}$ is to being a weak solution of (1.3). Let us first briefly remind the reader of some notation: $\bar{\rho}^{n+1}_{\tau,R}$ is given by (5.12), $\rho^{n+1}_{\tau,R}$ is given by (5.13), and $\tilde{\rho}^{n+1}_{\tau,R} := \lim_{t\uparrow t_{n+1}} \rho_{\tau,R}(t)$.

Lemma 5.2.2. *(The Approximate Equation Satisfied by the Time Interpolation) Let* $\varphi \in C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ *with time support in* $[-T,T]$*. Then*

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} \rho_{\tau,R}(t)\left[\partial_t\varphi(t) - (-\Delta)^s\varphi(t) - \nabla\Psi\cdot\nabla\varphi(t)\right]\,\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^d} \rho^0\varphi(0)\,\mathrm{d}x \\ = \mathcal{R}(\tau,R)\end{aligned}$$

where $\mathcal{R}$ satisfies

$$\mathcal{R}(\tau, R) := \sum_{n=0}^{N-1} \int_{\mathbb{R}^d} \left(\tilde{\rho}_{\tau,R}^{n+1} - \bar{\rho}_{\tau,R}^{n+1}\right) \varphi(t_{n+1}) \, \mathrm{d}x \tag{5.15}$$

$$+ \sum_{n=0}^{N-1} \left[\int_{\mathbb{R}^d} \left(\bar{\rho}_{\tau,R}^{n+1} - \rho_{\tau,R}^{n+1}\right) \varphi(t_{n+1}) \, \mathrm{d}x \right.$$

$$\left. - \tau \int_{\mathbb{R}^d} \rho_{\tau,R}^{n+1} \nabla \Psi \cdot \nabla \varphi(t_{n+1}) \, \mathrm{d}x \right] \tag{5.16}$$

$$- \sum_{n=1}^{N-1} \left[\int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \left[\rho_{\tau,R}(t) \nabla \Psi \cdot \nabla \varphi(t) - \rho_{\tau,R}^{n} \nabla \Psi \cdot \nabla \varphi(t_n)\right] \, \mathrm{d}x \, \mathrm{d}t \right] \tag{5.17}$$

$$- \int_0^{\tau} \int_{\mathbb{R}^d} \Phi_s(t) * \rho^0 \nabla \Psi \cdot \nabla \varphi(t) \, \mathrm{d}x \, \mathrm{d}t. \tag{5.18}$$

Remark 5.2.3. *The errors above are characterized as follows:*

1. *The error (5.15) is from the approximation of the initial data $\tilde{\rho}_{\tau}^{n}$ by $\bar{\rho}_{\tau,R}^{n}$, as discussed above, so that we can run the variational method (5.3).*

2. *The error (5.16) comes from approximating solutions of the transport equation with minimizers of the variational problem (5.3) (or equivalently, is due to the fact that the variational problem (5.3) is essentially an implicit Euler scheme on the characteristic equation of the transport equation).*

3. *Finally, the last two terms (5.17) and (5.18) in $\mathcal{R}(\tau, R)$ are from the error in the splitting itself, because we want to replace*

$$\sum_{n=0}^{N-1} \tau \int_{\mathbb{R}^d} \rho_{\tau,R}^{n+1} \nabla \Psi \cdot \nabla \varphi(t_{n+1}) \, \mathrm{d}x$$

with

$$\int_0^T \int_{\mathbb{R}^d} \rho_{\tau,R}(t) \nabla \Psi \cdot \nabla \varphi(t) \, \mathrm{d}x \, \mathrm{d}t.$$

Proof. This essentially follows by an integration by parts, and addition and subtraction of terms. Integrating by parts we obtain

$$\int_{t_n}^{t_{n+1}}\int_{\mathbb{R}^d}\rho_{\tau,R}(t)\partial_t\varphi(t)\,\mathrm{d}x\,\mathrm{d}t=\int_{t_n}^{t_{n+1}}\int_{\mathbb{R}^d}\rho_{\tau,R}(t)(-\Delta)^s\varphi(t)\,\mathrm{d}x\,\mathrm{d}t$$
$$+\int_{\mathbb{R}^d}\tilde{\rho}_{\tau,R}^{n+1}\varphi(t_{n+1})-\rho_{\tau,R}^{n}\varphi(t_n)\,\mathrm{d}x. \tag{5.19}$$

Now write (5.19) as

$$\int_{\mathbb{R}^d}\tilde{\rho}_{\tau,R}^{n+1}\varphi(t_{n+1})-\rho_{\tau,R}^{n}\varphi(t_n)\,\mathrm{d}x=\int_{\mathbb{R}^d}\rho_{\tau,R}^{n+1}\varphi(t_{n+1})-\rho_{\tau,R}^{n}\varphi(t_n)\,\mathrm{d}x \tag{5.20}$$
$$+\int_{\mathbb{R}^d}\left(\tilde{\rho}_{\tau,R}^{n+1}-\rho_{\tau,R}^{n+1}\right)\varphi(t_{n+1})\,\mathrm{d}x, \tag{5.21}$$

and then (5.21) can be written as

$$\int_{\mathbb{R}^d}\left(\tilde{\rho}_{\tau,R}^{n+1}-\rho_{\tau,R}^{n+1}\right)\varphi(t_{n+1})\,\mathrm{d}x=\int_{\mathbb{R}^d}\left(\tilde{\rho}_{\tau,R}^{n+1}-\bar{\rho}_{\tau,R}^{n+1}\right)\varphi(t_{n+1})\,\mathrm{d}x \tag{5.22}$$
$$+\int_{\mathbb{R}^d}\left(\bar{\rho}_{\tau,R}^{n+1}-\rho_{\tau,R}^{n+1}\right)\varphi(t_{n+1})\,\mathrm{d}x. \tag{5.23}$$

Finally, (5.23) is

$$\int_{\mathbb{R}^d}\left(\bar{\rho}_{\tau,R}^{n+1}-\rho_{\tau,R}^{n+1}\right)\varphi(t_{n+1})\,\mathrm{d}x$$
$$=\int_{\mathbb{R}^d}\left(\bar{\rho}_{\tau,R}^{n+1}-\rho_{\tau,R}^{n+1}\right)\varphi(t_{n+1})\,\mathrm{d}x-\tau\int_{\mathbb{R}^d}\rho_{\tau,R}^{n+1}\nabla\Psi\cdot\nabla\varphi(t_{n+1})\,\mathrm{d}x \tag{5.24}$$
$$+\tau\int_{\mathbb{R}^d}\rho_{\tau,R}^{n+1}\nabla\Psi\cdot\nabla\varphi(t_{n+1})\,\mathrm{d}x. \tag{5.25}$$

Summing (5.20), we obtain

$$-\int_{\mathbb{R}^d}\rho^0\varphi(0)\,\mathrm{d}x. \tag{5.26}$$

Summing (5.22), we obtain

$$\sum_{n=0}^{N-1}\int_{\mathbb{R}^d}\left(\tilde{\rho}_{\tau,R}^{n+1}-\bar{\rho}_{\tau,R}^{n+1}\right)\varphi(t_{n+1})\,\mathrm{d}x. \tag{5.27}$$

Summing (5.24), we obtain

$$\sum_{n=0}^{N-1} \int_{\mathbb{R}^d} \left(\bar{\rho}_{\tau,R}^{n+1} - \rho_{\tau,R}^{n+1}\right) \varphi(t_{n+1}) \,\mathrm{d}x - \tau \int_{\mathbb{R}^d} \rho_{\tau,R}^{n+1} \nabla\Psi \cdot \nabla\varphi(t_{n+1}) \,\mathrm{d}x. \tag{5.28}$$

Finally, we can sum (5.25) and write

$$\begin{aligned}
&\sum_{n=0}^{N-1} \tau \int_{\mathbb{R}^d} \rho_{\tau,R}^{n+1} \nabla\Psi \cdot \nabla\varphi(t_{n+1}) \,\mathrm{d}x \\
&\quad = \sum_{n=0}^{N-2} \tau \int_{\mathbb{R}^d} \rho_{\tau,R}^{n+1} \nabla\Psi \cdot \nabla\varphi(t_{n+1}) \,\mathrm{d}x \,(\text{since } \nabla\varphi(T) = 0), \\
&\quad = \sum_{n=1}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \rho_{\tau,R}^{n} \nabla\Psi \cdot \nabla\varphi(t_n) \,\mathrm{d}x \,\mathrm{d}t.
\end{aligned} \tag{5.29}$$

Then (5.29) can be written as

$$\begin{aligned}
&\sum_{n=1}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \rho_{\tau,R}^{n} \nabla\Psi \cdot \nabla\varphi(t_n) \,\mathrm{d}x \,\mathrm{d}t \\
&\quad = \int_0^T \int_{\mathbb{R}^d} \rho_{\tau,R}(t) \nabla\Psi \cdot \nabla\varphi(t) \,\mathrm{d}x \,\mathrm{d}t \\
&\qquad - \sum_{n=1}^{N-1} \left[\int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \left[\rho_{\tau,R}(t) \nabla\Psi \cdot \nabla\varphi(t) - \rho_{\tau,R}^{n} \nabla\Psi \cdot \nabla\varphi(t_n)\right] \right] \mathrm{d}x \,\mathrm{d}t \\
&\qquad - \int_0^\tau \int_{\mathbb{R}^d} \Phi_s(t) * \rho^0 \nabla\Psi \cdot \nabla\varphi(t) \,\mathrm{d}x \,\mathrm{d}t,
\end{aligned}$$

which completes the proof. □

Lemma 5.2.4. *The remainder $\mathcal{R}(\tau, R)$ satisfies*

$$|\mathcal{R}(\tau, R)| \leq C \left(\tau + R^{-2s} + \sum_{n=0}^{N-1} W_2(\bar{\rho}_{\tau,R}^{n+1}, \rho_{\tau,R}^{n+1})^2 \right)$$

where C is independent of τ and R.

Proof. We first estimate (5.15). Write

$$\left| \int_{\mathbb{R}^d} \left(\tilde{\rho}_{\tau,R}^{n+1} - \bar{\rho}_{\tau,R}^{n+1}\right) \varphi(t_{n+1}) \,\mathrm{d}x \right| \leq \sup_{t \in [0,T]} \|\varphi(t)\|_{L^\infty(\mathbb{R}^d)} \left\|\tilde{\rho}_{\tau,R}^{n+1} - \bar{\rho}_{\tau,R}^{n+1}\right\|_{L^1(\mathbb{R}^d)}.$$

Then recalling $\bar{\rho}_{\tau,R}^{n+1} = \rho_{\tau,R}^n * (\Phi_s^\tau 1_{B_R}) \,/\, \|\Phi_s^\tau 1_{B_R}\|_{L^1(\mathbb{R}^d)}$, we estimate

$$\begin{aligned}
&\int_{\mathbb{R}^d} \left|\tilde{\rho}_{\tau,R}^{n+1} - \bar{\rho}_{\tau,R}^{n+1}\right| \mathrm{d}x \\
&\quad = \int_{\mathbb{R}^d} \left|\rho_{\tau,R}^n * \left(\Phi_s^\tau 1_{\mathbb{R}^d\setminus B_R}\right) + \rho_{\tau,R}^n * (\Phi_s^\tau 1_{B_R}) \left(1 - \frac{1}{\int_{B_R} \Phi_s^\tau(z)\,\mathrm{d}z}\right)\right| \mathrm{d}x \\
&\quad \le \int_{\mathbb{R}^d} \left[\rho_{\tau,R}^n * \left(\Phi_s^\tau 1_{\mathbb{R}^d\setminus B_R}\right) + \rho_{\tau,R}^n * (\Phi_s^\tau 1_{B_R}) \left(\frac{1}{\int_{B_R} \Phi_s^\tau(z)\,\mathrm{d}z} - 1\right)\right] \mathrm{d}x \\
&\quad = \int_{\mathbb{R}^d} \left[\left(\Phi_s^\tau 1_{\mathbb{R}^d\setminus B_R}\right)(y) + (\Phi_s^\tau 1_{B_R})(y) \left(\frac{1}{\int_{B_R} \Phi_s^\tau(z)\,\mathrm{d}z} - 1\right)\right] \underbrace{\left[\int_{\mathbb{R}^d} \rho_{\tau,R}^n(x-y)\,\mathrm{d}x\right]}_{1} \mathrm{d}y \\
&\quad = \int_{\mathbb{R}^d} \left[\Phi_s^\tau(y) 1_{\mathbb{R}^d\setminus B_R}(y) + \Phi_s^\tau(y) 1_{B_R}(y) \left(\frac{1}{\int_{B_R} \Phi_s^\tau(z)\,\mathrm{d}z} - 1\right)\right] \mathrm{d}y \\
&\quad = 2\int_{\mathbb{R}^d\setminus B_R} \Phi_s^\tau(y)\,\mathrm{d}y \\
&\quad \le C\tau \int_{\mathbb{R}^d\setminus B_R} \frac{1}{|y|^{d+2s}}\,\mathrm{d}y, \quad \text{provided } R \ge \tau^{1/2s} \text{ (by (3.3))} \\
&\quad \le \frac{C\tau}{sR^{2s}}.
\end{aligned} \tag{5.30}$$

Thus

$$\left|\sum_{n=0}^{N-1} \int_{\mathbb{R}^d} \left(\tilde{\rho}_{\tau,R}^{n+1} - \bar{\rho}_{\tau,R}^{n+1}\right) \varphi(t_{n+1})\,\mathrm{d}x\right| \le CT \sup_{t\in[0,T]} \|\varphi(t)\|_{L^\infty(\mathbb{R}^d)} R^{-2s}. \tag{5.31}$$

Next we estimate (5.16) using the estimate from (4.12)

$$\begin{aligned}
&\left|\int_{\mathbb{R}^d} \left(\bar{\rho}_{\tau,R}^{n+1} - \rho_{\tau,R}^{n+1}\right) \varphi(t_{n+1})\,\mathrm{d}x - \tau \int_{\mathbb{R}^d} \rho_{\tau,R}^{n+1} \nabla\Psi \cdot \nabla\varphi(t_{n+1})\,\mathrm{d}x\right| \\
&\qquad \le \sup_{t\in[0,T]} \left\|D^2\varphi(t)\right\|_{L^\infty(\mathbb{R}^d)} \frac{1}{2} W_2(\bar{\rho}_{\tau,R}^{n+1}, \rho_{\tau,R}^{n+1})^2,
\end{aligned}$$

so that

$$\begin{aligned}
&\left|\sum_{n=0}^{N-1} \left[\int_{\mathbb{R}^d} \left(\bar{\rho}_{\tau,R}^{n+1} - \rho_{\tau,R}^{n+1}\right) \varphi(t_{n+1})\,\mathrm{d}x - \tau \int_{\mathbb{R}^d} \rho_{\tau,R}^{n+1} \nabla\Psi \cdot \nabla\varphi(t_{n+1})\,\mathrm{d}x\right]\right| \\
&\qquad \le \sup_{t\in[0,T]} \left\|D^2\varphi(t)\right\|_{L^\infty(\mathbb{R}^d)} \frac{1}{2} \sum_{n=0}^{N-1} W_2(\bar{\rho}_{\tau,R}^{n+1}, \rho_{\tau,R}^{n+1})^2.
\end{aligned} \tag{5.32}$$

Finally, we estimate (5.17) and (5.18). We can dispense with (5.18) by noting that

$$\begin{aligned}\left|\int_0^\tau \int_{\mathbb{R}^d} \Phi_s(t) * \rho^0 \nabla\Psi \cdot \nabla\varphi(t)\,\mathrm{d}x\,\mathrm{d}t\right| &\le \int_0^\tau \|\nabla\Psi\cdot\nabla\varphi(t)\|_{L^\infty(\mathbb{R}^d)} \left\|\Phi_s(t) * \rho^0\right\|_{L^1(\mathbb{R}^d)}\,\mathrm{d}t \\ &\le \tau \sup_{t\in[0,T]} \|\nabla\Psi\cdot\nabla\varphi(t)\|_{L^\infty(\mathbb{R}^d)}. \end{aligned} \tag{5.33}$$

We can write (5.17) as

$$\begin{aligned}&\int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \rho_{\tau,R}(t)\nabla\Psi\cdot\nabla\varphi(t) - \rho_{\tau,R}^n \nabla\Psi\cdot\nabla\varphi(t_n)\,\mathrm{d}x\,\mathrm{d}t \\ &\quad= \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \left(\rho_{\tau,R}(t) - \rho_{\tau,R}^n\right) \nabla\Psi\cdot\nabla\varphi(t)\,\mathrm{d}x\,\mathrm{d}t && (5.34)\\ &\qquad+ \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \rho_{\tau,R}^n \nabla\Psi\cdot\nabla\left[\varphi(t)-\varphi(t_n)\right]\,\mathrm{d}x\,\mathrm{d}t. && (5.35)\end{aligned}$$

For (5.34), we have $\rho_{\tau,R}(t) - \rho_{\tau,R}^n = \int_{t_n}^t \partial_u \rho_{\tau,R}(u)\,\mathrm{d}u = -\int_{t_n}^t (-\Delta)^s \rho_{\tau,R}(u)\,\mathrm{d}u$, so that

$$\begin{aligned}&\int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \left(\rho_{\tau,R}(t) - \rho_{\tau,R}^n\right) \nabla\Psi\cdot\nabla\varphi(t)\,\mathrm{d}x\,\mathrm{d}t \\ &\quad= \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \int_{t_n}^t -(-\Delta)^s \rho_{\tau,R}(u)\nabla\Psi\cdot\nabla\varphi(t)\,\mathrm{d}u\,\mathrm{d}x\,\mathrm{d}t \\ &\quad= \int_{t_n}^{t_{n+1}} \int_{t_n}^t \int_{\mathbb{R}^d} -\rho_{\tau,R}(u)(-\Delta)^s\left[\nabla\Psi\cdot\nabla\varphi(t)\right]\,\mathrm{d}x\,\mathrm{d}u\,\mathrm{d}t.\end{aligned}$$

Since by assumption $\Psi \in C^{2,1}(\mathbb{R}^d)$, then $\nabla\Psi\cdot\nabla\varphi(t) \in C_c^{1,1}(\mathbb{R}^d)$, and we have from Proposition 2.2.5 that its fractional Laplacian is bounded in $\mathbb{R}^d$. Therefore

$$\begin{aligned}&\left|\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}^d} \left(\tilde\rho_{\tau,R}^{n+1} - \rho_{\tau,R}(t)\right) \nabla\Psi\cdot\nabla\varphi(t)\,\mathrm{d}x\,\mathrm{d}t\right| \\ &\quad\le \sup_{t\in[0,T]} \|(-\Delta)^s \nabla\Psi\cdot\nabla\varphi(t)\|_{L^\infty(\mathbb{R}^d)} \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{t_n}^t \underbrace{\int_{\mathbb{R}^d} \rho_{\tau,R}(u)\,\mathrm{d}x}_{1}\,\mathrm{d}u\,\mathrm{d}t \\ &\quad\le \sup_{t\in[0,T]} \|(-\Delta)^s \nabla\Psi\cdot\nabla\varphi(t)\|_{L^\infty(\mathbb{R}^d)} \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} (t-t_n)\,\mathrm{d}t \\ &\quad\le \sup_{t\in[0,T]} \|(-\Delta)^s \nabla\Psi\cdot\nabla\varphi(t)\|_{L^\infty(\mathbb{R}^d)} \frac{1}{2} T\tau. \end{aligned} \tag{5.36}$$

For (5.35), we have from a Taylor expansion $\varphi(t) - \varphi(t_n) = \partial_t \varphi(t_\lambda)(t - t_n)$ for $t_\lambda \in$

(t_n, t_{n+1}),

$$\begin{aligned}
&\left|\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}\int_{\mathbb{R}^d}\rho^n_{\tau,R}\nabla\Psi\cdot\nabla\left[\varphi(t)-\varphi(t_n)\right]\,\mathrm{d}x\,\mathrm{d}t\right| \\
&\qquad=\left|\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}\int_{\mathbb{R}^d}\rho^n_{\tau,R}(t-t_n)\nabla\Psi\cdot\nabla\partial_t\varphi(t_\lambda)\,\mathrm{d}x\,\mathrm{d}t\right| \\
&\qquad\le\sup_{t\in[0,T]}\|\nabla\Psi\cdot\nabla\partial_t\varphi(t)\|_{L^\infty(\mathbb{R}^d)}\left\|\rho^n_{\tau,R}\right\|_{L^1(\mathbb{R}^d)}\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}}(t-t_n)\,\mathrm{d}t \\
&\qquad\le\sup_{t\in[0,T]}\|\nabla\Psi\cdot\nabla\partial_t\varphi(t)\|_{L^\infty(\mathbb{R}^d)}\frac{1}{2}T\tau. \qquad (5.37)
\end{aligned}$$

Combining (5.31), (5.32), (5.33), (5.36) and (5.37), we obtain the desired estimate for $\mathcal{R}$. □

It remains to obtain an estimate on $\sum_{n=0}^{N-1}W_2(\bar\rho^{n+1}_{\tau,R},\rho^{n+1}_{\tau,R})^2$ in terms of the parameters τ and R. The usual way to obtain such an estimate is to observe from the variational problem, that since $\bar\rho^n_{\tau,R}$ is admissible and $\rho^n_{\tau,R}$ is a minimizer,

$$\frac{1}{2\tau}W_2(\bar\rho^n_{\tau,R},\rho^n_{\tau,R})^2\le\int_{\mathbb{R}^d}\bar\rho^n_{\tau,R}\Psi\,\mathrm{d}x-\int_{\mathbb{R}^d}\rho^n_{\tau,R}\Psi\,\mathrm{d}x. \qquad (5.38)$$

Disregarding the splitting scheme for the moment, if we were simply iterating the variational method, then $\bar\rho^n_{\tau,R}=\rho^{n-1}_{\tau,R}$, and thus (5.38) would become

$$\frac{1}{2\tau}W_2(\rho^{n-1}_{\tau,R},\rho^n_{\tau,R})^2\le\int_{\mathbb{R}^d}\rho^{n-1}_{\tau,R}\Psi\,\mathrm{d}x-\int_{\mathbb{R}^d}\rho^n_{\tau,R}\Psi\,\mathrm{d}x,$$

of which the right-hand side is a telescoping sum. Therefore

$$\frac{1}{2\tau}\sum_{n=1}^{N}W_2(\rho^{n-1}_{\tau,R},\rho^n_{\tau,R})^2\le\int_{\mathbb{R}^d}\rho^0\Psi\,\mathrm{d}x-\int_{\mathbb{R}^d}\rho^N_{\tau,R}\Psi\,\mathrm{d}x\le\int_{\mathbb{R}^d}\rho^0\Psi\,\mathrm{d}x,$$

and we have the desired estimate. However, in our case a difficulty arises because of the splitting. We want to estimate the potential energy of $\bar\rho^n_{\tau,R}$ in terms of the potential energy of $\rho^{n-1}_{\tau,R}$, even though $\bar\rho^n_{\tau,R}$ doesn't 'see' the potential Ψ. A sufficient way to overcome this difficulty is to impose the requirement that $D^2\Psi\in L^\infty(\mathbb{R}^d)$, or $\Psi\in C^{1,1}(\mathbb{R}^d)$. We then obtain the following.

Lemma 5.2.5. *When $\Psi \in C^{1,1} \cap C^2(\mathbb{R}^d)$, we have the estimate*

$$\begin{aligned}\sum_{n=1}^{N} & W_2(\bar{\rho}^n_{\tau,R}, \rho^n_{\tau,R})^2 \\ & \leq C \left[\tau \int_{\mathbb{R}^d} \rho^0 \Psi \, \mathrm{d}x + T \left\| D^2 \Psi \right\|_{L^\infty(\mathbb{R}^d)} \left(\tau^{1/s} + \tau R^{2-2s} \right) \right]. \end{aligned} \tag{5.39}$$

Proof. We have (since $\rho^n_{\tau,R}$ is a minimizer)

$$\frac{1}{2\tau} W_2(\bar{\rho}^n_{\tau,R}, \rho^n_{\tau,R})^2 \leq \int_{\mathbb{R}^d} \bar{\rho}^n_{\tau,R} \Psi \, \mathrm{d}x - \int_{\mathbb{R}^d} \rho^n_{\tau,R} \Psi \, \mathrm{d}x. \tag{5.40}$$

Appealing to estimate (5.5), we can obtain the inequality

$$\int_{\mathbb{R}^d} \bar{\rho}^n_{\tau,R} \Psi \, \mathrm{d}x \leq \int_{\mathbb{R}^d} \rho^{n-1}_{\tau,R} \Psi \, \mathrm{d}x + \frac{\|D^2 \Psi\|_{L^\infty(\mathbb{R}^d)}}{2} \frac{\int_{B_R} |x|^2 \Phi^\tau_s(x) \, \mathrm{d}x}{\int_{B_R} \Phi^\tau_s(x) \, \mathrm{d}x}. \tag{5.41}$$

Substituting (5.41) into (5.40), summing over n, and recalling $N = \frac{T}{\tau}$ and $\Psi \geq 0$, we obtain

$$\sum_{n=1}^{N} W_2(\bar{\rho}^n_{\tau,R}, \rho^n_{\tau,R})^2 \leq 2\tau \int_{\mathbb{R}^d} \rho^0 \Psi \, \mathrm{d}x + T \left\| D^2 \Psi \right\|_{L^\infty(\mathbb{R}^d)} \frac{\int_{B_R} |x|^2 \Phi^\tau_s(x) \, \mathrm{d}x}{\int_{B_R} \Phi^\tau_s(x) \, \mathrm{d}x}. \tag{5.42}$$

Finally all that remains is to estimate $\frac{\int_{B_R} |x|^2 \Phi^\tau_s(x) \, \mathrm{d}x}{\int_{B_R} \Phi^\tau_s(x) \, \mathrm{d}x}$. Here we make use of an estimate on the fractional heat kernel Φ^τ_s, which we recall from (3.3) is

$$\Phi^\tau_s(x) \leq C \begin{cases} \tau^{-d/2s} & |x| \leq \tau^{1/2s} \\ \tau \frac{1}{|x|^{d+2s}} & |x| > \tau^{1/2s}, \end{cases}$$

with Φ^τ_s bounded below by the same estimate, but with C replaced by C^{-1} (in the following, we do not differentiate C from C^{-1}). Then

$$\begin{aligned} \int_{B_R} \Phi^\tau_s(x) \, \mathrm{d}x &\geq C \left[\tau^{-d/2s} \int_0^{\tau^{1/2s}} r^{d-1} \, \mathrm{d}r + \tau \int_{\tau^{1/2s}}^{R} \frac{r^{d-1}}{r^{d+2s}} \, \mathrm{d}r \right] \\ &\geq C(1 + \tau R^{-2s}) \geq C > 0, \end{aligned}$$

while

$$\int_{B_R} |x|^2 \Phi_s^\tau(x)\,\mathrm{d}x \le C\left[\tau^{-d/2s}\int_0^{\tau^{1/2s}} r^2 r^{d+1}\,\mathrm{d}r + \tau\int_{\tau^{1/2s}}^{R} \frac{r^2}{r^{d+2s}} r^{d-1}\,\mathrm{d}r\right]$$
$$\le C\left[\tau^{1/s} + \tau R^{2-2s}\right].$$

Substituting these estimates into (5.42), we obtain (5.39). □

Thus it follows by Lemmas 5.2.4 and 5.2.5 that

$$|\mathcal{R}(\tau,R)| \le C\left(\tau + R^{-2s} + \tau^{1/s} + \tau R^{2-2s}\right).$$

We then are able to obtain the following.

Corollary 5.2.6. *Set $R = \tau^{-1/2}$. Then $\{\rho_{\tau,R}\}$ is a sequence in τ, which we denote by $\{\rho_\tau\}$,*

$$\left|\mathcal{R}(\tau,\tau^{-1/2})\right| \le C\left(\tau + \tau^s + \tau^{1/s}\right),$$

and therefore

$$\lim_{\tau\to 0}\int_0^T \int_{\mathbb{R}^d} \rho_\tau(t)\left[\partial_t\varphi(t) - (-\Delta)^s\varphi(t) - \nabla\Psi\cdot\nabla\varphi(t)\right]\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^d}\rho^0\varphi(0)\,\mathrm{d}x = 0.$$

Remark 5.2.7. *Setting $R = \tau^{-\alpha}$ where $0 < \alpha < 1/(2-2s)$ is also sufficient to ensure the remainder vanishes as $\tau \to 0$.*

5.3 Convergence to a Weak Solution

We now seek to obtain compactness of the sequence $\{\rho_\tau\}_{\tau\downarrow 0}$ w.r.t. some suitable topology. Since $\|\rho_\tau(t)\|_{L^1(\mathbb{R}^d)} = 1$ for all t, one may immediately wonder if it is possible to obtain weak L^1 compactness. However, this turns out to be difficult, as it is well known [8, 22] that additional properties such as tightness and uniform integrability are needed, and it is not at all clear in this case how one might proceed. Fortunately, compactness w.r.t. the weak topology on L^p ($1 < p < \infty$, or weak-$*$ if $p = \infty$) is much easier to obtain. We recall the following sufficient condition.

Proposition 5.3.1. *[10] Let $1 < p \le \infty$ and suppose $\{f_j\}_{j\in\mathbb{N}}$ is a sequence of real valued functions, $f_j : \mathbb{R}^d \to \mathbb{R}$. If $\|f_j\|_{L^p(\mathbb{R}^d)} \le C$ uniformly in j, then $\{f_j\}_{j\in\mathbb{N}}$ is relatively weakly compact in $L^p(\mathbb{R}^d)$.*

Thus by imposing a sufficient assumption on the potential Ψ, we are led to the following.

Lemma 5.3.2. *Let $\tau > 0$ be small enough so that $\det(I + \tau D^2\Psi(x)) \le 1 + \alpha\tau$, for some fixed $\alpha > \|D^2\Psi\|_{L^\infty(\mathbb{R}^d)}$. If $\rho^0 \in L^p(\mathbb{R}^d)$ for $1 < p < \infty$, then for every $t \in (0, T)$,*

$$\|\rho_\tau(t)\|^p_{L^p(\mathbb{R}^d)} \le e^{(p-1)\alpha T} \left\|\rho^0\right\|^p_{L^p(\mathbb{R}^d)} \tag{5.43}$$

and for $p = \infty$,

$$\|\rho_\tau(t)\|_{L^\infty(\mathbb{R}^d)} \le e^{\alpha T} \left\|\rho^0\right\|_{L^\infty(\mathbb{R}^d)}. \tag{5.44}$$

Proof. The fact that $\det(I + \tau D^2\Psi(x)) \le 1 + \alpha\tau$ for τ small enough assures us that

$$\|\rho^n_\tau\|^p_{L^p(\mathbb{R}^d)} \le (1 + \tau\alpha)^{p-1} \|\bar{\rho}^n_\tau\|^p_{L^p(\mathbb{R}^d)}$$

and since,

$$\|\bar{\rho}^n_\tau\|^p_{L^p(\mathbb{R}^d)} = \left\| \frac{\Phi_s 1_{B_R} * \rho^{n-1}_\tau}{\|\Phi_s 1_{B_R}\|_{L^1(\mathbb{R}^d)}} \right\|^p_{L^p(\mathbb{R}^d)} \le \left\|\rho^{n-1}_\tau\right\|^p_{L^p(\mathbb{R}^d)},$$

we have

$$\|\rho^n_\tau\|^p_{L^p(\mathbb{R}^d)} \le (1 + \tau\alpha)^{n(p-1)} \left\|\rho^0\right\|^p_{L^p(\mathbb{R}^d)}.$$

Then for $1 < p < \infty$ and $t \in (t_n, t_{n+1})$,

$$\begin{aligned}
\|\rho_\tau(t)\|^p_{L^p(\mathbb{R}^d)} &= \|\Phi_s(t - t_n) * \rho^n_\tau\|^p_{L^p(\mathbb{R}^d)} \\
&\le \|\rho^n_\tau\|^p_{L^p(\mathbb{R}^d)} \\
&\le (1 + \tau\alpha)^{n(p-1)} \left\|\rho^0\right\|^p_{L^p(\mathbb{R}^d)} \\
&\le (1 + \tau\alpha)^{\frac{T}{\tau}(p-1)} \left\|\rho^0\right\|^p_{L^p(\mathbb{R}^d)},
\end{aligned}$$

and (5.43) follows by noting $\lim_{\tau\downarrow 0}(1 + \alpha\tau)^{\frac{T}{\tau}(p-1)} = e^{\alpha(p-1)T}$; (5.44) is obtained in a similar manner. $\square$

Lemma 5.3.3. *If $\rho^0 \in L^p(\mathbb{R}^d)$, $1 < p \le \infty$, then there exists a non-relabelled subsequence $\{\rho_\tau\}_{\tau\downarrow 0}$ and a $\rho \in L^1 \cap L^p(\mathbb{R}^d \times (0, T))$ such that $\rho_\tau \rightharpoonup \rho$ in $L^p(\mathbb{R}^d \times (0, T))$ (or $\rightharpoonup^*$ if $p = \infty$), and*

$$\int_0^T \int_{\mathbb{R}^d} \rho(t) \left[\partial_t\varphi(t) - (-\Delta)^s\varphi(t) - \nabla\Psi \cdot \nabla\varphi(t)\right] \mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^d} \rho^0\varphi(0)\,\mathrm{d}x = 0.$$

Moreover, if $\rho^0 \ge 0$, then $\rho(x, t) \ge 0$ for a.e. $(x, t) \in \mathbb{R}^d \times (0, T)$.

Proof. By Proposition 5.3.1 and Lemma 5.3.2, we deduce the existence of a $\rho \in L^p(\mathbb{R}^d \times (0,T))$ such that $\rho_\tau \rightharpoonup \rho$ in $L^p(\mathbb{R}^d \times (0,T))$. Then by appealing to Corollary 5.2.6 and the fact that $\partial_t \varphi(t) - (-\Delta)^s \varphi(t) - \nabla\Psi \cdot \nabla\varphi(t) \in L^{p'}(\mathbb{R}^d \times (0,T))$ for every $\varphi \in C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ with time support in $[-T,T]$ (where $1/p + 1/p' = 1$), we have

$$\int_0^T \int_{\mathbb{R}^d} \rho(t) \left[\partial_t \varphi(t) - (-\Delta)^s \varphi(t) - \nabla\Psi \cdot \nabla\varphi(t)\right] \mathrm{d}x \, \mathrm{d}t + \int_{\mathbb{R}^d} \rho^0 \varphi(0) \, \mathrm{d}x = 0.$$

If $\rho^0 \geq 0$, then it follows by definition of ρ_τ that $\rho_\tau(x,t) \geq 0$ for every $(x,t) \in \mathbb{R}^d \times (0,T)$. Therefore, combined with the weak convergence of ρ_τ to ρ in $L^p(\mathbb{R}^d \times (0,T))$, we deduce $\int_0^T \int_{\mathbb{R}^d} \rho(x,t)\lambda(x,t) \, \mathrm{d}x \, \mathrm{d}t \geq 0$ for all step-functions $\lambda : \mathbb{R}^d \times (0,T) \to \mathbb{R}$ with compact support. Thus $\rho(x,t) < 0$ is strictly forbidden on any subset of $\mathbb{R}^d \times (0,T)$ with positive measure, and hence $\rho(x,t) \geq 0$ for a.e. $(x,t) \in \mathbb{R}^d \times (0,T)$.

Finally, we have $\rho \in L^1(\mathbb{R}^d \times (0,T))$, since, for each $R > 0$, 1_{B_R} is in $L^{p'}(\mathbb{R}^d \times (0,T))$, and

$$\begin{aligned}\int_0^T \int_{B_R} \rho(x,t) \, \mathrm{d}x \, \mathrm{d}t &= \int_0^T \int_{B_R} \rho_\tau(x,t) \, \mathrm{d}x \, \mathrm{d}t + \int_0^T \int_{B_R} \left[\rho(x,t) - \rho_\tau(x,t)\right] \mathrm{d}x \, \mathrm{d}t \\ &\leq T + \int_0^T \int_{B_R} \left[\rho(x,t) - \rho_\tau(x,t)\right] \mathrm{d}x \, \mathrm{d}t.\end{aligned}$$

Thus by the weak convergence of ρ_τ to ρ,

$$\int_0^T \int_{B_R} \rho(x,t) \, \mathrm{d}x \, \mathrm{d}t \leq T.$$

Using the fact that $\rho(x,t) \geq 0$ a.e. (x,t), we may apply the monotone convergence theorem to obtain in the limit $R \to \infty$,

$$\int_0^T \int_{\mathbb{R}^d} \rho(x,t) \, \mathrm{d}x \, \mathrm{d}t \leq T.$$

□

According to Lemma 5.3.3, we are very close to proving that the candidate ρ is indeed a weak solution to (1.3) according to our definition. We still need to show however, that $\rho(t)$ is a probability density for a.e. t.

Lemma 5.3.4. *Let $T < \infty$, $\Psi \in C^{1,1}(\mathbb{R}^d)$, and suppose there exists $\rho \in L^1(\mathbb{R}^d \times$*

$(0,T))$ *satisfying*

$$\int_0^T \int_{\mathbb{R}^d} \rho(t) \left[\partial_t \varphi(t) - (-\Delta)^s \varphi(t) - \nabla\Psi \cdot \nabla\varphi(t)\right] \mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^d} \rho^0 \varphi(0)\,\mathrm{d}x = 0 \tag{5.45}$$

for all $\varphi \in C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ *with time support in* $[-T,T]$*. Then* $\int_{\mathbb{R}^d} \rho(t)\,\mathrm{d}x = \int_{\mathbb{R}^d} \rho^0\,\mathrm{d}x = 1$ *for a.e.* $t \in (0,T)$*.*

Proof. Let $\eta : [0,\infty) \to [0,\infty)$ be a smooth function satisfying $\eta(r) = 1$ for $r \le 1$ and $\eta(r) = 0$ for $r > 2$; for instance,

$$\eta(r) = \begin{cases} 1 & 0 \le r \le 1 \\ e^{1 - \frac{1}{1-(r-1)^2}} & 1 < r \le 2 \\ 0 & r > 2 \end{cases} \tag{5.46}$$

and let $\eta_R \in C_c^\infty(\mathbb{R}^d)$ be defined by $\eta_R(x) = \eta\left(\frac{|x|}{R}\right)$ for $R > 0$. Then for all $\theta \in C_c^\infty(-T,T)$, we have from (5.45) (with $\varphi(x,t) = \eta_R(x)\theta(t)$)

$$\begin{aligned} &\int_0^T \theta'(t) \int_{\mathbb{R}^d} \rho(x,t)\eta_R(x)\,\mathrm{d}x\,\mathrm{d}t + \theta(0) \int_{\mathbb{R}^d} \rho^0(x)\eta_R(x)\,\mathrm{d}x \\ &\quad = \int_0^T \theta(t) \int_{\mathbb{R}^d} \rho(x,t) \left[(-\Delta)^s \eta_R(x) + \nabla\Psi(x) \cdot \nabla\eta_R(x)\right] \mathrm{d}x\,\mathrm{d}t. \end{aligned} \tag{5.47}$$

Noting that $\lim_{R\to\infty} \eta_R(x) = 1$ pointwise on $\mathbb{R}^d$, assume for the moment that

$$\lim_{R\to\infty} (-\Delta)^s \eta_R(x) = 0 \quad \text{and} \quad \lim_{R\to\infty} \nabla\Psi(x) \cdot \nabla\eta_R(x) = 0, \quad \text{pointwise on } \mathbb{R}^d,$$

and that by the dominated convergence theorem we may pass these limits inside the integrals in the above display. (We will rigorously justify these assertions later.) Then we obtain in the limit $R \to \infty$

$$\int_0^T \theta'(t) \int_{\mathbb{R}^d} \rho(x,t)\,\mathrm{d}x\,\mathrm{d}t + \theta(0) \int_{\mathbb{R}^d} \rho^0(x)\,\mathrm{d}x = 0, \tag{5.48}$$

for every $\theta \in C_c^\infty(-T,T)$. In particular, for every $\gamma \in C_c^\infty(0,T)$ we have

$$\int_0^T \gamma'(t) \int_{\mathbb{R}^d} \rho(x,t)\,\mathrm{d}x\,\mathrm{d}t = 0,$$

from which we can deduce [10] that

$$\int_{\mathbb{R}^d} \rho(x,t)\,\mathrm{d}x = C, \quad \text{a.e. } t \in (0,T). \tag{5.49}$$

The constant is fixed by appealing to (5.48), since substitution of (5.49) in (5.48) implies

$$C\int_0^T \theta'(t)\,\mathrm{d}t + \theta(0)\int_{\mathbb{R}^d} \rho^0(x)\,\mathrm{d}x = 0,$$

from which we deduce

$$C = \int_{\mathbb{R}^d} \rho^0(x)\,\mathrm{d}x.$$

Now we rigorously prove the asserted limits above and application of the dominated convergence theorem. First, it is immediate that

$$\lim_{R\to\infty} \nabla\Psi(x)\cdot\nabla\eta_R(x) = 0,$$

pointwise for $x \in \mathbb{R}^d$, for, we simply wait until $|x| < R$ for R large enough, and then $\nabla\eta_R(x) = 0$. To apply the dominated convergence theorem in (5.47), we need a L^∞ bound for $\nabla\Psi\cdot\nabla\eta_R$. Since $\nabla\eta_R(x) = \frac{x}{|x|R}\eta'\left(\frac{|x|}{R}\right)$ for $R < |x| < 2R$ (and 0 elsewhere), and $|\eta'(r)| \le C$ for all $r \in (1,2)$, then

$$|\nabla\eta_R(x)| \le \frac{C}{R}.$$

Now using the fact that $\Psi \in C^{1,1}(\mathbb{R}^d)$ and recalling $\nabla\Psi(x)\cdot\nabla\eta_R(x)$ is non-zero only when $R < |x| < 2R$, we have

$$\begin{aligned} |\nabla\Psi(x)\cdot\nabla\eta_R(x)| &\le C(1+|x|)\,|\nabla\eta_R(x)| \\ &\le C\frac{1+2R}{R} \end{aligned} \tag{5.50}$$

and thus $\|\nabla\Psi\cdot\nabla\eta_R\|_{L^\infty(\mathbb{R}^d)} \le C$ for all R.

Now we show $\lim_{R\to\infty}(-\Delta)^s\eta_R(x) = 0$ pointwise by an application of the dominated convergence theorem to pass the limit inside the integral representation of $(-\Delta)^s\eta_R(x)$. Indeed, since in some neighbourhood B_r of the origin,

$$\frac{|\eta_R(x+z)+\eta_R(x-z)-2\eta_R(x)|}{|z|^{d+2s}} \le \left\|D^2\eta_R\right\|_{L^\infty(\mathbb{R}^d)}|z|^{2-d-2s}, \quad z \in B_r, \tag{5.51}$$

and outside B_r,

$$\frac{|\eta_R(x+z)+\eta_R(x-z)-2\eta_R(x)|}{|z|^{d+2s}} \le \frac{4}{|z|^{d+2s}}, \quad z \in \mathbb{R}^d \backslash B_r, \tag{5.52}$$

it suffices by the dominated convergence theorem to bound $\|D^2\eta_R\|_{L^\infty(\mathbb{R}^d)}$ uniformly in R. In particular, it suffices to have the second-order partial derivatives $|\partial_{ij}^2\eta_R(x)| \le C$. By direct computation, when $R < |x| < 2R$,

$$\partial_{ij}^2\eta_R(x) = \frac{x_i x_j}{|x|^2 R^2}\eta''\left(\frac{|x|}{R}\right) - \frac{x_i x_j}{|x|^3 R}\eta'\left(\frac{|x|}{R}\right),$$

so that

$$\left|\partial_{ij}^2\eta_R(x)\right| \le \frac{C}{R^2} \le C. \tag{5.53}$$

Hence $\lim_{R\to\infty}(-\Delta)^s\eta_R(x) = 0$.

Finally we remark that passage of the limits through the integrals in (5.47) follows by the uniform (w.r.t. R) L^∞ bounds on $\nabla\Psi\cdot\nabla\eta_R$ (from (5.50)) and $(-\Delta)^s\eta_R$ (from (5.51), (5.52), and (5.53)), together with the fact that $\rho \in L^1(\mathbb{R}^d \times (0,T))$. □

5.3.1 Proof of the Main Result

Combining the results obtained in the previous sections, we can now prove that the constructed splitting scheme weakly converges in L^p to a weak solution of (1.3).

Theorem 5.3.5. *Let $s \in (0,1)$, $\rho^0 \in L^1 \cap L^p(\mathbb{R}^d)$ for some $1 < p \le \infty$, and $\Psi \in C^{1,1}\cap C^{2,1}(\mathbb{R}^d)$, with $\Psi \ge 0$ and $\int_{\mathbb{R}^d}\rho^0\Psi\,\mathrm{d}x < \infty$. Define the sequence $\{\rho_\tau\}_{\tau\downarrow 0}$ according to (5.14) with $R = \tau^{-1/2}$. Then there exists a non-relabeled subsequence $\{\rho_\tau\}_{\tau\downarrow 0}$ and a $\rho \in L^1 \cap L^p(\mathbb{R}^d \times (0,T))$, such that*

$$\rho_\tau \rightharpoonup \rho, \ \textit{weakly in } L^p(\mathbb{R}^d \times (0,T)),$$

where ρ is a weak solution of (1.3).

Proof. By Lemma 5.3.3, there is a $\rho \in L^1 \cap L^p(\mathbb{R}^d \times (0,T))$ and a non-relabeled subsequence $\{\rho_\tau\}_{\tau\downarrow 0}$ such that $\rho_\tau(t) \rightharpoonup \rho(t)$ weakly in $L^p(\mathbb{R}^d)$ for a.e. $t \in (0,T)$, with

$$\int_0^T\int_{\mathbb{R}^d}\rho(t)\left[\partial_t\varphi(t) - (-\Delta)^s\varphi(t) - \nabla\Psi\cdot\nabla\varphi(t)\right]\mathrm{d}x\,\mathrm{d}t + \int_{\mathbb{R}^d}\rho^0\varphi(0)\,\mathrm{d}x = 0 \tag{5.54}$$

for all $\varphi \in C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ with time support in $[-T, T]$. Moreover, by Lemma 5.3.4, we deduce that $\rho(t)$ is a probability density for a.e. t. Thus ρ is a weak solution to (1.3), in the sense of Definition 5.2.1. $\square$

Chapter 6

Conclusion

6.1 Concluding Remarks

We have shown that by splitting (1.3) into two parts (fractional heat and transport), one of which (the transport equation) can be solved by running a gradient flow of an energy functional (the potential energy) w.r.t. the 2-Wasserstein distance, it is possible to recover a solution to the original PDE (1.3). More precisely, we have shown how to construct a function from solutions of the fractional heat equation and minimizers of the transport variational problem, that weakly converges to a weak solution of (1.3).

We remark on some of the difficulties. Without imposing the strong assumption that $D^2\Psi$ is bounded (i.e. $\Psi \in C^{1,1}(\mathbb{R}^d)$), it is not clear how one may obtain compactness of the sequence $\{\rho_\tau\}_{\tau\downarrow 0}$. Although the fractional heat equation preserves the L^p norm, the same is not true in general for the transport equation. When we compute the L^p norm of $\rho_\tau(t)$, we find ourselves needing to estimate

$$\int_{\mathbb{R}^d} \left[\bar{\rho}_\tau^n \left(x + \tau\nabla\Psi(x)\right) \det(I + \tau D^2\Psi(x))\right]^p \, \mathrm{d}x$$

in terms of $\int_{\mathbb{R}^d} \left[\bar{\rho}_\tau^n\left(x\right)\right]^p \, \mathrm{d}x$, and it is not clear if any such estimate is available without the required assumption on Ψ. This problem persists even if we take $s = 1$ - the Gaussian case- and thus we suspect that one may be able to relax the assumption with a finer analysis.

Since we are free to choose our time interpolation, one may also wonder if there is another way to construct ρ_τ which might be easier to work with. In other works (see [18]) the interpolation was constructed by dividing each interval (t_n, t_{n+1}) in half, and defining on each subinterval the time interpolation to be (e.g. in our case) the solution of the fractional heat equation on $(t_n, t_{n+1/2})$, and then the solution of the transport equation on $(t_{n+1/2}, t_{n+1})$. By showing that the sequence $\{\rho_\tau\}_{\tau\downarrow 0}$ has a uniform spatial, and uniform temporal, modulus of continuity, e.g.

$$\int_{\mathbb{R}^d} [\rho_\tau(x+h,t) - \rho_\tau(x,t)]^p \, \mathrm{d}x \leq \nu(h),$$

ν not depending on τ, they were able to obtain compactness in the strong topology of $L^p_{loc}(\mathbb{R}^d \times (0,T))$ (see [18]). Some initial experimentation with this approach proved unsuccessful.

6.2 Open Questions

6.2.1 Regularity and Uniqueness

Regularity

An immediate question that comes to mind is whether the weak solution is actually a regular solution. In [19], they sketch a bootstrap argument that shows a weak solution of the classical Fokker-Planck equation (1.1) is in fact a smooth solution. One may hope that the proof can be extended when we replace the Laplacian by the fractional Laplacian. The answer, however, is negative, and this is due to the non-local nature of the fractional Laplacian. More precisely, to imitate their proof, one would need to compute $(-\Delta)^s[\eta\xi]$ where $\eta, \xi \in C_c^\infty(\mathbb{R}^d)$. The lack of a 'product rule' is what impedes further progress, and we therefore leave the question of regularity open. Our conjecture is that the regularity of the fractional Fokker-Planck equation (1.3) is the same as the classical one (1.1).

Uniqueness

Provided that a weak solution is in fact a classical smooth solution, it is possible to prove uniqueness by imitating the proof given in [19]. Let us sketch it here.

Suppose ρ_1 and ρ_2 are solutions of (1.3), smooth enough so that the following computations make sense, and let $\rho := \rho_1 - \rho_2$. Let $\delta > 0$, and $\phi_\delta(z) := (z^2 + \delta^2)^{1/2}, z \in \mathbb{R}$ a smooth convex approximation to the absolute value function. Then by a straightforward computation, (where for clarity we suppress the (x,t) dependence of ρ) we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\phi_\delta(\rho) &+ (-\Delta)^s\left[\phi_\delta(\rho)\right] - \operatorname{div}\left(\phi_\delta(\rho)\nabla\Psi\right)\\ &= (-\Delta)^s\left[\phi_\delta(\rho)\right] - \phi_\delta'(\rho)(-\Delta)^s\rho + \left[\rho\phi_\delta'(\rho) - \phi_\delta(\rho)\right]\Delta\Psi.\end{aligned}$$

Now, we introduce the following lemma from [14].

Lemma 6.2.1. *Let $\varphi \in C^2(\mathbb{R})$ be a convex function. Then for all $s \in (0,1)$,*

$$(-\Delta)^s\left[\varphi(f)\right] \le \varphi'(f)(-\Delta)^s f.$$

Proof. The proof is taken from [14]. Since φ is convex, then $\varphi(b)-\varphi(a) \ge \varphi'(a)(b-a)$ for all $a, b \in \mathbb{R}$. Therefore it follows that

$$\begin{aligned}&\varphi(f(x+z)) - \varphi(f(x)) \ge \varphi'(f(x))(f(x+z) - f(x))\\ &\varphi(f(x+z)) - \varphi(f(x)) - \nabla[\varphi(f)](x)\cdot z \ge \varphi'(f(x))\left[f(x+z) - f(x) - \nabla f(x)\cdot z\right]\end{aligned}$$

for every $x, z \in \mathbb{R}^d$. By appealing to the integral representation (2.1) of $(-\Delta)^s$ we conclude the result. $\square$

Then identifying ϕ_δ with φ and ρ with f in the lemma above, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_\delta(\rho) + (-\Delta)^s\left[\phi_\delta(\rho)\right] - \operatorname{div}\left(\phi_\delta(\rho)\nabla\Psi\right) \le \left[\rho\phi_\delta'(\rho) - \phi_\delta(\rho)\right]\Delta\Psi.$$

Multiply by a nonnegative $\eta \in C_c^\infty(\mathbb{R}^d)$, integrate over $\mathbb{R}^d$, and integrate by parts, to get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\phi_\delta(\rho(t))\eta &+ \int_{\mathbb{R}^d}\phi_\delta(\rho(t))(-\Delta)^s\left[\eta\right] + \int_{\mathbb{R}^d}\phi_\delta(\rho(t))\nabla\Psi\cdot\nabla\eta\\ &\le \int_{\mathbb{R}^d}\left[\rho\phi_\delta'(\rho(t)) - \phi_\delta(\rho(t))\right]\Delta\Psi\eta.\end{aligned}$$

Now integrate over $(0,t)$ for some $t \in (0,\infty)$ to obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d} \phi_\delta(\rho(t))\eta \,\mathrm{d}x - \int_{\mathbb{R}^d} \phi_\delta(\rho(0))\eta \,\mathrm{d}x \\
&\quad + \int_0^t \int_{\mathbb{R}^d} \phi_\delta(\rho(t))(-\Delta)^s \eta \,\mathrm{d}x \,\mathrm{d}t + \int_0^t \int_{\mathbb{R}^d} \phi_\delta(\rho(t)) \nabla\Psi \cdot \nabla\eta \,\mathrm{d}x \,\mathrm{d}t \\
&\quad \leq \int_0^t \int_{\mathbb{R}^d} \left[\rho\phi_\delta'(\rho(t)) - \phi_\delta(\rho(t))\right] \eta\Delta\Psi \,\mathrm{d}x \,\mathrm{d}t.
\end{aligned}$$

In the above, we remark that we have assumed

$$\lim_{t\to 0} \int_{\mathbb{R}^d} \phi_\delta(\rho(t))\eta \,\mathrm{d}x = \int_{\mathbb{R}^d} \phi_\delta(\rho(0))\eta \,\mathrm{d}x$$

for every $\delta > 0$ and every $\eta \in C_c^\infty(\mathbb{R}^d)$.

Letting δ tend to 0 gives

$$\int_{\mathbb{R}^d} |\rho(t)|\eta - \int_{\mathbb{R}^d} |\rho(0)|\eta + \int_0^t \int_{\mathbb{R}^d} |\rho(t)|(-\Delta)^s\eta + \int_0^t \int_{\mathbb{R}^d} |\rho(t)|\nabla\Psi \cdot \nabla\eta \leq 0$$

since $\rho\phi_\delta'(\rho(t)) - \phi_\delta(\rho(t)) \to 0$ as $\delta \to 0$. Now replace (as we did in Lemma 5.3.4, and as in [19]) η with

$$\eta_R(x) = \eta_1\left(\frac{x}{R}\right), \quad \text{where } \eta_1(x) = 1 \text{ for } |x| \leq 1, \eta_1(x) = 0 \text{ for } |x| \geq 2.$$

Letting $R \to \infty$ as in Lemma 5.3.4 (assuming $\Psi \in C^{1,1}(\mathbb{R}^d)$), we find

$$\int_{\mathbb{R}^d} |\rho(t)| \leq \int_{\mathbb{R}^d} |\rho(0)|,$$

which implies uniqueness.

6.2.2 Extension of the Method

Finally we conclude by suggesting that a direction for future work is to see if this method of combining splitting with gradient flow can be extended to other PDE's for which the gradient flow part is not as simple as that which we have studied here.

Bibliography

[1] Martial Agueh. Local existence of weak solutions to kinetic models of granular media. *Preprint*.

[2] Nathaël Alibaud. Entropy formulation for fractal conservation laws. *J. Evol. Equ.*, 7(1):145–175, 2007.

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Birkhäuser Verlag, Basel, 2005.

[4] David Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge University Press, Cambridge, 2004.

[5] Piotr Biler and Grzegorz Karch. Generalized Fokker-Planck equations and convergence to their equilibria. In *Evolution equations (Warsaw, 2001)*, volume 60 of *Banach Center Publ.*, pages 307–318. Polish Acad. Sci., Warsaw, 2003.

[6] Patrick Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1999.

[7] R. M. Blumenthal and R. K. Getoor. Some theorems on stable processes. *Trans. Amer. Math. Soc.*, 95:263–273, 1960.

[8] Haim Brezis. *Functional Analysis, Sobolev Spaces, and Partial Differential Equations*. Springer, New York, 2011.

[9] Catherine Choquet and Marie-Christine Néel. From particles scale to anomalous or classical convection-diffusion models with path integrals. *Discrete Contin. Dyn. Syst. Ser. S*, 7(2):207–238, 2014.

[10] Bernard Dacorogna. *Introduction to the Calculus of Variations*. Imperial College Press, London, 2004.

[11] S. Daneri and G. Savaré. Lecture Notes on Gradient Flows and Optimal Transport. *ArXiv e-prints*, September 2010.

[12] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *ArXiv e-prints*, April 2011.

[13] J. Droniou, T. Gallouet, and J. Vovelle. Global solution and smoothing effect for a non-local regularization of a hyperbolic equation. *J. Evol. Equ.*, 3(3):499–521, 2003.

[14] Jérôme Droniou and Cyril Imbert. Fractal first-order partial differential equations. *Arch. Ration. Mech. Anal.*, 182(2):299–331, 2006.

[15] M. Erbar. Gradient flows of the entropy for jump processes. *ArXiv e-prints*, April 2012.

[16] Lawrence C. Evans. *Partial Differential Equations.* American Mathematical Society, Providence, R.I, 1998.

[17] Ivan Gentil and Cyril Imbert. The Lévy-Fokker-Planck equation: Φ-entropies and convergence to equilibrium. *Asymptot. Anal.*, 59(3-4):125–138, 2008.

[18] Helge Holden, Kenneth H. Karlsen, Knut-Andreas Lie, and Nils Henrik Risebro. *Splitting methods for partial differential equations with rough solutions.* EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich, 2010. Analysis and MATLAB programs.

[19] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[20] Ioannis Karatzas. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York, 1991.

[21] David Kinderlehrer and Adrian Tudorascu. Transport via mass transportation. *Discrete Contin. Dyn. Syst. Ser. B*, 6(2):311–338 (electronic), 2006.

[22] Pablo Pedregal. *Parametrized Measures and Variational Principles.* Birkhäuser Verlag, Basel, 1997.

[23] J. Perrin and F. Soddy. *Brownian Movement and Molecular Reality.* Dover Books on Physics Series. Dover Publications, 2005.

[24] Luis Silvestre. Regularity of the obstacle problem for a fractional power of the Laplace operator. *Comm. Pure Appl. Math.*, 60(1):67–112, 2007.

[25] T. H. Solomon, Eric R. Weeks, and Harry L. Swinney. Chaotic advection in a two-dimensional flow: Lévy flights and anomalous diffusion. *Physica D*, pages 70–84, 1994.

[26] I. Tristani. Fractional Fokker-Planck equation. *ArXiv e-prints*, December 2013.

[27] Enrico Valdinoci. From the long jump random walk to the fractional Laplacian. *Bol. Soc. Esp. Mat. Apl. S$\vec{e}$MA*, (49):33–44, 2009.

[28] Cédric Villani. *Topics in Optimal Transportation*. American Mathematical Society, Providence, RI, 2003.