LÉVY MOTION AND COLD ATOMS

by

WESLEY W. ERICKSON

A DISSERTATION

Presented to the Department of Physics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2020

ProQuest Number: 28000459

ProQuest.

ProQuest 28000459

DISSERTATION APPROVAL PAGE

Student: Wesley W. Erickson

Title: Lévy Motion and Cold Atoms

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Physics by:

Michael Raymer	Chair
Daniel A. Steck	Advisor
Eric Corwin	Core Member
Marina Guenza	Institutional Representative

and

Kate Mondloch	Interim Vice Provost and Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2020

DISSERTATION ABSTRACT

Wesley W. Erickson

Doctor of Philosophy

Department of Physics

June 2020

Title: Lévy Motion and Cold Atoms

Lévy processes are a universal model for characterizing the behavior of extreme events and anomalously diffusive systems. They are also important in modeling the transport of laser-cooled atoms. This dissertation contributes to the understanding of Lévy processes themselves, and to their application in laser-cooled atomic dynamics.

Lévy processes are models of systems that contain extreme events. However, since any particular event could, upon closer inspection, be composed of multiple smaller events, what makes an event "extreme"? To answer this question it is useful to consider the *history* of the event, by studying Lévy processes conditioned to have a fixed final state at some later time. We find that events that greatly exceed a particular threshold are more likely to be composed of many small events and a *single* large event, rather than a series of comparably sized events. Analysis of the source of this threshold helps clarify the Gaussian limit of Lévy processes,

and serves as the foundation for a useful distinction between "short" and "long" steps. These ideas also suggest possible experimental techniques that may be employed for cold atoms.

Studies of anomalous diffusion in laser-cooled atoms have typically focused on the expansion profiles of clouds of cold atoms. However, the results of experiments and theoretical models have not been in close agreement, suggesting that existing models are still incomplete. To address this, it is important to develop alternative experimental approaches. One such approach is to use a single atom as a probe of the diffusion, which allows for the collection of boundary crossing statistics, such as the time between when an atom enters and leaves an imaging region. Through simulations, we find that distributions of these statistics develop peaks that correspond to atomic Lévy flights, in direct contrast to featureless power-law distributions for atoms undergoing Brownian motion. We find that characterizing these distributions gives information on the anomalous-diffusion exponent and typical velocities. Furthermore, these distributions serve as the basis for a method to directly detect Lévy flights at the level of a single atom.

CURRICULUM VITAE

NAME OF AUTHOR: Wesley W. Erickson

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:
    University of Oregon, Eugene, OR
    Reed College, Portland, OR

DEGREES AWARDED:
    Doctor of Philosophy, Physics, 2020, University of Oregon
    Master of Science, Physics, 2014, University of Oregon
    Bachelor of Arts, Physics, 2012, Reed College

AREAS OF SPECIAL INTEREST:
    Laser-cooled Atoms
    Computational Physics
    Stochastic Processes

PROFESSIONAL EXPERIENCE:

    Graduate Employee, University of Oregon, 2017-2020

    Graduate Teaching Fellow, University of Oregon, 2012-2017

    Senior Reactor Operator, Reed Research Reactor, 2010-2012

    Reactor Operator, Reed Research Reactor, 2009-2010

GRANTS, AWARDS AND HONORS:

    International High Performance Computing Summer School, Toronto, ON,
     Summer 2015

    Research Experience for Undergraduates, University of Arkansas,
     Summer 2010

PUBLICATIONS:

Wesley W. Erickson and Daniel A. Steck, "The Anatomy of an Extreme
Event: What Can We Infer About the History of a Heavy-Tailed Random
Walk?", *submitted for publication (arXiv:2002.03849)*, (2020).

ACKNOWLEDGEMENTS

I owe a great many thanks to a great many people who have helped me throughout the completion of this dissertation. While I would like to thank everyone by name, below I only mention several, but you all have my sincerest appreciation.

First and foremost, I would like to thank my advisor Daniel Steck, who introduced me to truly fascinating and beautiful ideas about the natural world, many of which will continue to hold my curiosity for years to come. More importantly, Dan has been an endlessly optimistic and supportive mentor, and his patience and encouragement were essential as I completed this dissertation.

I also want to thank my committee members, Michael Raymer, Eric Corwin, and Marina Guenza, whose feedback and challenging questions helped me clarify and strengthen this work, as well as Steven van Enk, who provided several useful ideas for framing some of the results on conditioned processes.

Next, I want to thank all the members of Steck Lab, Eryn Cook, Jonathan Mackrory, Paul Martin, and Richard Wagner, who were all welcoming when I joined the lab, generous in showing me the ropes, and often up for an interesting conversation.

Also many thanks to the friends who made my graduate school experience full of wonderful memories: Mayra Amezcua, Sripoorna "SP" Bharadwaj, Gabriel Barello, George de Coster, Herbie Grotewohl, Erik Keever, Sofiane Merkouche, Chris Newby, Dileep Reddy, Rudy Resch, and Richard Wagner. We had many exciting events and adventures, both in physics but just as often outside it, and graduate school was all the better for it.

I also wish to thank the many curious and supportive members of my family. There are too many to name, but relaxing with you all during the holidays made it easy to return determined. But I do want to especially thank my parents, who have stuck with me for quite a long education. Your unwavering support has made everything easier, I can't thank you enough.

And last but certainly not least, I want to thank my girlfriend Jun, who has been thoughtful and encouraging at every turn. Thank you for everything!

TABLE OF CONTENTS

LIST OF FIGURES

Figure   Page

xvii

CHAPTER I

INTRODUCTION

Universality refers to an observed tendency for a broad variety of different systems to display remarkably similar behavior. For example, given the right circumstances, both the motion of a molecule suspended in a fluid [2], and the motion of a star gravitationally interacting with other stars [3], can be effectively modeled by Brownian motion, despite the vastly different scales, composition, and forces involved. While the microscopic details of these systems differ completely, these details are indiscernible when the system is viewed at larger scales as the behaviors converge to the universal ones.

This example also alludes to the property of scale invariance, where a system at one scale behaves similarly to itself when viewed at a larger scale. It is this property that is often present in universal systems, as the same mechanism that relates different systems can also relate a system to itself. These universal and self-similar relations often arise as power-law scaling relationships, characterized by universal exponents that are independent of the system parameters. Universal techniques are generalizable to many systems, and the resulting models are more convenient for analysis, even if the true physical system is complex or difficult to measure directly.

As there are many classes and subclasses of systems that exhibit universal properties, it is essential to demonstrate concrete examples of universal processes as well as specific applications of these principles to interesting and useful physical systems. In this thesis we will discuss the Lévy processes, stochastic processes constructed from a general type of random walk, which is a natural model for

extreme events and anomalous diffusion. As Lévy processes are a generalizable mathematical tool, we will touch on many subjects where they are applicable, before focusing on how they can be applied in the physics of laser-cooled atomic systems.

1.1. From Random walks to a Universal Stochastic Process

A convenient starting point for the development of random walks is the year 1900, when Louis Bachelier completed his thesis "The Theory of Speculation," which studied the pricing of stock options by introducing the first mathematical model of what would later be known as Brownian motion [4, 5]. Only five years later, the first use of the term *random walk* appeared in a letter to *Nature* in discussion of a model of the migratory behavior of mosquitoes, which used the same statistics as those in Bachelier's thesis [6]. And, later that same month, Einstein's paper on Brownian motion was published [2], explaining that the persistent irregular motion of particles suspended in a fluid (as earlier observed with pollen grains by Robert Brown [7]) could be explained via atomic theory and thermal molecular motion, again with the same random-walk statistics.

It is fascinating that such similar behavior was observed in the dynamics of stock prices, animal migratory behavior, and diffusive atomic motion. As might be expected, the commonality of these systems is not accidental: each system used some form of the central limit theorem as the mechanism causing convergence to Brownian motion in the asymptotic limit of long-times. However, while the long time limit was used here, it is natural to ask what the limit looks like at short times. Within 3 years of Einstein's results, Perrin published experimental results confirming Einstein's theory, and thus the atomic nature of matter [8].

His experiment involved tracing the trajectories of particles of resin suspended in solution. Additionally, he speculated on the actual trajectory of these particles, and noted "...how near the mathematicians are to the truth, in refusing, by logical instinct, to admit the pretended geometrical demonstrations, which are regarded as experimental evidence for the existence of a tangent at each point of a curve" [8]. In other words, a discrete random walk is only a partial sample of the "true" mathematical path, an idea which foreshadows the continuous limit of Brownian motion, known as the Wiener process.

The Wiener process is the continuous limit of a Gaussian random walk, and though it does serve as a model for physical Brownian motion, it has some peculiarities. The process has no inertia, it is nowhere differentiable, it is fractal-like and scale-free, and it maintains the issues of Gaussian diffusion that allow for arbitrarily large velocities. So, while the Wiener process can serve as natural model for diffusion, it has a more important role as a mathematical tool and as a foundation for many other stochastic processes. More complex processes can be constructed using the Wiener process (e.g. the Ornstein–Uhlenbeck process, which is in a sense a Wiener process with inertia), and it is omnipresent in stochastic equations (e.g. Langevin equations and Feynman-Kac formulas can be written using the Wiener process).

1.2. Lévy Statistics and Extreme Events

In the previous section we discussed the broad applicability of Brownian motion and the Wiener process, both of which are intrinsically Gaussian random processes. However, despite their effectiveness for modeling certain systems, they are often applied beyond their strict regime of applicability. The central limit

theorem that underpins the prevalence of Gaussian statistics is not applicable in the random variables with correlations, non-stationary distributions, or heavy tailed distributions, and applying it in these situations can be highly problematic, though it remains a common practice.

One such example is in the Black-Scholes model [9] used in the pricing of financial derivatives. The model assumes that the log of stock prices fluctuate as stochastic Wiener processes with a fixed variance. That this model is heavily used, even today, may be somewhat surprising considering that stock prices often have strong correlations, are non-stationary, and are heavy tailed [10]. As such, the Black-Scholes pricing model has sometimes been considered by financial experts as putting "the wrong numbers in the wrong formula to get the right price" [11].

On the other hand, the Black-Scholes model may not accurately predict the "right price" after all. The author of the book *The Black Swan* [12], where the title refers to events that are rare and difficult to predict, was highly critical of the uses of Black-Scholes models and Gaussian risk management by financial institutions, and wrote "...the financial ecology is swelling into gigantic, incestuous, bureaucratic banks (often Gaussianized in their risk measurement)—when one falls, they all fall." This portent of disaster was quite timely considering the devastating 2007-2008 financial crisis that began in the months after the book was published. Other misapplications of Gaussian statistics exist, typically with the error being an underestimate of the importance of large but rare events.

With the potential failures of Gaussian statistics in mind, it is important to consider alternative models that can account for extreme events. While there is no perfect model, extreme events are commonly associated with probability distributions with "heavy" power-law tails. Lévy processes (specifically, *stable*

4

Lévy processes [13, 14, 15]) are a natural extension to Wiener processes to incorporate heavy-tailed statistics.

Lévy statistics have found a wide range of uses, including modeling atmospheric dynamics, heart beat rhythm [16], stellar dynamics, and even quantum decay [17]. One particularly interesting result is that Lévy flights have been shown to be an optimal foraging strategy under certain conditions [18], partially explaining how the observed flight patterns of the albatross can be modeled with Lévy flights [19], random walks with heavy tailed step distributions. Based on the examples discussed here, "rare" and "extreme" events of Lévy statistics are actually quite prevalent.

1.3. Anomalous Diffusion of Laser Cooled Atoms

Lévy processes also play a particularly important role in the understanding of the diffusion of cold atoms, though as we will see, detailed investigations of anomalous diffusion did not occur until relatively late in the development of laser cooling systems. In this section we overview some of the history of laser-cooling systems, highlighting studies involving diffusive properties of these systems, as well as how Lévy statistics are an effective description for more recently studied regimes of anomalous diffusion.

One of the early and most effective cooling systems developed for neutral atoms is "optical molasses" [20, 21]. This cooling system relies on both radiation pressure [22], where atoms interacting with light experience a force in the direction the light is propagating, as well as the Doppler effect, where an atom in motion will observe a shift in frequency of incident light. A cooling effect is achieved by setting up counter-propagating beams that are red-detuned from the atomic

resonance of the atom species to be cooled. If an atom in motion is interacting with both beams, the beam that opposes the atom's motion will be blue-shifted into resonance and therefore interact more strongly. This in turn increases the radiation pressure from the beam that opposes the atom's motion, effectively damping the motion. As damping the motion reduces the kinetic energy, this can be reasonably be considered a cooling system for classical definitions of temperature[1]. This general cooling effect is known as Doppler cooling, while the setup with counter-propagating cooling beams is referred to as optical molasses as it damps motion in all directions akin to the viscous damping in molasses.

The motion of atoms in optical molasses does not damp completely away, however, as photons emitted through spontaneous emission give the atom random momentum kicks, ultimately driving diffusive motion. For simple Doppler cooling, the competing cooling effect and the heating effect of spontaneous emission reach an equilibrium temperature known as the Doppler temperature [23]. This model suggests that a simple diffusion model is appropriate to describe atomic motion in optical molasses, and that the atomic motion can be used to characterize laser cooled atoms, especially for the purpose of measuring temperatures [21].

However, it became apparent that such simple models were inadequate for atomic species with hyperfine structure, when temperatures were measured far below the Doppler temperature. This led to the development of more sophisticated quantum models to explain the observed sub-Doppler cooling effect. The primary effect leading to the unexpectedly cold temperatures was found to be Sisyphus cooling, which is also the cooling force considered in this dissertation (see Chapter IV for more details on Sisyphus cooling). The more detailed quantum

---

[1]In the case of a single atom, various definitions of temperature may not be well defined, but these "cooling" systems still damp the atom's motion.

models also led to the development of new cooling techniques such as velocity-selective coherent population trapping [24]. Interestingly, the cooling through velocity-selective coherent population trapping is a diffusive process rather than a friction force. The technique involves optically pumping slowly moving atoms into a dark state, where the atoms effectively decouple from the driving lasers due to destructive quantum interference. While in the dark state the atoms no longer experience heating due to spontaneous emission, and so remain moving slowly. Atoms which are not in a dark state still experience the typical cooling and heating effects, until their momentum happens to be small enough such that the atom is pumped into the dark state, where atoms tend to stick. Because the dark state corresponds to small atomic momentum, the atoms are effectively cooled. In addition to being an interesting diffusion-based cooling mechanism, this also led to the first use of Lévy statistics in the description of laser-cooled atoms, as the time that the atom remains in the dark state is random, with a power-law-tailed distribution [25]. Significant work has been done on analyzing the Lévy statistics for this type of cooling mechanism [26]. Our focus here, however, will be on Sisyphus cooling.

The full quantum models for Sisyphus cooling can be complex [27, 28, 29]. As semiclassical models [30, 31] were developed, they were placed on a rigorous footing through both perturbative calculations [32] as well as through numerical simulations [31]. Semiclassical approximations treat the laser fields and atomic position classically, while maintaining the quantum nature of the atomic state. These models were often simplified further by a process of adiabatic elimination, where changes in the atomic state are assumed to occur much more quickly than motion of the atom, which allowed for investigation of the diffusive behavior of

Sisyphus cooling. These semiclassical models were used to develop a detailed theory of optical molasses, including the velocity distribution (a particular heavy-tailed distribution known as the Tsallis distribution [33]), as well as a more thorough investigation of the spatial diffusion properties [30].

These earlier studies still primarily focused on regimes of normal diffusion; the anomalous diffusion that we will discuss shortly was referred to as *décrochage*, or disintegration, and was generally considered undesirable as the density of confined atoms would rapidly drop in these regimes. The rational here was due to a major motivation behind laser cooling at the time, which was to reach Bose-Einstein condensation. This required high atomic densities through the use of magneto optical traps [34]; the dense atoms would then be "boiled" off through evaporative cooling techniques [35] to reach condensation, so high atom counts were critical. However, it was found in Ref. [30] the diffusion constant diverges in the regime of large detuning and low intensity laser fields, and upon closer investigation it was found that this divergence was a transition to a regime that could be described by anomalous diffusion and the onset of Lévy flights [31]. Unlike the velocity-selective coherent population trapping, these Lévy flights[2] occur in space, leading to anomalous diffusion. While regular diffusion has the property that the standard deviation of the diffusing particles will grow as $\sqrt{t}$, anomalous diffusion refers to any process where the growing width of an ensemble distribution scales as t^μ for some $\mu \neq 1/2$.

While the onset of anomalous diffusion had been observed in Sisyphus cooling [30], more statistical signatures of anomalous diffusion and Lévy-flight

---

[2]Technically these are known as Lévy walks, which specify that the particle velocity is always finite. The distinction is not completely trivial (see [36]), but Lévy processes can be used to describe Lévy flight behavior in the same sense that a Gaussian diffusion can be used to describe physical processes even though Gaussian diffusion has no maximum velocity.

statistics were observed by studying the motion of a single trapped ion [37]. These results were in qualitative agreement with the theoretical work in Ref. [31], with the onset of anomalous diffusion occurring for shallow optical potential depth, though the optical potential depth for the onset of the diffusion did not match theoretical values.

More recent experiments used clouds of atoms to investigate anomalous diffusion in Sisyphus cooling [38], which again found the onset of anomalous diffusion for shallow optical potentials. However, this cloud based experiment characterized the anomalous diffusion exponent using multiple methods, with results that suggested the presence of multiple distinct anomalous diffusion exponents. For simple models of anomalous diffusion, the parameter that describes the power law growth of the standard deviation should match the parameter that describes the power-law tail of the position distribution. However, this experiment found disagreement between these measured exponents. Theoretical work studying the onset of anomalous diffusion in more depth [39] also found that the anomalous diffusion should be characterized by a single universal exponent, indicating that current models may be missing some key component. Moreover, a phase transition to a regime where the anomalous diffusion exponent is fixed at $\alpha = 3/2$, known as the Obukov-Richardson phase, has been predicted but not yet observed experimentally. This phase corresponds to integrated Brownian motion, and is in the limit where the cooling force is negligible but the heating effect due to spontaneous emission is still present.

It has been proposed that some of these differences between theory and experiment may be due to loss effects via inelastic collisions and molecule formation [39], especially since the losses themselves may contribute to the

appearance of anomalous diffusion [40]. To circumvent such potential issues, and to enable alternative experimental probes of Lévy dynamics of laser-cooled atoms, in this dissertation I investigate the possibility of using experiments with *single* atoms to probe the stochastic dynamics. Single-atom experiments would open up the possibility of studying statistics not accessible to experiments with atomic clouds, such as boundary-crossing and escape statistics. Measurements of these statistics would complement existing measurements, which have been limited to characterizations of the scaling behavior of ensemble densities.

1.4. Organization of this Dissertation

This dissertation is organized as follows: Chapter II introduces some basic concepts of stochastic processes, including the Wiener and Lévy processes and some important statistics. Building on this foundation, Chapter III explores conditioned stochastic processes, first introducing the well known Brownian bridge in Section 3.1, before focusing on the less common Lévy bridge in Section 3.2. The new results of this dissertation begin in Section 3.2.3, where I find that a conditioned displacement of the final state for the Lévy bridge experiences a bifurcation when the conditioned length exceeds a certain threshold. The displacement at which this bifurcation occurs defines a novel length scale which naturally characterizes many features of the conditioned Lévy process. Furthermore, I show that it has interesting uses such as detecting extreme events (Section 3.2.5), generating Lévy bridges (Section 3.2.6), explaining properties of first passage times (Section 3.2.7), and parameter inference of a Lévy process from data (Section 3.2.8). Finally, Chapter IV reviews existing theory on the onset of anomalous diffusion in Sisyphus cooling. Original work begins in Section 4.5 with

a formula that characterizes the extent of power-law scaling which is useful in the design of experiments studying Lévy behavior with laser-cooled atoms. New results continue in Section 4.6, where through simulations I establish the existence of a peak in the boundary-crossing statistics that is connected to the underlying Lévy-type behavior, and derive scaling relations for its extent. These results serve as a starting point for development of future single-atom experiments.

CHAPTER II

STOCHASTIC PROCESSES

This chapter reviews basic, well-known background material for handling stochastic processes, which serves as a foundation for understanding the conditioned processes investigated in Chapter III, and for understanding the dynamics of laser-cooled atoms in Chapter IV. There are a variety of good general references for this material [14, 15, 41, 42, 43], as well as some with specific focus on Lévy processes [16, 44, 45, 46].

Understanding the transport properties of cold atoms in Sisyphus cooling requires modeling of random processes at several different levels. First, the underlying physical process involves spontaneous emission, an inherently stochastic process that gives random momentum kicks to the cooled atoms. Next, the motion of the atom itself is the time integral of the momentum, and so is a sum of all the incremental random kicks. Finally, the overall diffusive process incorporates all possible trajectories each atom could take. These levels are each described by different machineries in the language of stochastic processes: The underlying physical process is modeled as an Itō stochastic differential equation for the momentum; the motion of the atom is the stochastic integral of the momentum equation; and the aggregate diffusive process is the evolution of a probability distribution through a master equation. And so, in this chapter we discuss some of the essential aspects of stochastic processes necessary for the understanding of cold-atom transport.

The subject of stochastic processes is a broad one. It covers the analysis of random systems, lending itself well to describing many physical systems like

molecular diffusion and electrical noise, as well as less physical systems like fluctuations in the stock market and animal birth-death rates [14]. Despite this breadth in application, there are only a few fundamental random processes from which more complex stochastic systems can be constructed. In Section 2.1 we will discuss the Wiener process, an idealized mathematical description of Brownian motion. It is the one of the most widely used stochastic processes due to a universality among stochastic processes, where seemingly disparate systems ultimately have identical limiting behavior. The source of the mechanism that causes this universality for Wiener process is the central limit theorem, one of the most celebrated results of probability theory. We will also discuss the basics of stochastic differential equations, some important examples, as well as several statistics for the Wiener process. Next, in Section 2.2 we will discuss Lévy processes. These can be thought of as an extension to the Wiener processes by relaxing a continuity requirement to allow for systems involving rare and dramatic events, which appear as discontinuous "jumps" in the system's evolution. We will also discuss a generalization of the central limit theorem that naturally leads a certain class of Lévy processes known as the α-stable Lévy processes. We will show that these processes are particularly important for models of anomalous diffusion, despite their apparently pathological behaviors like discontinuous sample paths which lead to the possibility for Lévy processes to jump across an absorbing region without being absorbed.

2.1. Wiener Processes

The Wiener process is a continuous-time random walk that is an idealized mathematical description of Brownian motion. While true physical processes

differ somewhat from the Wiener processes, the Wiener process is a convenient mathematical object, and remains an excellent approximation for many systems exhibiting Brownian motion. To show why this is, we will start by deriving the central limit theorem (Section 2.1.1), which underpins the the importance of the Wiener process. Following this we will give an informal definition of the Wiener process as a limit of a discrete random walk (Section 2.1.2), as well some basic properties (Section 2.1.3). Finally, we will introduce several tools of stochastic calculus, including stochastic differential equations (Section 2.1.5), master equations (Section 2.1.6), methods for sampling stochastic processes (Section 2.1.7), and certain key statistics for the Wiener process (Section 2.1.8).

2.1.1. Central Limit Theorem

One of the most interesting and useful results in the study of probability and statistics is the central limit theorem (CLT), where the sum of many random numbers from an unknown distribution has an asymptotically Gaussian distribution. This somewhat counter-intuitive result, where the combination of minimally described random numbers results in a precise statistical description, is broadly applicable. It explains quite generally why so many measured statistics appear Gaussian and is an important component of the explanation for regular diffusive behavior in cold atoms discussed in both the introduction and in Section 4.3.

While there are many ways of presenting the CLT, we will show a brief proof using a random walk representation similar to proofs in Refs. [43, 47], which will be a useful formalism for motivating the Wiener process in the following section. A simple, one-dimensional random walk is a series of random steps Δx_i drawn from a

distribution $f_{\Delta x}$, such that after n steps the random walker has a position given by $x_n = \sum_{i=1}^{n} \Delta x_i$.

Using this notation, a basic statement of the CLT is as follows: If one takes n independent random variables Δx_i drawn from an *arbitrary* distribution $f_{\Delta x}$ with a known mean $\mu = \langle \Delta x \rangle$ and variance $\sigma^2 = \langle (\Delta x - \mu)^2 \rangle$, the sum of these numbers, x_n, has the asymptotic, *specific* distribution given by the Gaussian

$$f_{x_n}(x) = \frac{1}{\sqrt{2\pi n \sigma^2}} e^{-(x-\mu)^2/2n\sigma^2}. \tag{2.1}$$

The proof for this is simplified if we use the shifted and scaled variables

$$\Delta z_i = \frac{(\Delta x_i - \mu)}{\sqrt{n}\sigma} \tag{2.2}$$

which have zero mean and variance $1/n$. The motivation for this choice will become clear later, but we are essentially scaling out the expected linear growth of variance in n. Using these variables Δz_i, the analogous sum to x_n is

$$z_n = \sum_{i=1}^{n} \Delta z_i. \tag{2.3}$$

As this is the sum of independent random variables, the probability distribution will be given by an iterated convolution of the individual distributions. Since each step has the same distribution, the probability distribution for z_n is given by

$$f_{z_n}(x) = [f_{\Delta z} * f_{\Delta z} * \ldots * f_{\Delta z}](x), \tag{2.4}$$

15

where we have $n-1$ convolutions. From the convolution theorem, we can write the Fourier transform for f_{z_n} as a product of Fourier transforms of $f_{\Delta x}$, which is

$$\mathcal{F}[f_{z_n}](k) = \mathcal{F}[f_{\Delta z}](k)\mathcal{F}[f_{\Delta z}](k)\ldots\mathcal{F}[f_{\Delta z}](k), \tag{2.5}$$

where we have defined the Fourier transform by

$$\mathcal{F}[f_{\Delta z}](k) = \int_{-\infty}^{\infty} dx\, f_{\Delta z}(x)e^{ikx}. \tag{2.6}$$

We now can expand the exponential as a sum to find

$$\mathcal{F}[f_{\Delta z}](k) = \sum_{m=0}^{\infty} \frac{1}{m!}(ik)^m \int_{-\infty}^{\infty} dx\, f_{\Delta z}(x)x^m, \tag{2.7}$$

where we see that the integral is just the mth moment of $f_{\Delta z}$. Since the mean is zero and the variance is $1/n$ this expansion can be written as

$$\mathcal{F}[f_{\Delta z}](k) = 1 - \frac{k^2}{2n} + \mathcal{O}(k^3). \tag{2.8}$$

Putting this into our expression in Eq. (2.5) gives us

$$\mathcal{F}[f_{z_n}](k) = \left(1 - \frac{k^2}{2n} + \mathcal{O}(k^3)\right)^n. \tag{2.9}$$

We now can see that terms of order k^m are proportional to $n^{-m/2}$, which justifies keeping only to lowest order in n in the limit $n \to \infty$. Then, in this limit we can use the relation $e^x = \lim_{n\to\infty}(1 + x/n)^n$ to find

$$\mathcal{F}[f_{z_n}](k) = e^{-k^2/2}, \tag{2.10}$$

16

which through the inverse Fourier transform gives us

$$f_{z_n}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{2.11}$$

Finally, by unscaling x_n as in Eq. (2.2) to we find the desired result in Eq. (2.1).

2.1.2. Definitions

The CLT gives good reason to expect many systems to exhibit Gaussian behavior above some length or time scale, even if the underlying process is not fundamentally Gaussian. The Wiener process, however, represents an underlying process that explicitly consists only of infinitesimal Gaussian increments at all scales. They can be used as an ideal mathematical object that is conveniently manipulated to find exact results for many statistics, or as a particularly useful model for many real-world systems when the true underlying non-Gaussian behavior is experimentally inaccessible or irrelevant.

Informally, a Wiener process $W(t)$ can be defined as the continuous time analog of a discrete Gaussian random walk with n steps occurring at time intervals $\Delta t = t/n$. The final position of such a walk is given by

$$x_n = \sum_i^n \Delta x_i, \tag{2.12}$$

where the Δx_i are independently drawn from a Gaussian probability distribution

$$f_{\Delta x_i}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}. \tag{2.13}$$

As the sum of independent Gaussian random variables is also a Gaussian random variable, we have the distribution for final position given by

$$f_{x_n}(x) = \frac{1}{\sqrt{2\pi n \sigma^2}} e^{-x^2/2n\sigma^2}. \tag{2.14}$$

To properly translate this discrete walk into a continuous form, we would like the final position x_n to correspond to the final position $W(t)$ as we take the limit $n \to \infty$. That is, we want

$$f_{W(t)}(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} \tag{2.15}$$

to match Eq. (2.14), as $t \propto n$. However, as we increase n we see that we must scale σ such that the distribution for f_{x_n} does not diverge and maintains the appropriate scaling behavior. The choice for the variance of the individual steps Δx_i that maintains Eq. (2.14) is $\sigma^2 = \Delta t$. Taking this assumption we now can define the Wiener process as the limit

$$W(t) = \lim_{n \to \infty} \sum_{i=1}^{n} \Delta x_i. \tag{2.16}$$

Note that similarly, we have $W(t) = \lim_{n \to \infty} \sum_{i=1}^{n} \sigma^2$. Notationally we can write Eq. (2.16) as a stochastic integral

$$W(t) = \int_0^t dW(s), \tag{2.17}$$

where dW is an infinitesimal increment of the Wiener process. The Wiener increment dW can be thought of as a Gaussian random variable with a variance dt. This is because the variance of $W(t)$ must grow linearly in time to be consistent with the CLT and with $f_{W(t)}$, so each of the increments dW must

18

accordingly have variance dt. While we have not proven this explicitly, this relation is known as Itō's lemma [15] and can be written as

$$dW^2 = dt. \tag{2.18}$$

Using the independence of Wiener increments, as well as Itō's lemma, we can write the variance of the Wiener process as

$$\langle W^2(t) \rangle = \left\langle \int_0^t dW(s) \int_0^t dW(s') \right\rangle = \left\langle \int_0^t dW^2(s) \right\rangle = \int_0^t dt = t, \tag{2.19}$$

which demonstrates some convenient manipulations commonly used to simplify stochastic integrals. Note that we are using the angle brackets to indicate an ensemble average.

One remaining tool for Itō calculus is a generalization of the chain rule. For a function $y(x,t)$ the typical chain rule can be written as

$$dy = \left(\frac{\partial y}{\partial x} \right) dx + \left(\frac{\partial y}{\partial t} \right) dt \tag{2.20}$$

which can be thought of as a Taylor expansion of $y(x + dx, t + dt)$, dropping terms higher than first order. However, if dx contains a term proportional to a Wiener increment, we can see that we must keep terms containing dW^2 since they are actually first order in dt. This is done simply by adding the next term from the Taylor expansion, giving us the Itō chain rule

$$dy = \left(\frac{\partial y}{\partial x} \right) dx + \left(\frac{\partial y}{\partial t} \right) dt + \frac{1}{2} \left(\frac{\partial^2 y}{\partial x^2} \right) dx^2. \tag{2.21}$$

This covers a basic introduction to Wiener processes, which are a natural tool due to the CLT. In the next section we discuss some essential properties of the Wiener process.

2.1.3. Fundamental Properties

The Wiener process has several fundamental mathematical properties, which we briefly discuss here; more detailed explanations can be found in Ref. [13, 44, 47].

1. **Stationary Increments.** Also known as time-homogeneity, this states that that all Wiener process increments only depend on the current state and not on any other variables. That is, all elements of $\{W(t_0 + t) - W(t_0) : t \geq 0\}$ are independent of t_0.

2. **Independence of increments.** For any choice of $n \geq 1$ and $0 \leq t_0 < t_1 < \ldots < t_n$, all increments from $W(t_0), W(t_1) - W(t_0), \ldots, W(t_n) - W(t_{n-1})$ must be statistically independent.

3. **Infinite Divisibility.** Processes that have both stationary and independent increments are called infinitely divisible, so this property also holds for the Wiener process. As the term indicates, the process $W(t)$ can be split into a sum of N independent identically distributed random variables.

4. **Martingale.** The martingale property is the requirement that for all $t > t_0 > 0$, the expectation value $\langle W(t - t_0) \rangle = W(t_0)$. This is essentially a statement that there is no deterministic drift.

5. **Continuity.** The Wiener process must satisfy the standard continuity requirement: For all t and for all $\epsilon > 0$ there exists a $\delta > 0$ such that $|W(t + \delta) - W(t)| < \epsilon$.

2.1.4. Qualitative Behavior

In addition to the formal properties, there are many peculiar qualitative behaviors of the Wiener process that are important to understand.

Sample Path Behavior. As stochastic processes depend on random elements with multiple possible outcomes, there are many solutions for a given stochastic process. Each possible solution to a stochastic process is known as a *sample path*. As an example we have plotted 10 sample paths for the Wiener process in Fig. 2.1. As we can see, the individual sample paths can vary greatly, and consistent results will only occur only in distributions of possible solutions. However, a single sample path is still useful for a qualitative understanding of the motion, as the sample paths can actually represent the microscopic behavior of a physical process. Furthermore, for real-world data an ensemble is not always available, and the information available from analyzing a single sample path can still be valuable.

Boundary Crossing Behavior. The fractal-like nature of the Wiener process has some unusual properties when considering the crossing of a boundary. First, if a Wiener process crosses some threshold, one might naïvely assert that there must then be a well defined first crossing point. However, this interpretation has issues for a Wiener process considered at a finite resolution. While one resolution may have a crossing time at t_1, any of the possible refinements (i.e. higher resolution sample paths that hit the same points and have consistent
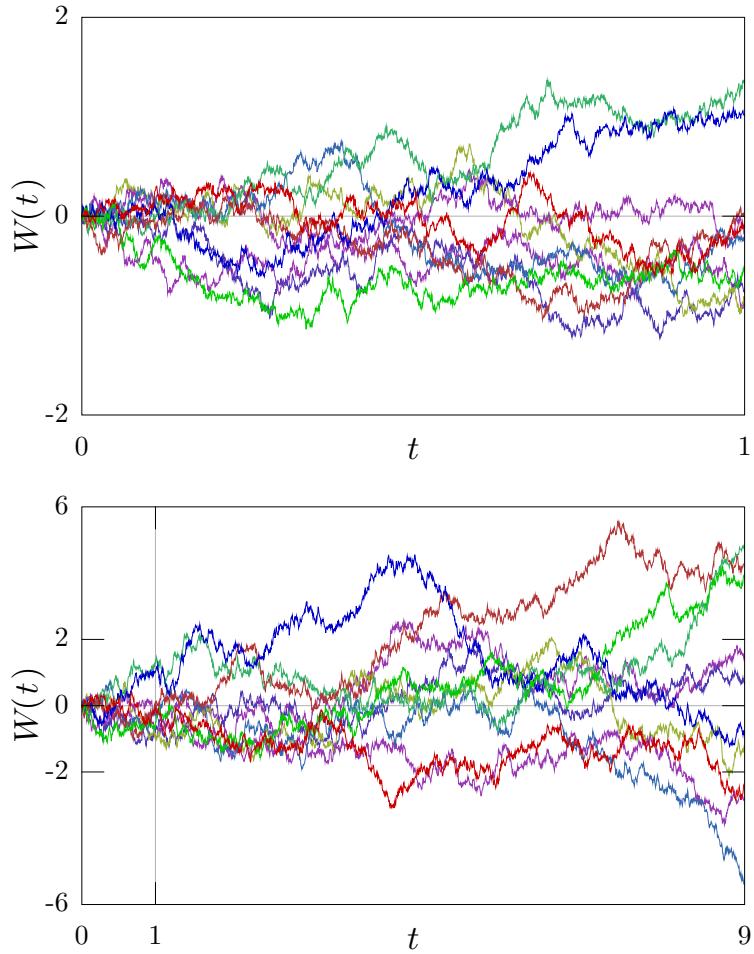
FIGURE 2.1. Plots of 10 sample paths of a Wiener process. Both plots show
the same sample paths (indicated by colors), while scale of each plot is chosen to
indicate the qualitative self-similarity at at different scales.

statistics) of this process may have different crossing times at a distinct time

$t_2 \neq t_1$ or even multiple crossing times. In fact, a true Wiener process that crosses

a boundary at least once will (almost surely) have an infinite number of crossings

in the vicinity, and any given crossing for a finite resolution sample path will

inevitably have an arbitrarily large number of nearby crossings when the sample

paths are refined.

Fortunately, the assertion that there must be a well defined first crossing

point is actually correct for the true Wiener process [48], which allows us to

define well behaved distributions like the first passage time, which we explore in Section 2.1.8. Still, there are some cases that are somewhat problematic. For example, an *excursion* is defined as a process that starts and ends at the origin. For a Wiener process that starts at the origin, the first passage time at the origin is simply $t = 0$ since every crossing will have an infinite number of nearby crossings. This can be problematic because excursions can play a useful role in analyzing random walk behavior (see for instance Ref. [39], where this issue is mitigated by considering boundaries with a size $\epsilon > 0$ and take the limit $\epsilon \to 0$ as needed).

Fractal and Scale Free Behavior. Fractals are a broad class of self-similar geometric objects, where the object at one scale has similarities to itself at other scales. Since the Wiener process is fundamentally random, it is incredibly unlikely for a Wiener sample path to have any *exact* self-similarity. However, the Wiener process does have statistical self-similarity which firmly places it as an example of a random fractal. This self-similar behavior can be seen in Fig. 2.1, where each plot shows the same set of sample paths, but shown over different interval lengths and at different scales. Following the variance scaling of the Wiener process, the time axis is scaled by s while the length axis is scaled by $\sqrt{s}$, where $s = 3$. Thus, with the appropriate scaling the qualitative appearance of these plots is similar, even if the individual trajectories are quite different.

2.1.5. Stochastic Differential Equations

As we have seen, the Wiener process is a useful description of many random processes due to the CLT. However, it also is useful in describing a class of random systems beyond basic Brownian motion through appropriate stochastic differential

equations. Thus the Wiener process is not only fundamental because of the CLT, it also is a key building block for many other random processes.

The Itō stochastic differential equation, given by

$$dx = h(x,t)dt + g(x,t)dW, \tag{2.22}$$

where $h(x,t)$ is the drift coefficient and $g(x,t)$ is the diffusion coefficient, describes the evolution of a process $x(t)$ that can be computed through a stochastic integral. As is true for all SDEs, there are many possible solutions to this equation.

We will now discuss two important examples that can be written in the form of Eq. (2.22) that are useful for the analysis and interpretation of the diffusion in cold atoms in Chapter IV.

2.1.5.1. Ornstein–Uhlenbeck Process

The Ornstein–Uhlenbeck process (a Wiener process plus linear damping) is a common model for the momentum of a massive particle that results in physically observed Brownian motion. As such, the integrated Ornstein–Uhlenbeck process asymptotically approaches a Wiener process, but the particles in this model have inertia and do not instantaneously change direction as happens for the true Wiener process. It is also a successful model for some regimes cold atom diffusion, so it is a natural system to use for comparison to diffusive regimes that have non-linear damping.

The Ornstein–Uhlenbeck process (dp) and the corresponding particle position (x) is described by the pair of stochastic differential equations

$$dp = -\gamma p \, dt + \sigma_p \, dW, \tag{2.23}$$

$$dx = \frac{p}{m} \, dt. \tag{2.24}$$

where γ is a damping constant and σ_p is a momentum diffusion constant. For convenience we use the scaled variables $t \to t/\gamma$, $p \to pm$, along with the definition $\sigma = \sigma_p/m\sqrt{\gamma}$, which gives the dimensionless equations

$$dp = -p \, dt + \sigma \, dW, \tag{2.25}$$

$$dx = pdt. \tag{2.26}$$

The typical method to solve this system is to notice that the homogeneous form of Eq. (2.25) (i.e. $\sigma = 0$) has the solution

$$p(t) = p_0 e^{-t}, \tag{2.27}$$

which in turn motivates the transformation

$$y = pe^t. \tag{2.28}$$

To find the increment dy, we apply the Itō chain rule in Eq. (2.21) resulting in

$$dy = \left(e^t\right)\left(-p \, dt + \sigma \, dW\right) + \left(pe^t\right) dt \tag{2.29}$$

$$= e^t \sigma \, dW. \tag{2.30}$$

We can now integrate this to find $y(t)$,

$$y(t) = p_0 + \sigma \int_0^t e^s dW(s), \tag{2.31}$$

where p_0 is an integration constant. Finally, we transform back via Eq. (2.28) to find

$$p(t) = p_0 e^{-t} + \sigma \int_0^t e^{s-t} dW(s). \tag{2.32}$$

This has a nice form with a transient $p_0 e^{-t}$ that will vanish at long times, and a second term which is simply the integral of a Gaussian random variable (and thus is also a Gaussian with a variance that depends on t). As the mean is clearly zero, we just need to calculate the variance of this second term to fully characterize it, which can by found by

$$\left\langle \left(\sigma \int_0^t e^{s-t} dW(s) \right)^2 \right\rangle = \left(\sigma \int_0^t e^{s-t} dW(s) \right) \left(\sigma \int_0^t e^{s'-t} dW(s') \right) \tag{2.33}$$

$$\left\langle \left(\sigma \int_0^t e^{s-t} dW(s) \right)^2 \right\rangle = \sigma^2 e^{-2t} \int_0^t e^{2s} ds \tag{2.34}$$

$$\left\langle \left(\sigma \int_0^t e^{s-t} dW(s) \right)^2 \right\rangle = \frac{\sigma^2}{2} \left(1 - e^{-2t} \right). \tag{2.35}$$

While this characterizes the momentum, we still have not solved for the position evolution. To find the position as a function of time, we must integrate Eq. (2.26):

$$x(t) = x_0 + \int_0^t p_0 e^{-s} ds + \sigma \int_0^t ds' e^{-s'} \int_0^{s'} e^s dW(s). \tag{2.36}$$

The second term is easy to integrate giving

$$\int_0^t p_0 e^{-s} ds = p_0 (1 - e^{-t}). \tag{2.37}$$

26

Since the third term in Eq. (2.36) is a Gaussian random variable with zero mean, we only need to compute the variance σ_x^2 which is given by

$$\sigma_x^2 = \sigma^2 \left\langle \left[\int_0^t ds_1 e^{-s_1} \int_0^{s_1} e^{s_2} dW(s_2) \right] \left[\int_0^t ds_3 e^{-s_3} \int_0^{s_3} e^{s_4} dW(s_4) \right] \right\rangle. \qquad (2.38)$$

Since the stochastic increments dW are independent random variables, all contributions to the integral will vanish in the average unless the increments $dW(s_2)$ and $dW(s_4)$ are the same increment. This allows us to rewrite this integral with the replacement $s_4 \to s_2$ giving us

$$\sigma_x^2 = \sigma^2 \int_0^t ds_1 \int_0^t ds_3 \int_0^{\min(s_1,s_3)} e^{(2s_2 - s_1 - s_3)} ds_2, \qquad (2.39)$$

where we have dropped all uncorrelated products of dW. Notice that the limit for the s_2 integral is the minimum of s_1 and s_3, as the integration limits must be truncated to the range where each $dW(s_2)$ has a complementary $dW(s_4)$. To continue simplifying this integral, it helps to think about it as an integral over a subsection of a $t \times t \times t$ cube, where we only include the volume of a "quarter-pyramid"; the subsection of the cube selected by the integration limits is equivalent to the following integral:

$$\sigma_x^2 = \sigma^2 \int_0^t ds_2 \int_{s_2}^t ds_3 \int_{s_2}^t ds_1 e^{2s_2 - s_1 - s_3}. \qquad (2.40)$$

Evaluating this, we have

$$\sigma_x^2 = \sigma^2 \left(t - \frac{3}{2} + 2e^{-t} - \frac{1}{2} e^{-2t} \right). \qquad (2.41)$$

27

This gives us the expected asymptotic linear growth in the variance as well as an exponentially damped transient. However, it is interesting that there are two different transients proportional to e^{-t} and to e^{-2t}.

Finally, we can combine the results from Eqs. (2.37) and (2.41) into the probability density

$$f(x,t) = \frac{1}{\sqrt{2\pi(\sigma_x(t))^2}}\exp\left(-\frac{(x - x_0 - p_0(1 - e^{-t}))^2}{2(\sigma_x(t))^2}\right), \qquad (2.42)$$

which is just a Gaussian with a particular time-dependent mean and variance.

2.1.5.2. Integrated Brownian Motion

One other commonly used process that will be relevant for later analysis is integrated Brownian motion. This is equivalent to the Ornstein–Uhlenbeck Process with a vanishing damping coefficient, and so we have the equations

$$dp = \sigma\, dW, \qquad (2.43)$$

$$dx = pdt, \qquad (2.44)$$

which we can see leads to

$$dx = \sigma \int_0^t W(s)dt. \qquad (2.45)$$

This has the well known variance scaling of $\langle x^2 \rangle \sim t^3$, which can be seen via

$$x(t) = \sigma \int_{s=0}^{t} W(s)ds \tag{2.46}$$

$$x(t) = \sigma \int_{s=0}^{t} \int_{u=0}^{s} dW(u)ds \tag{2.47}$$

$$x(t) = \sigma \int_{u=0}^{t} dW(u) \int_{s=u}^{t} ds \tag{2.48}$$

$$x(t) = \sigma \int_{u=0}^{t} (t-u)dW(u), \tag{2.49}$$

where we have changed the order of integration. Since we have a sum of Gaussian random variables, we only need to compute the variance

$$\langle x(t)^2 \rangle = \sigma^2 \int_0^t (t-u)dW(u) \int_0^t (t-u)dW(u) \tag{2.50}$$

$$\langle x(t)^2 \rangle = \sigma^2 \int_0^t (t-u)^2 dt \tag{2.51}$$

$$\langle x(t)^2 \rangle = \frac{\sigma^2}{3}t^3, \tag{2.52}$$

which gives us the expected scaling behavior.

2.1.6. Master Equations

Master equations are differential equations that describe the evolution of a probability density. While the Itō SDEs can be thought of as describing the evolution of an individual particle, master equations can be used to describe either the probability of different possible outcomes for a single particle, or evolution of the density of an ensemble of particles. In both cases the trajectory of individual particles is not readily apparent and so this picture is more suited as a macroscopic description of a system.

The equivalent master equation to the Itō SDE in Eq. (2.22) is known as the Fokker–Planck equation

$$\partial_t \rho(x,t) = -\partial_x[h(x,t)\rho(x,t)] + \frac{1}{2}\partial_x^2[D(x,t)\rho(x,t)], \qquad (2.53)$$

where $D(x,t) = g(x,t)^2$ is the diffusion coefficient.

As a simple example, we can look at the solution in the case of the pure Wiener process, where we have $h(x,t) = 0$, $D(x,t) = \sigma^2$. The Fokker Planck equation then reduces to the heat equation

$$\frac{\partial}{\partial t}\rho(x,t) = \frac{\sigma^2}{2}\frac{\partial}{\partial x^2}\rho(x,t). \qquad (2.54)$$

The propagator [the solution given $\rho(x,t=0) = \delta(x)$] is just the Gaussian kernel

$$\rho_G(x,t) = \frac{1}{\sqrt{2t\sigma^2}}e^{-x^2/2t\sigma^2} \qquad (2.55)$$

which allows us to express the solution for an arbitrary initial distribution $\rho(x,t=0)$ as a convolution:

$$\rho(x,t) = \int_{-\infty}^{\infty} \rho_G(x-x',t)\rho(x,t=0)dx'. \qquad (2.56)$$

2.1.7. Sampling and Simulation

This section outlines standard numerical methods used for integrating stochastic differential equations involving a Wiener process. These methods are used for producing Brownian bridge statistics in Chapter III, as well as the diffusion simulations of the Sisyphus cooling force in Chapter IV. Methods of

numerically solving ODEs often involve discretizing a continuous process into finite time steps Δt and incrementally solving for the state at each discrete time $n\Delta t$. These methods can be employed for the numerical simulations of SDEs, but there are some additional challenges compared to simulations for ODEs. First, since there are a multitude of solutions, to properly sample an SDE we must collect a series of sample paths until the statistics settle and the behavior is well characterized. This can be substantially more computationally expensive than simulations for ODEs. Second, the addition of Gaussian noise at each discrete time step leads to poorer performance than similar algorithms for ODEs. Finally, certain numerical algorithms for SDEs only have *weak convergence*, where the simulation only has correct results for ensemble averages, while the individual paths may not reflect the true process. If the behavior of the individual paths is important (as is the case for statistics like first passage times in Section 2.1.8), then we require the algorithm to have *strong convergence*.

Despite these additional concerns, some algorithms for SDEs are very similar to algorithms for ODEs. To numerically simulate an Itō SDE of the form

$$dx = h(x,t)dt + g(x,t)dW \qquad (2.57)$$

the simplest approach is to use the Euler–Maruyama method [14, 49]. This is done by approximating the state $x(t)$ at discrete time steps Δt by $x_n \approx x(t_0 + n\Delta t)$ where the x_n are iteratively generated through

$$x_{n+1} = f(x_n, n\Delta t)\Delta t + g(x_n, n\Delta t)\sqrt{\Delta t}\, Z_n. \qquad (2.58)$$

The Z_n are independent Gaussian random numbers with zero mean and unit variance. While similar to the Euler method for ODEs, the single-path error for the Euler–Maruyama method is of order $\sqrt{\Delta t}$. This is worse than the linear error for the Euler method, or Δt^4 for the (fourth order) Runge–Kutta method, so a small time increment can be necessary. However, with the availability of powerful graphics processors, the Euler–Maruyama method was adequate for our simulations with sufficiently small time steps. More complex and higher order methods for SDEs discussed in Refs. [49, 50, 51] may be useful if the Euler–Maruyama method has problematic convergence.

2.1.8. Boundary Crossing Statistics

So far we have discussed stochastic differential equations, which describe the evolution of a stochastic trajectory, as well as the Fokker–Planck equation, which describes the evolution of a probability distribution for an ensemble of stochastic trajectories. Both of these describe the evolution of the system in time, while now we move to some more descriptive statistics that can be derived from these models. In this section we derive crossing probabilities and first-passage distributions for the Wiener process. These are used in Section 4.6.1 for modeling escape times of laser-cooled atoms.

The first statistic we consider is the crossing probability. We define the crossing probability $P_{\text{cross}}(t)$ as the probability that a particle has crossed a boundary located a distance d from the origin at a time t. To derive the crossing probability one can imagine tagging particles as they cross a boundary and calculate the fraction of tagged particles to find P_{cross}; while this is a useful interpretation, to derive P_{cross}, we will use a more convenient approach by defining

an absorbing boundary and calculating the survival probability P_{surv}. Then the crossing probability is just the probability that the particles did not survive, or $P_{\text{cross}} = 1 - P_{\text{surv}}$.

To calculate the survival probability we employ a trick known as the method of images[1], where a required boundary condition (in this case the absorbing boundary) is satisfied by adding equal but oppositely weighted "anti-particle" on the opposite side of the boundary, offset by the same distance d. This guarantees that at all times t the absorbing boundary condition $f(d, t) = 0$ is satisfied. Since we know the evolution of both particles [Eq. (2.55)], the solution to the probability density is just the superposition

$$f(x, t) = \frac{1}{\sqrt{2t\sigma^2}} \left(e^{-x^2/2t\sigma^2} - e^{-(x-2d)^2/2t\sigma^2} \right), \tag{2.59}$$

which completely accounts for the absorbing boundary when restricted to the domain $x < d$. The survival probability is just the integral

$$P_{\text{surv}}(t) = \int_{-\infty}^{d} dx \, f(x, t), \tag{2.60}$$

which has the solution

$$P_{\text{surv}}(t) = \text{erf}\left(\frac{d}{\sqrt{2t\sigma^2}} \right), \tag{2.61}$$

which in turn gives us

$$P_{\text{cross}}(t) = \text{erfc}\left(\frac{d}{\sqrt{2t\sigma^2}} \right), \tag{2.62}$$

_____

[1]There are many other ways to derive the survival probability. A "brute force" method using the backwards Fokker-Planck equation is shown in Ref. [15], while a clever method using the reflection principle is shown in Ref. [41].
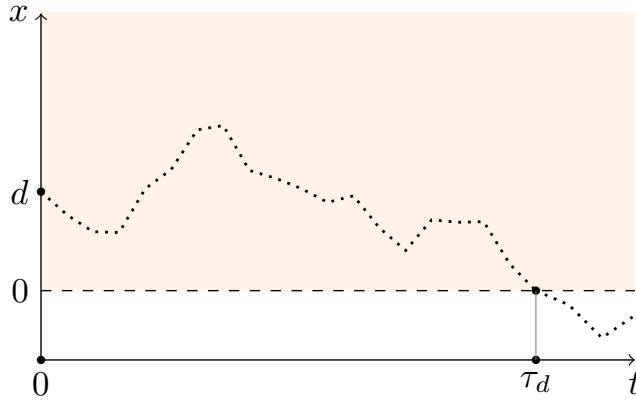
FIGURE 2.2. Schematic diagram illustrating first passage times. The process starts at d and ends when it hits the dashed line at $x = 0$.

where erf and erfc are the error function and complementary error function, respectively. In the limit $t \to \infty$, we see that $P_{\mathrm{cross}} \to 1$, showing that a random walker will always *eventually* cross a given boundary. This can be related to the "gambler's ruin problem"—even with fair odds, a persistent gambler with fixed size bets will eventually go bankrupt, and be unable to continue playing. It also is sensible that as $d \to 0$, we also have $P_{\mathrm{cross}} \to 1$, with the interpretation that if a random walker starts at a boundary, it must cross it immediately.

A related concept to the crossing probability is the first passage time, defined as the first time a particle crosses a threshold, as illustrated in Fig. 2.2. Since a particle crossing the boundary gets "tagged" or absorbed only the *first* time it crosses, we can see that the distribution of first passage times f_{τ_d} is then just the derivative of the crossing probability

$$f_{\tau_d}(t) = \frac{\partial}{\partial t'} P_{\mathrm{cross}}(t') \bigg|_{t'=t} = \frac{d}{\sqrt{2\pi t^3 \sigma^2}} \, e^{-d^2/2t\sigma^2}. \tag{2.63}$$

34

This distribution is shown in Fig. 2.3, which shows the same function plotted with linear scaling (top plot) and log-log scaling (bottom plot). The cutoff at short times is due to the low probability of the particle to exit immediately due to the Gaussian nature of the increment (exponentially damped tails). The density maximum is located at $d^2/3\sigma^2$, which serves as a transition to a $t^{-3/2}$ power law, as indicated on the log-log plot.

An extension of the first passage time for a single boundary is the escape-time distribution from an interval of length L, when the particle is offset from one of the boundaries by a distance d and offset from the other boundary by $L - d$, as illustrated in Fig. 2.4. The derivation follows similarly to that of the first passage time, starting by calculating the escape probability P_{escape}. This again uses the method of images, but since there are two boundaries, the images themselves also require images, and so on. The result is

$$P_{\text{escape}}(t) = \sum_{j=-\infty}^{\infty} (-1)^j \operatorname{sgn}(j + 0^+) \operatorname{erfc}\left(\frac{|jL + d|}{\sqrt{2t\sigma^2}} \right), \qquad (2.64)$$

and we can see that as $L \to \infty$ this reduces to the crossing probability in Eq. (2.62) when one boundary has a negligible effect. Now we can define the escape time distribution f_{τ_e} by

$$f_{\tau_e}(t) = \frac{\partial}{\partial t'} P_{\text{escape}}(t') \bigg|_{t'=t} \qquad (2.65)$$

and find

$$f_{\tau_e}(t) = \frac{1}{\sqrt{2\pi\sigma^2 t^3}} \sum_{j=-\infty}^{\infty} (-1)^j \operatorname{sgn}(j + 0^+) |jL + d| e^{-|jL+d|^2/2t\sigma^2}. \qquad (2.66)$$
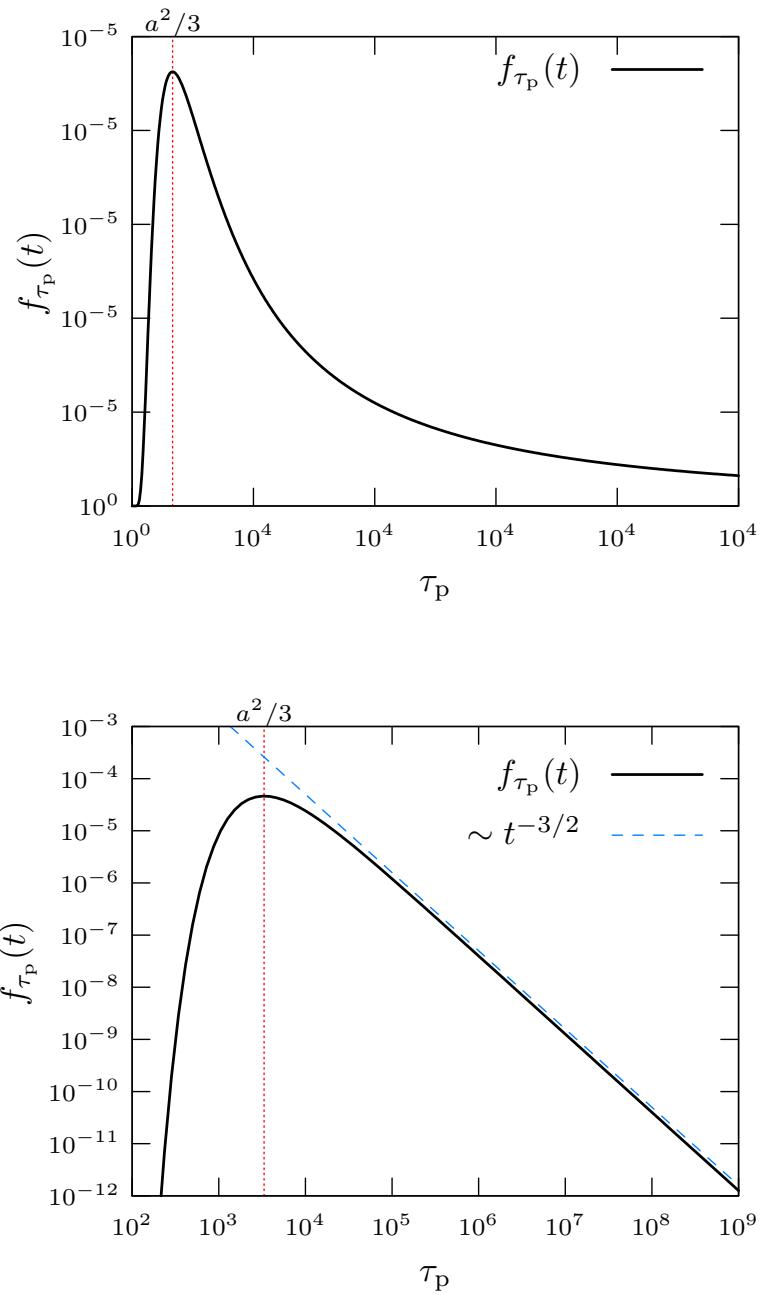
FIGURE 2.3. Plots of the first passage time distributions for a Wiener process with a single boundary. The top plot is scaled normally, while the bottom plot is log-log scaled to emphasize the $t^{-3/2}$ power law tail.

36
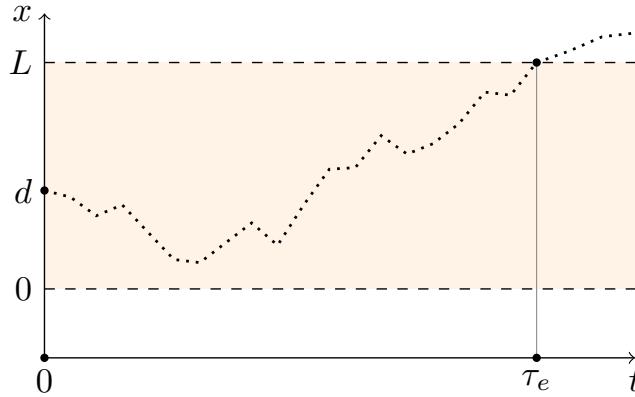
FIGURE 2.4. Schematic diagram illustrating escape times. The process starts at d and ends when it hits either dashed line at $x = 0$ of $x = L$.

This is plotted in Fig. 2.5 for two values of d. For $d = L/100$ we see that the power law present for the first passage times still occurs at $\tau_e > d^2/3\sigma^2$ but now has a cutoff at $(L-d)^2/3\sigma^2$. This essentially corresponds to the time it takes for the width of the particle density function to expand to a size $L - d$ so that the further absorbing boundary has a significant impact on the particle's survival. The power law also retains $t^{-3/2}$ scaling, despite the presence of an additional absorbing boundary. Note that this discussion is true for $d < L/2$; for $d > L/2$ the roles of $L - d$ and d are swapped. The plot for $d = L/2$ shows no visible power law since, as we will see later in this section, the power law is caused by the interaction with a *single* absorbing boundary, while when $d = L/2$ the particle is equally likely to interact with either boundary.

An important relation apparent in the above first-passage distributions is the power-law time dependence. For both the single-boundary first-passage distribution in Eq. (2.63) and the two-boundary escape-time distribution in Eq. (2.66), the asymptotic time behavior scales as $t^{-3/2}$. These are two examples of a universal scaling relation attributed to Sparre Andersen [48]. This scaling
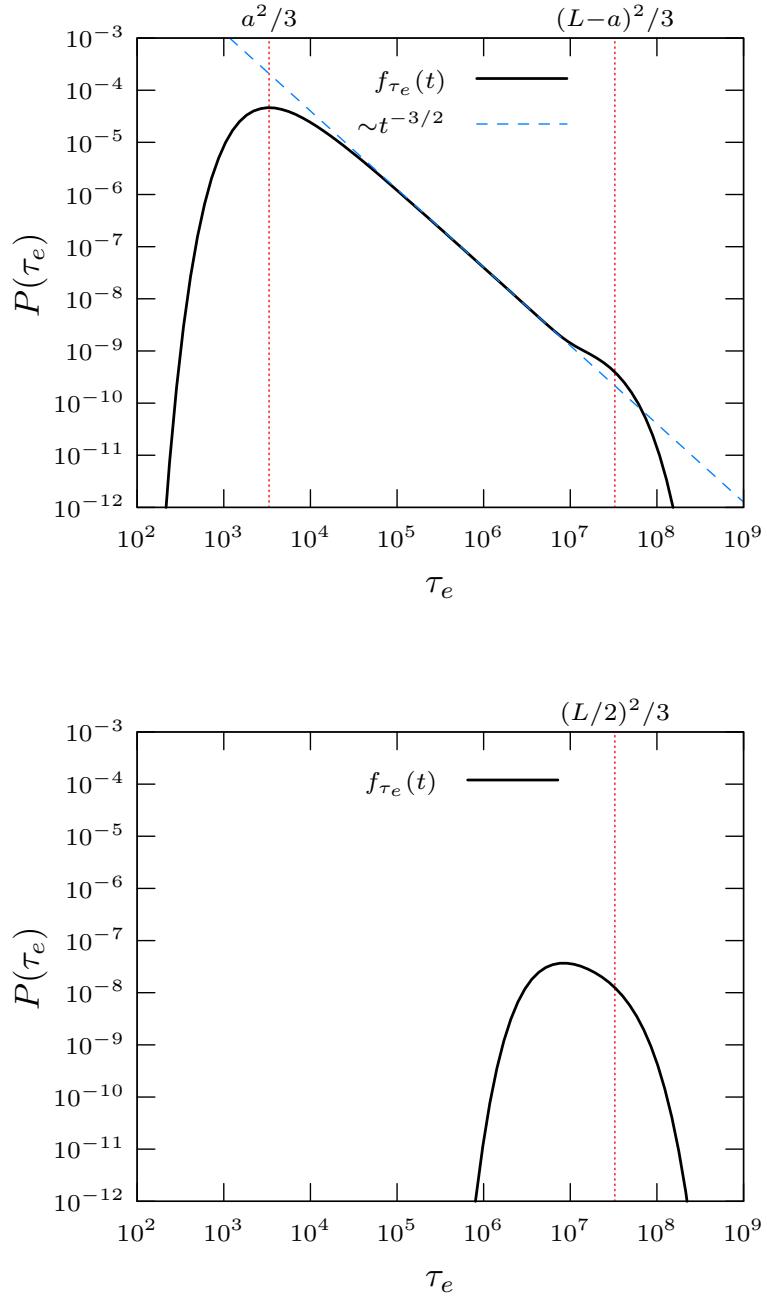
FIGURE 2.5. Escape time plots for a Wiener process with two absorbing boundaries. The top plot is shown for $a = L/100$, which has a clear power law extending over a wide range. The bottom plot is for $a = L/2$ and has no observable power law.

behavior was proved originally for discrete random walks [52, 53], but it also holds in the continuous limit [16]. Surprisingly, these universal relations are for *all* random walks with symmetric step distributions and are not limited to Gaussian processes.

2.2. Lévy Processes

Lévy processes are a natural extension of Wiener processes that encompass continuous-time random walks with drifts, heavy tailed step distributions, and most dramatically, discontinuous sample paths. They are defined to include *all* possible independent and identically distributed noise processes. Similarly to Wiener processes, Lévy processes are important in understanding a systems from a wide range of domains, including ecology [19], financial systems [13] such as option pricing [54, 55] or security returns [56], transport in fluid flows [57] and chaotic systems [58], optimal stochastic search strategies [18, 59] which can explain animal foraging behavior [60, 61], and some more obscure areas like the statistics of cricket games [62] or the process by which humans learn to balance sticks [63]. Also as we will discuss in more detail in Chapter IV, they have an important role in laser-cooled atoms [26, 31, 37, 38, 39, 64, 65, 66].

Since the CLT implies a certain universality of the Wiener process, it is reasonable to suspect there may be a natural extension to the CLT that implies a similar universality for the Lévy processes. This intuition turns out to be mostly correct: one extension to the CLT is known as the generalized central limit theorem (gCLT), and like the CLT it characterizes the how the distribution of sums of random variables converges to a limiting Lévy random process, with a couple caveats: First, the gCLT applies to sums of random variables with power-law-tailed step distributions, or Lévy flights in the language of random walks. Second, the limiting distribution of the gCLT is an α-stable distribution, which corresponds to the α-stable Lévy processes, a subset of the Lévy processes rather than a generic Lévy process. It is worth noting that many of the cited uses of the

Lévy processes predominately use α-stable Lévy processes, so the gCLT and the α-stable Lévy processes are particularly useful.

With this in mind, we will start this section by introducing the Lévy process in general (Section 2.2.1), but we will quickly restrict ourselves to using the α-stable Lévy processes (Section 2.2.2) because of the broad applicability due to the gCLT. We will then cover methods for simulation of α-stable Lévy processes (Section 2.2.4), and discuss boundary crossing statistics, including leapover distributions and first passage distributions (Section 2.2.5).

2.2.1. Lévy Process Properties and Definitions

A Lévy process $\mathcal{L}(t)$ can be defined as stochastic process that satisfies the following properties[2] [44]:

1. **Stationary Increments.** Also known as time-homogeneity, this states that that all Lévy process increments only depend on the current state and not on any other variables. That is, all elements of $\{\mathcal{L}(t_0 + t) - \mathcal{L}(t_0) : t \geq 0\}$ are independent of t_0.

2. **Independence of increments.** For any choice of $n \geq 1$ and $0 \leq t_0 < t_1 < \ldots < t_n$, all increments from $\mathcal{L}(t_0), \mathcal{L}(t_1) - \mathcal{L}(t_0), \ldots, \mathcal{L}(t_n) - \mathcal{L}(t_{n-1})$ must be statistically independent.

3. **Stochastic Continuity.** For any $\epsilon > 0$ $P(|L(s + t) - L(s)| > \epsilon) \to 0$ as $t \to 0$. This is a limited form of continuity that allows for the presence of

---

[2]There are other formulations of Lévy processes that are sometimes included or excluded in definitions of the Lévy process. This list does not include the convention of a fixed origin at $L(0) = 0$.

discontinuous jumps, as long as they occur at a random time. Deterministic discontinuities are still excluded by this form of continuity.

4. **Cadlag.** This requirement states that Lévy processes must be right continuous with left limits. This simply means that the limiting value for the process at points of discontinuity is taken to be the value *after* the jump.

As we can see, the formal definition of the Lévy process has similar properties to the Wiener process (Section 2.1.2), with the most exceptional difference being a relaxation of the Wiener process' continuity requirement, which is replaced by stochastic continuity. This allows for jump discontinuities as long as the probability for a jump goes to zero as the time interval for the jump goes to zero. The other difference is the lack of the martingale property which allows for a Lévy process to have deterministic drift.

Given this definition, it can be shown that any Lévy process has a distribution $f(x,t)$ that can be written as the Fourier transform [44]:

$$\mathcal{F}[f](k,t) = \exp\left[(t - t_0)\left(ika - \frac{b^2}{2}k^2 + \int dx' W_0(x')\left(e^{ikx' - 1 - ikx'\Theta(1 - x'^2)}\right)\right)\right].$$
(2.67)

This is known as the Lévy–Khintchine representation, with the drift coefficient a, the diffusion rate b, and a jump transition density W_0. While a and b are constants, the transition density (or Lévy measure) $W_0(x)$ represents the expected number of jumps of size x per unit time and satisfies

$$\int_{\substack{-\infty \\ (x \neq 0)}}^{\infty} dx\,(1 \wedge x^2)W_0(x) < \infty,$$
(2.68)

where we use the notation

$$
(1 \wedge x^2) = \begin{cases} 1 & |x| > 1 \\ x^2 & |x| < 1. \end{cases}
$$

(2.69)

The interpretation we gave for the parameters in Eq. (2.67) is more clear if one considers each term in the exponential separately. Recall that a distribution for the sum of random variables is just a convolution of the individual distributions, which in Fourier space is equivalent to a product of the Fourier transforms of the individual distributions. Therefore, by reversing this chain of logic, each term in the exponential directly corresponds to an independent random process: the ika term corresponds to a linear drift at, the $-b^2 k^2/2$ term becomes a Wiener process $bW(t)$, and the W_0 term is the jump process $P_{W_0}(t)$ (also known as a compound Poisson process). Thus, the Lévy process can be written as

$$
\mathcal{L}(t) = at + bW(t) + P_{W_0}(t).
$$

(2.70)

This representation is known Lévy–Itō decomposition. This decomposition allows all Lévy processes to be described as the sum of a drift term, a Wiener process, and a jump process, which is convenient and initially seems quite intuitive. However, we will see in Chapter III that this decomposition does not capture some of the strong similarities between Wiener processes and the stable Lévy processes.

2.2.2. Stable Lévy Processes

We will now define the symmetric α-stable Lévy process, which as we will see in Section 2.2.3 is an important limiting process due to the gCLT. The symmetric

43

α-stable Lévy process is a subset of the Lévy processes, and so can be defined by specifying choices for the drift coefficient a, the diffusion rate b, and the transition density W_0. The α-stable processes are pure power law jump processes, meaning both the drift and diffusion rate vanish, while the transition density is a pure power law

$$W_0(x) = \frac{A}{|x|^{\alpha+1}} \tag{2.71}$$

with the stability parameter α and jump rate parameter A. Note that α must satisfy $0 < \alpha < 2$ for Eq. (2.68) to hold. While these choices of a, b and W_0 along with Eq. (2.67) are sufficient to define the α-stable process, a more convenient form uses the scale parameter σ, defined via the relation

$$A = \sin(\pi\alpha/2)\Gamma(\alpha+1)\sigma^\alpha/\pi. \tag{2.72}$$

Using this, Eq. (2.67) can be simplified into the form

$$f_\alpha(x;t) = \frac{1}{\pi} \int_0^\infty dk \, \cos(kx) e^{-t\sigma^\alpha k^\alpha} \tag{2.73}$$

where again σ is the scale parameter and α is the stability parameter with $0 < \alpha < 2$.

In general there is no simple expression for the integral in Eq. (2.73), but this form is still convenient as it can be easily compared to the Fourier transform of a standard Gaussian distribution. We can see that as $\alpha \to 2$ the distribution is exactly the Gaussian

$$f_2(x;t) = \frac{1}{\sqrt{4\pi\sigma^2 t}} e^{-x^2/4\sigma^2 t} \tag{2.74}$$

which has a standard deviation $\sqrt{2t}\sigma$. The only other α with a simple expression is $\alpha = 1$ which reduces to the Cauchy distribution given by

$$f_1(x;t) = \frac{\sigma t}{\pi \left(x^2 + \sigma^2 t^2\right)}.$$ (2.75)

We can see that the Cauchy distribution has asymptotic $|x|^{-2}$ tails, and similar behavior is present for all for all $\alpha < 2$, where the densities have heavy tails with power-law behavior scaling as $|x|^{-(1+\alpha)}$, matching the transition density in Eq. (2.71).

As an example, sample paths for selected α are shown in Fig. 2.6. Long jumps associated with rare events in the tails of the distribution are present for all α, but the jumps tend to be increasingly important for smaller alpha, a reflection of the heavier tails.

2.2.3. Generalized Central Limit Theorem

The generalized Central Limit Theorem (gCLT) is an expansion of the CLT to account for sums of random variables drawn from a distribution an infinite variance, where the infinite variance is due to the presence of power law tails. More precisely, if a step distribution has power law tails that decay like $f_{\Delta x}(x) \sim |x|^{-(\alpha+1)}$ with $0 < \alpha < 2$, then the gCLT states that a sum of these random variables will approach an α-stable distribution, defined by

$$f_\alpha(x) = \frac{1}{\pi} \int_0^\infty dk \, \cos(kx - k\mu) e^{-\sigma^\alpha k^\alpha}.$$ (2.76)

The gCLT can be motivated with a few observations: First, we can see that the step distribution for $\alpha \geq 2$ has a finite second moment, and so would approach
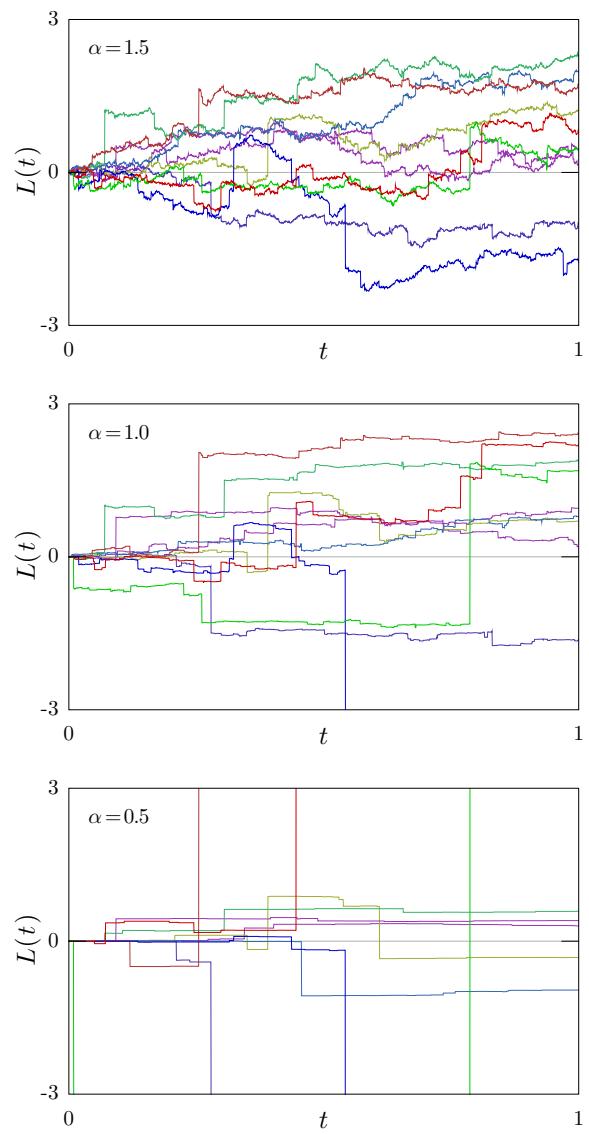
FIGURE 2.6. Plot of 10 sample paths of the α-stable Lévy process for 3 values of α.

46

a Gaussian distribution due to the regular CLT as expected. However, when $\alpha < 2$ the requirements for the CLT would not be satisfied as $\int dx\, f_{\Delta x}(x) x^2$ diverges. But while the second moment diverges for $0 < \alpha < 2$, the distribution remains sensible as a probability distribution since $\int dx\, f_{\Delta x}(x)$ is still finite. Thus we can anticipate the existence of some sort of limiting distribution for f_{x_n} even if the standard CLT fails.

Proving the gCLT is similar to proving the CLT, and can be seen in more detail in Refs. [17, 41]. Again, similar to the CLT, the gCLT gives strong grounding for the universality of the α-stable Lévy process. As for actually applying the gCLT in practice it is important to understand how the processes typically converge which we discuss in the next section.

2.2.3.1. Slow Convergence to Gaussian

One concern about applying the gCLT to a system is that it is often difficult to guaranteed that a power law tail holds indefinitely at all scales. However, it turns out that even with a power law that only spans a finite range, the gCLT is still useful for modeling behavior. The purpose of this section is to help build intuition for how stochastic processes can be described by the asymptotic Gaussian and stable Lévy distributions at finite times. It is particularly useful for interpretation of many of the simulations shown in Chapter IV. To see how systems can approach the asymptotic limits, we look at an example of applying both the CLT and the gCLT to a system with heavy but truncated tails.
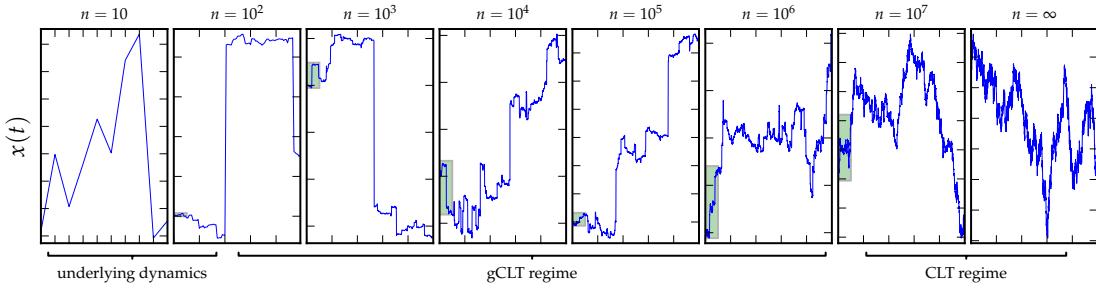
FIGURE 2.7. Trajectories from a single truncated Cauchy random walk with steps drawn from Eq. (2.77) with $a = 10^5$. Each plot moving to the right contains the entire motion of the previous plot, highlighted in green. For $n \ll a$ the motion appears like a Cauchy random walk, while for $n \gg a$ the motion appears Brownian.

Consider a random walk with n independent steps drawn from a truncated Cauchy distribution defined by

$$
f_{\text{tc}}(\Delta x) = \begin{cases} \dfrac{1}{2\tan^{-1}(a)} \dfrac{1}{1 + (\Delta x)^2} & |\Delta x| > a \\ 0 & |\Delta x| < a \end{cases} \tag{2.77}
$$

where the $1/2\tan^{-1}(a)$ is just a normalization factor and a is a parameter determining the truncation cutoff. This distribution has a finite second moment $\sigma_0 = -1 + a/\tan^{-1}(a)$, and so in principle the CLT is applicable. However, the extent of the tails can be large (controlled by a), limiting the rate of convergence to a Gaussian. In practice there can be a large range of n where the gCLT should be applied instead. To see this qualitatively, we have plotted the trajectory of a single random walker with $a = 10^5$ for various values of n, shown in Fig. 2.7. When $n \gg a$, the motion appears to be Brownian motion (Brownian motion shown as $n = \infty$ for comparison), however, for $n < a$ we notice that the motion appears like a Cauchy random walk with large jumps.

To see this more quantitatively, we can compare how the CLT limiting distribution, given by

$$f_{\text{CLT}}(x,t) = \frac{1}{\sqrt{2\pi\sigma(t)^2}} e^{-x^2/2\sigma(t)^2}, \tag{2.78}$$

compares to the gCLT limiting distribution

$$f_{\text{gCLT}}(x,t) = \frac{\gamma(t)}{\pi\left(\gamma(t)^2 + x^2\right)} \tag{2.79}$$

where $\sigma(t) = \sigma_0\sqrt{t}$ and $\gamma(t) = t$. The latter should be valid only as $a \to \infty$. To see the comparison, in Fig. 2.8 we show the position distribution of 10^5 random walks with steps drawn from Eq. (2.77) with $a = 10^5$ for various numbers of steps. For $n \ll a$ the simulated distribution tends to overlap with the Lévy distribution, while for $n \gg a$ the simulated distribution tends to overlap with the Gaussian distribution. This "slow" convergence to Gaussian behavior due to the presence of power laws has been noted before [67, 68].

2.2.4. Sampling and Simulation

This section covers standard methods for numerical simulation of stable Lévy processes, as well as how to sample from stable Lévy distributions. These techniques are used for all simulations in Chapter III that involve stable Lévy processes. To simulate an α-stable Lévy process, we write the process as a discrete random walk with positions given by the relation

$$x_{n+1} = x_n + (K_\alpha)\left(\Delta t\right)^{1/\alpha} X_\alpha(n) \tag{2.80}$$
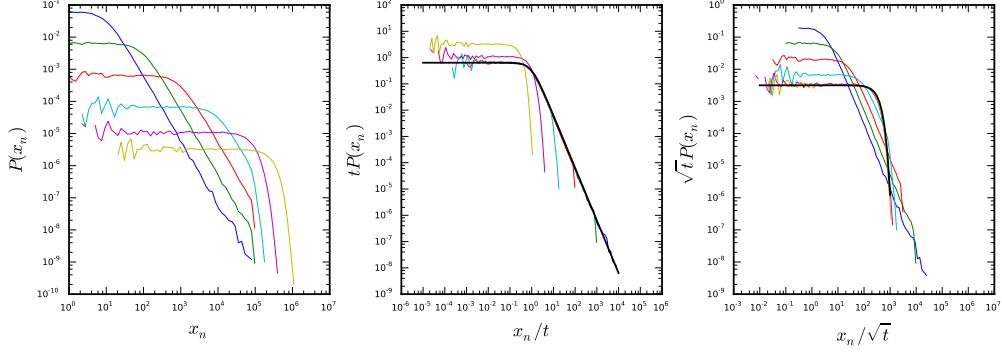
FIGURE 2.8. Demonstration of the applicability of the gCLT even for step distributions with a finite variance. Number of steps taken is indicated by the line color: 10^1 (blue), 10^2 (green), 10^3 (red), 10^4 (teal), 10^5 (purple) 10^6 (yellow). The second plot is scaled such that the distributions overlap with Eq. (2.79) (second plot, black), while the third plot is scaled such that the distributions overlap with Eq. (2.78) (third plot, black).

where K_α is an anomalous diffusion coefficient and $X_\alpha(n)$ is an α-stable random variable.

As for generating the stable random variables X_α, many numerical tools and packages often do not have built in methods. However, there is a straightforward formula that transforms a uniform random number U on the interval $\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$ and an exponential random variable W with mean 1 into a stable random variable through the transformation

$$X = (1 + \zeta^2)^{1/2\alpha} \frac{\sin(\alpha(U + \xi))}{(\cos(U))^{1/\alpha}} \left(\frac{\cos(U) - \alpha(U + \xi)}{W}\right)^{(1-\alpha)/\alpha}, \qquad (2.81)$$

where

$$\zeta = -\beta \tan(\pi\alpha/2) \qquad (2.82)$$

and

$$\xi = \arctan(-\zeta)/\alpha. \qquad (2.83)$$

This is valid for all $0 < \alpha < 1$ and $1 < \alpha < 2$ [69]. When $\alpha = 1$ the alternative formula

$$X = \frac{2}{\pi} \left(\left(\frac{\pi}{2} + \beta U \right) \tan(U) - \beta \log \left(\frac{\pi W \cos(U)}{\pi + 2\beta U} \right) \right) \qquad (2.84)$$

is used, while when $\alpha = 2$ the variable X is a Gaussian random variable. While this method is straightforward, a more numerically efficient algorithm is also outlined in Ref. [69].

2.2.5. Boundary Crossing Statistics

In this section we will briefly discuss some important and counterintuitive boundary crossing statistics for the Lévy processes. In Section 2.1.8 we discussed some of these statistics for the Wiener process, which included crossing probabilities and first passage times, however, results for Lévy processes are much more limited.

2.2.5.1. First Passage Times

The derivations for first-passage distributions in Section 2.1.8 relied on the method of images. However, it has been shown that this approach does not work for the stable Lévy processes [70], and ultimately leads to erroneous results. The breakdown of the method of images is due to the discontinuous nature of Lévy processes. In the Brownian case, the image method works by guaranteeing zero net probability flux across an absorbing boundary by effectively forcing the boundary to have zero net probability at all times (see Section 2.1.8). Meanwhile, in the Lévy case, a particle can effectively leap across a boundary, so that zero net probability at a single point does *not* prevent a particle from crossing the boundary.

51
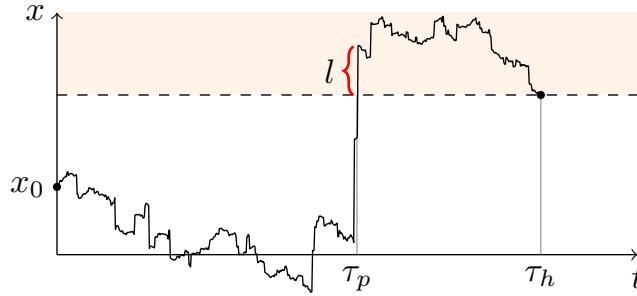
FIGURE 2.9. Schematic diagram illustrating the leapover distance l, as well as the distinction between first passage times τ_p and first hitting times τ_h for a Lévy process.

Because of this difficulty, there is no simple expression for the first passage distributions for Lévy processes, and instead we only have the asymptotic $\tau^{-3/2}$ behavior for a single boundary, as was already guaranteed through the universal Sparre Andersen scaling [70].

2.2.5.2. Leapovers

As discussed previously, the discontinuous nature of Lévy processes leads to interesting behavior when assuming absorbing boundary conditions. For example, consider an α-stable Lévy process starting at some $x_0 > 0$ with an absorbing point at $x = 0$. While a Wiener process would remain positive until it eventually is absorbed, a Lévy process can hop between negative and positive values over the absorbing point at the origin many times before it is absorbed [70, 71]. Again, this is because of the discontinuous jumps permitted for Lévy processes.

This behavior also is applicable to an absorbing region, rather than an absorbing point. For an absorbing half-line region $(-\infty, 0]$, a Lévy process again starting at some $x_0 > 0$ will (almost surely) penetrate some distance l into the absorbing region before it is absorbed. This leapover length is illustrated in

Fig. 2.9. Also illustrated in this figure is the difference in between the first hitting time τ_h and the first passage time τ_p. While the first passage time is marked the moment the particle crosses the boundary, the first hitting time is not marked until the particle diffuses "into" the boundary.[3]

The leapover length can be characterized by the leapover distribution f_{L}. For an absorbing interval $(-\infty, 0]$ and a particle starting at $x_0 > 0$, the leapover distribution is given by [72, 73]

$$f_{\mathrm{L}}(l) = \frac{1}{\pi} \sin\left(\frac{\pi\alpha}{2}\right) \frac{x_0^{\alpha/2}}{l^{\alpha/2}(x_0 + l)}. \tag{2.85}$$

For a finite absorbing region $[-2L, 0]$ and a particle starting at $x_0 > 0$, the leapover distribution is given by

$$f_{\mathrm{L}}(x) = \frac{1}{\pi} \sin\left(\frac{\pi\alpha}{2}\right) \frac{|L^2 - (x_0 + L)^2|^{\alpha/2}}{|L^2 - (x + L)^2|^{\alpha/2}} \frac{1}{|x_0 - x|} \tag{2.86}$$
$$- \frac{\alpha - 1}{\pi L} \sin\left(\frac{\pi\alpha}{2}\right) \left(1 - \frac{(x + L)^2}{L^2}\right)^{-\alpha/2} \int_1^{|x_0+L|/L} du \, (u^2 - 1)^{\alpha/2 - 1}.$$

Plots for Eq. (2.86) are shown in Fig. 2.10. While the particle starts with $x_0 > 0$, it is somewhat remarkable that the probability is peaked near $x = -2L$. This can be explained by the possibility that the particle jumps across the region entirely, at which point it is more likely to enter the region near $x = -2L$ on the subsequent step.

---

[3]The criteria for a particle "hitting" the boundary can be quite subtle. While a δ absorbing point can always be approximated by an absorbing region with a finite width, the appropriate width of the region for a given time scale is not obvious. This is discussed more in Chapter III.
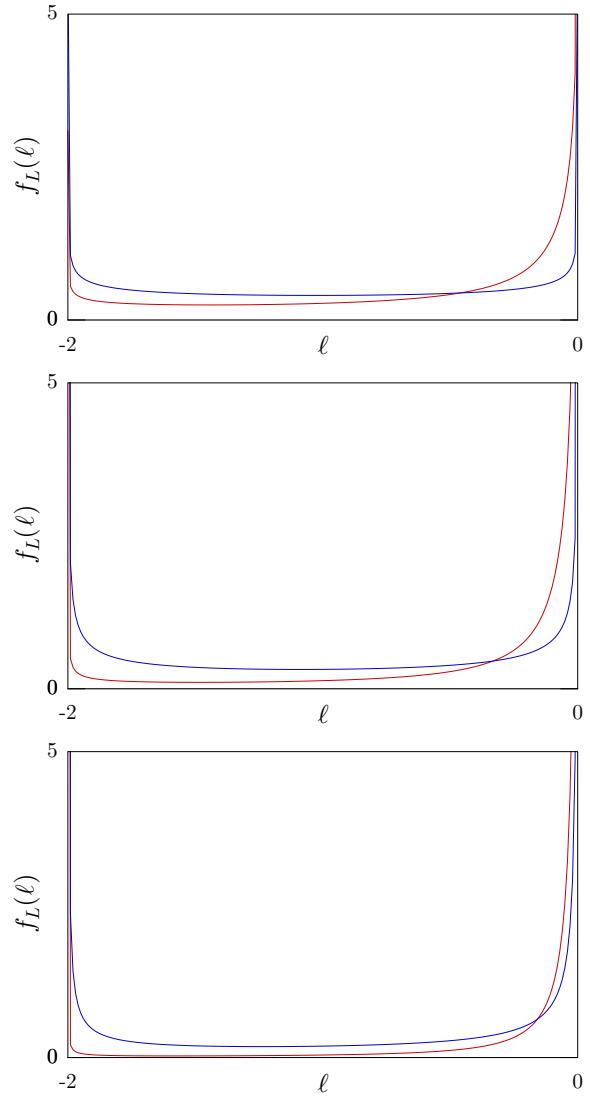
FIGURE 2.10. Leapover distributions (Eq. (2.86)) into the interval $(-2L, 0)$ with $L = 1$ and $x_0 = 0.1$ (red) and $x_0 = 10.0$ (blue).

54

CHAPTER III

CONDITIONED RANDOM PROCESSES

So far in our investigation of random processes we have assumed a stochastic process with initial condition and predicted some expected behavior through use of Ito SDEs, Fokker–Planck equations, and first passage distributions. However, in certain circumstances both the initial and final condition are known, while the history of how the transition occurred is unknown.

As an example, consider a particle within the boundaries of some imaging system, such that scattered light from the particle is detected by a single photodetector. A sudden drop in the photodetector signal indicates that the particle must have escaped the imaging boundaries. However, the *path* that the particle took is unknown. This example is particularly relevant to some simulations we discuss in Chapter IV, but to investigate this type of system generically, we study the conditioned evolution of the process. Studies of conditioned evolution lends insight into the likely history of an atom in this hypothetical experiment.

And so, in this chapter we will investigate conditioned random processes for both Wiener processes and Lévy processes. The most basic conditioned process is a stochastic bridge, where a process has known initial conditions and a known final condition after some evolution time. The stochastic bridges represent possible behavior of the particle between measurements, and they are important for both understanding the qualitative behavior of the motion as well as for computation of a variety of relevant statistics.

In Section 3.1, we introduce some well-known properties of the conditioned Wiener processes (Brownian bridges). Next, in Section 3.2, we introduce the conditioned Lévy process and briefly discuss existing literature, before moving on to some of the main results of this dissertation: In Section 3.2.3 we show that a bifurcation arises in the most probable history of a conditioned Lévy process, with the onset of the bifurcation occurring when the final displacement of the conditioned processes far exceeds a certain value. The value at which the bifurcation arises defines a novel length scale (the "bifurcation length") which we show has applications for detecting extreme events (Section 3.2.5), generating Lévy bridges (Section 3.2.6), explaining properties of first passage times (Section 3.2.7), and parameter inference of a Lévy process from data (Section 3.2.8).

3.1. Brownian Bridges

The most basic conditioned process is the the Brownian bridge—a continuous-time Gaussian stochastic process with a fixed starting location, a given evolution time, and a fixed final location. It also is one of the most widely used examples of a conditioned process and has been thoroughly studied [1] and applied in diverse areas, such as financial mathematics [74, 75], models of animal movements [76], Monte Carlo methods in quantum mechanics [77, 78], random interfaces and potentials [79, 80, 81], and extreme-value statistics [82].

3.1.1. Stretched Brownian Bridge

There are multiple useful ways of expressing a Brownian bridge, but a common and convenient definition is the "stretched" Wiener path. The standard Brownian bridge $B(t)$ is defined on the time interval $[0, 1]$ with both starting and

ending points at $x = 0$, and can be written as

$$B(t) = W(t) - tW(1), \qquad (3.1)$$

where the bridge is compressed back to the origin by gradually subtracting off the final position of an unconstrained Wiener process. In this definition, the bridge is essentially a Wiener process plus a linear drift, such that the linear drift and the Wiener process cancel out at the final time.

Surprisingly, this "stretch" method for generating bridges generates the correct increment statistics of a Wiener process. That is, the total drift needed to cancel the Wiener process at the final time is spread evenly over the time interval, so that the statistics of the infinitesimal increment end up being the same as the pure Wiener process. This aspect of the Brownian bridge can be seen by looking at the infinitesimal increment

$$dB(t) = dW(t) - dt\, W(1). \qquad (3.2)$$

Since the constituent parts of dB are Gaussian noise, we know that dB must also be Gaussian and will be characterized by its mean and variance. Then, $\langle dB(t) \rangle = 0$ since both $dW(t)$ and $W(1)$ are symmetric, while

$$\langle dB(t)^2 \rangle = \langle dW(t)^2 \rangle + 2dt\langle dW(t)W(1) \rangle + dt^2 \langle W(1)^2 \rangle \qquad (3.3)$$

which simplifies to $\langle dB(t)^2 \rangle = dt$ when we drop terms of higher order than dt in the continuum limit. Since the variance is dt, the statistics in this limit are identical to that of the Wiener process.

We also can extend the standard bridge to stretch to an arbitrary final time T at position L. In this case the stretched Brownian bridge is given by

$$B(t) = W(t) - \frac{t}{T}\left(W(T) - L\right), \qquad (3.4)$$

and calculating the statistics of the increment is similar to the standard bridge case, though we now have a non zero mean increment $\langle dB(t) \rangle = dt\frac{L}{T}$, while the variance remains $\langle dB(t)^2 \rangle = dt$ in the continuum limit.

3.1.2. Midpoint Brownian Bridge

While the "stretch" method is convenient for generating Brownian bridges, it does not follow obviously from the definition of a conditioned stochastic bridge. A more easily motivated method is to generate a Brownian bridge by iteratively sampling the midpoint of the time interval using conditioned distributions. To do this, we need to find the probability density for the position of the particle at time $T/2$, given that the particle has a known final position L at the final time T. For this intermediate position we use the notation $x_{1/2} := x(T/2)$. Finding the density of $x_{1/2}$, which we refer to as the the midpoint density, is a straightforward application of Bayes' Theorem, $P(A|B) = P(B|A)P(A)/P(B)$. Since we are interested in the probability density that the particle passes through $(x_{1/2}; T/2)$ (condition A) conditioned on the particle passing $(L; T)$ (condition B), the midpoint density can be written as

$$f_W(x_{1/2}; T/2|L; T) = \frac{f_W(L; T|x_{1/2}; T/2)f_W(x_{1/2}; T/2)}{f_W(L; T)} \qquad (3.5)$$

where $f_W(x, t)$ is the probability density for a Wiener process to be at position x at time t. Since the propagator only depends on the relative displacement and duration, the first term can be simplified to find

$$f_W(x_{1/2}; T/2|L; T) = \frac{f_W(L - x_{1/2}; T/2)f_W(x_{1/2}; T/2)}{f_W(L; T)} \qquad (3.6)$$

and by using Eq. (2.55) we find that this factors into

$$f_W(x_{1/2}; T/2|L; T) = f_W\left(x_{1/2} - L/2, T/4\right) \qquad (3.7)$$

which is just a Gaussian centered at $x = L/2$. This matches the intuition that at time $T/2$ the most likely position should be halfway between the starting and ending positions ($\langle x_{1/2} \rangle = L/2$), while the variance of the overall position of the process is somewhat constrained due to the conditioning (at time $T/2$ the free Wiener process has a variance $T/2$, while we see that the bridge has a reduced variance of $T/4$).

To use Eq. (3.7) to generate an entire bridge, we start by generating sample points for bridges at time $T/2$. Once this is done, the process is bisected into two bridges, at which point this process can be repeated to generate sample points in each subinterval of duration $T/4$. By iterating this sampling process we can generate bridges to arbitrarily fine resolution. This process was used to generate the Bridges shown in Fig. 3.1.

3.1.3. First Passage Times

Similarly to the free Wiener process, we can also develop first passage statistics for the Brownian bridge. For a particle conditioned to travel from
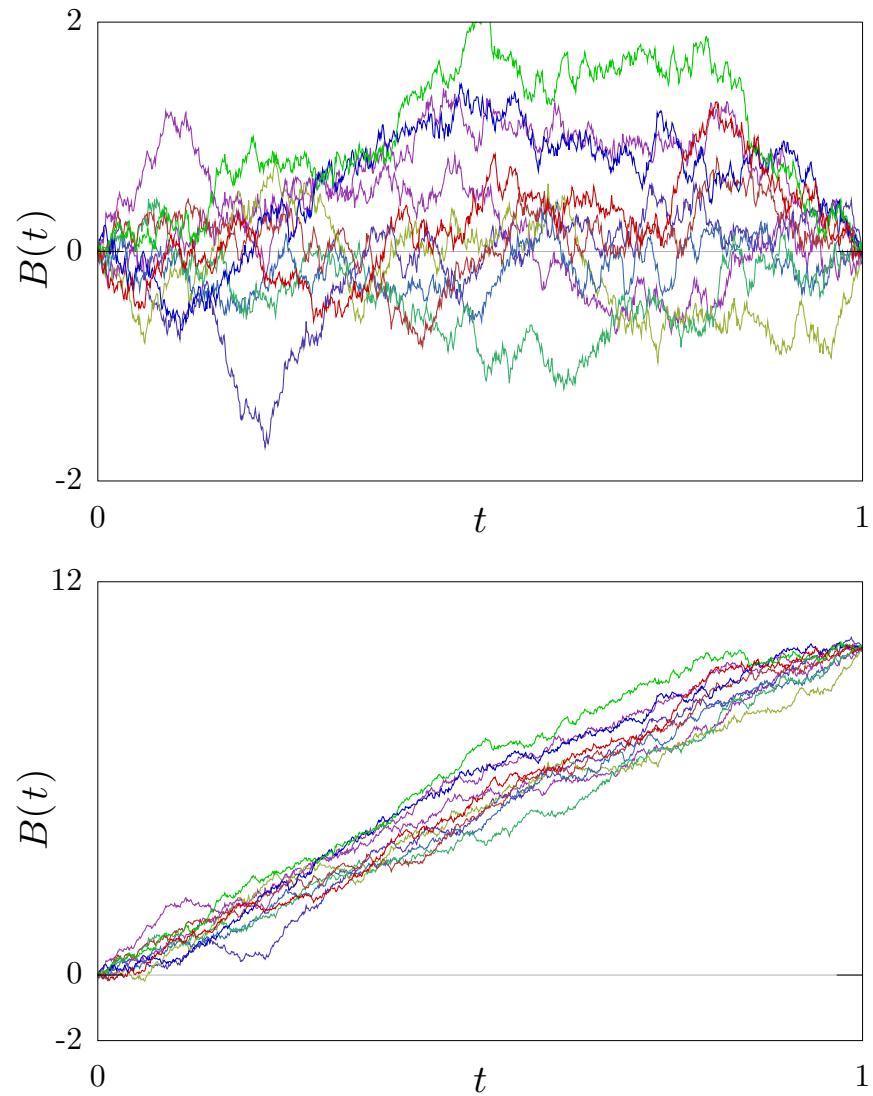
FIGURE 3.1. Examples of Brownian bridges generated through iterated sampling of the midpoint distribution using Eq. (3.7). The top plot has bridges with final displacement $L = 0$, while the bottom figure has bridges with $L = 10$.
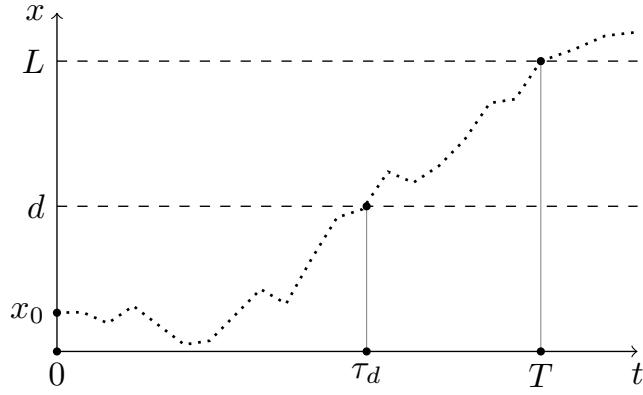
FIGURE 3.2. Schematic diagram illustrating the first passage time for the Brownian bridge.

position $x = 0$ to $x = L$ in a time t, the distribution of first passage times across a boundary at $x = d$ is given by [47]

$$f_{\tau_d}(x) = \frac{d\sqrt{t}}{\sqrt{2\pi x^3(t-x)}} e^{-(dt-Lx)^2/(2tx(t-x))}.$$ (3.8)

This process is illustrated in Fig. 3.2, which makes it clear that for $L > d$ the process is guaranteed to have a first passage time. For $L < d$ it is possible that no first passage occurs, and so Eq. (3.8) becomes unnormalized.

3.2. Lévy Bridges

Similar to the Brownian bridge, a Lévy bridge $B_\alpha(t)$ is defined as a Lévy process conditioned to have a particular final position. Unlike the Brownian bridge, however, there is relatively little literature related to Lévy bridges, which is somewhat surprising considering the many applications of Lévy statistics (as discussed in the introduction and in Section 2.2). Refs. [83, 84] characterize several functionals of Lévy bridges, through as the definition used in these references

61

constrains the bridge to return to the origin, the results have limited relation to this chapter. There are also some preliminary investigations into Lévy bridges that are not constrained to the origin in Ref. [85]; however, while the formulation was made for general Lévy processes, the primary results were specifically for constrained Gamma processes and Wiener processes; the results were also for computation of certain pricing statistics. The most relevant reference to the work in this chapter is Ref. [86], which demonstrated the construction of stable Lévy bridges for $\alpha = 1/2$ with the skewness parameter $\beta = 1$ (i.e. not a *symmetric* stable Lévy process). This reference mentions in passing that the frequency of long jumps varies as a scaling parameter is adjusted, which has some relation to our results in Section 3.2.5 and in Section 3.2.8, though the interpretation is quite different. It is also worth mentioning that some basic properties and constructions for more general Markovian bridges have also been considered [87, 88, 89], though as these processes do not even require stochastic continuity (see Section 2.2.1), they are unlikely to find practical use for describing spatial diffusion.

While Lévy bridges are defined similarly to Brownian bridges, much of the intuition developed for Brownian bridges does not carry over when studying conditioned Lévy processes. This has led to a somewhat dubious definition of a Lévy bridge [90], where the bridge is not even time symmetric. To address this directly, we begin this section with an explicit example of how a naïve extension of the Brownian bridge "stretch" method fails for Lévy bridges (Section 3.2.1). We then develop a proper Lévy bridge using conditioned distributions leading to a midpoint density that can be used to iteratively sample points of a Lévy bridge to arbitrary accuracy (Section 3.2.2).

The midpoint distribution also provides surprising insight into the behavior or Lévy processes. We find that unlike in the Brownian case, the Lévy midpoint distribution bifurcates when the bridge exceeds a critical length L_b (Section 3.2.3). This bifurcation length provides insight into the transition between the Lévy and Gaussian regimes, criteria for detecting "long" jumps (Section 3.2.5), correcting the stretched bridge method (Section 3.2.6), inferring the stability index α from a given sample path (Section 3.2.8), and in general can be useful for the analysis of extreme events.

3.2.1. Failure of Stretched Lévy Bridges

One of the first peculiarities of Lévy bridges is that the "stretch" method we used for the Wiener bridge does not generalize in an obvious way to the Lévy case [83]. This is also briefly indicated in Ref. [85], where the stretched Lévy bridge definition is briefly considered before being rejected.

The reason for the failure of the stretch method is somewhat subtle, since in Section 3.1.1 we saw that the method works for the Brownian case because the increment dt is small compared to $dt^{1/2}$ in the continuous limit. This argument initially seems like it should hold for the Lévy case as well, as $dt^{1/\alpha}$ is still small compared to dt for $0 < \alpha < 2$. However, the reasoning breaks since in some sense the drift term cannot simply be considered order dt due to the heavy tails of the final position. Additional details related to this argument can be seen in [83, 85].

To show that the stretched bridges do not generate correct statistics, we define the stretched Lévy bridge as

$$B_{\alpha\text{-stretch}}(t) = \mathcal{L}_\alpha(t) - \frac{t}{T}\left(\mathcal{L}_\alpha(T) - L\right). \tag{3.9}$$
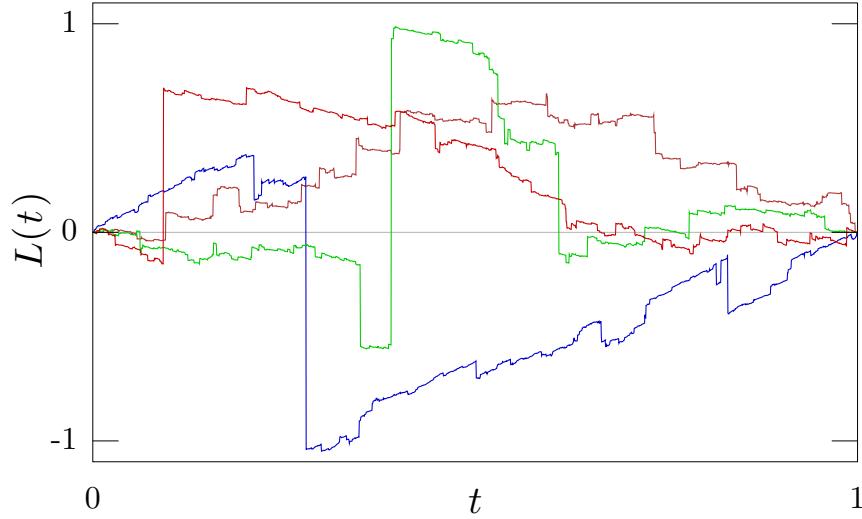
FIGURE 3.3. Examples of stretched Lévy bridges generated through Eq. (3.9) with $T = 1$ and $L = 0$. Some of the bridges have noticeable drift, indicating that the stretch method for Lévy bridges does not work properly.

Looking at bridges created using this approach in Fig. 3.3, we can immediately see that they are somewhat peculiar. Roughly speaking, bridges should "appear" to be identical to the free underlying process, except that they happen to end at a particular location. Instead, we have that even though the bridges have $B(T) = 0$ for final position, many of the bridges appear to have a strong drift present over the entire bridge. While the bridge does return to the origin as required (indicating that the mean increment is zero), this is apparently not sufficient, as the *typical* step still has a significant drift which causes the bridge to appear biased.

To check this more rigorously, we can consider an arbitrary function $F[\mathcal{X}]$ of a path $\mathcal{X}$. The distribution of this function over all unconstrained Lévy processes $f_{F[\mathcal{L}_\alpha]}(x)$ should match the distribution over constrained processes $f_{F[B_{\alpha\text{-stretch}}|L=L']}$ when integrated over all L' and weighted by the probability of the final position

64
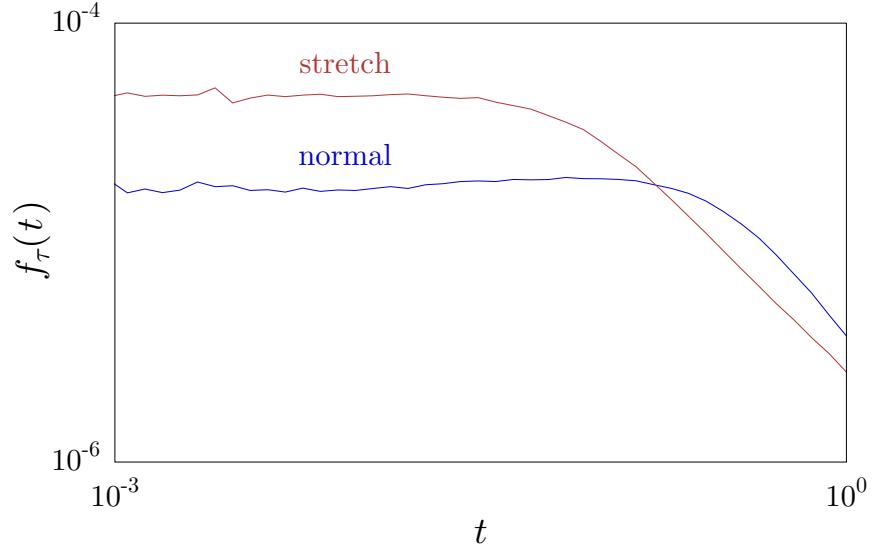
FIGURE 3.4. Test showing the failure the equality shown in Eq. (3.10) using the first passage time for F. The LHS of Eq. (3.10) is shown in blue, while the RHS is shown in red. Simulations are for $\alpha = 1$, with the passage boundary $d = 10$, time step $\Delta t = 10^{-4}$, bridge constrained time $T = 1$, and $N = 10^6$ trajectories.

$f_\alpha(L'; T)$. That is the equality

$$f_{F[\mathcal{L}_\alpha]}(x) = \int_{-\infty}^{\infty} dL'\, f_{F[B_{\alpha\text{-stretch}}|L=L']} f_\alpha(L'; T) \tag{3.10}$$

should hold if $B_{\alpha\text{-stretch}}$ is a proper bridge.

We can test this numerically by choosing F to be some easily observable statistic. As an example we choose F to be the first passage time, with a plot of the corresponding distributions shown in Fig. 3.4. The distribution $f_{F[\mathcal{L}_\alpha]}(x)$ is shown in blue while the distribution $\int_{-\infty}^{\infty} dL'\, f_{F[B_{\alpha\text{-stretch}}|L=L']} f_\alpha(L'; T)$ is shown in red, clearly indicating significant deviations. The stretch method has a large excess of short first passage times compared to the correct distribution.

3.2.2. Lévy Midpoint Density

While the stretched bridges do not work for Lévy bridges, the midpoint sampling approach from Section 3.1.2 does apply, since it follows directly from the definition of the Lévy bridge. As we did for the Brownian bridge, the Lévy bridge arrival point is specified as $x(T) = L$ for some arrival time $T > 0$. Then, the intermediate position $x_{1/2}$ has the midpoint density

$$f_\alpha(x_{1/2}; T/2|L; T) = \frac{f_\alpha(L; T|x_{1/2}; T/2)f_\alpha(x_{1/2}; T/2)}{f_\alpha(L; T)}, \qquad (3.11)$$

where the $f_\alpha(x_1; t/2)$ and $f_\alpha(x_2; t)$ factors are simply the free α-stable density from Eq. (2.73). As in the Brownian case, the $f_\alpha(L; T|x_{1/2}; T/2)$ term can be rewritten as $f_\alpha(L - x_{1/2}|T/2)$ since a Lévy process has stationary increments. This gives us the midstep density

$$f_\alpha(x_{1/2}; T/2|x = L; T) = \frac{f_\alpha(x_{1/2}; T/2)\, f_\alpha(L - x_{1/2}; T/2)}{f_\alpha(L; T)}, \qquad (3.12)$$

written in terms of a product of the unconditioned density $f_\alpha(x; t)$. Similar to the Brownian case, the distribution is symmetric about $x_{1/2} = L/2$. However, unlike the Brownian case, these terms do not factor, which leads to a number of important properties, most notably that $x_{1/2} = L/2$ is not always a maximum of the midstep distribution.

To use Eq. (3.12) for sampling Lévy bridges, we can use the same approach as the Brownian case. Once $x_{1/2}$ at time $T/2$ is sampled, the bridge is effectively bisected into two bridges each of duration $T/2$. This allows for the midpoints of the new bridges to be sampled, a process that may be iterated to sample the Lévy

66

bridge to any desired time resolution. Bridges generated using this method with $\alpha = 1$ are shown in Fig. 3.5 for both $L = 0$ and $L = 10$. We can see that the character of the bridges is completely different from the failed stretched bridges in Fig. 3.3, and no notable drifts are present for either L. This is even true for $L = 10$ which is somewhat surprising: a Brownian bridges with $L = 10$ does have significant drifts as seen in Fig. 3.1, while the Lévy bridge with $L = 10$ instead induces a long jump rather than a drift. This hints at some counterintuitive behavior and an interesting transition, since the Gaussian case should behave similarly to the Lévy case when α is large. In Fig. 3.6 we have simulated bridges for $\alpha = 1.9$ and $L = 2$, which seem to have the drift behavior that was present in the Brownian bridges in Fig. 3.1, so we can see that the behavior of inducing a long jump for $L > 0$ is not completely guaranteed for all L and α. To explain this behavior we will need to consider the midstep distribution in more depth.

3.2.3. Lévy Bridge Probability Bifurcations

We discussed in the previous section how the midstep distribution [Eq. (3.12)] is symmetric about $x_{1/2} = L/2$; however, there is no guarantee that the midstep distribution is maximized at $x_{1/2} = L/2$ like it is for $\alpha = 2$. In fact there is a transition with L where the maximum at $x_{1/2} = L/2$ becomes a minimum. This transition in the conditional probability density marks the shift from a unimodal to bimodal distribution. We refer to this shift as a probability bifurcation,[1] and it leads to many interesting and counter-intuitive effects that we will investigate throughout the rest of this chapter.

---

[1]The use of the term probability bifurcation has been used in a similar context in Ref. [91].
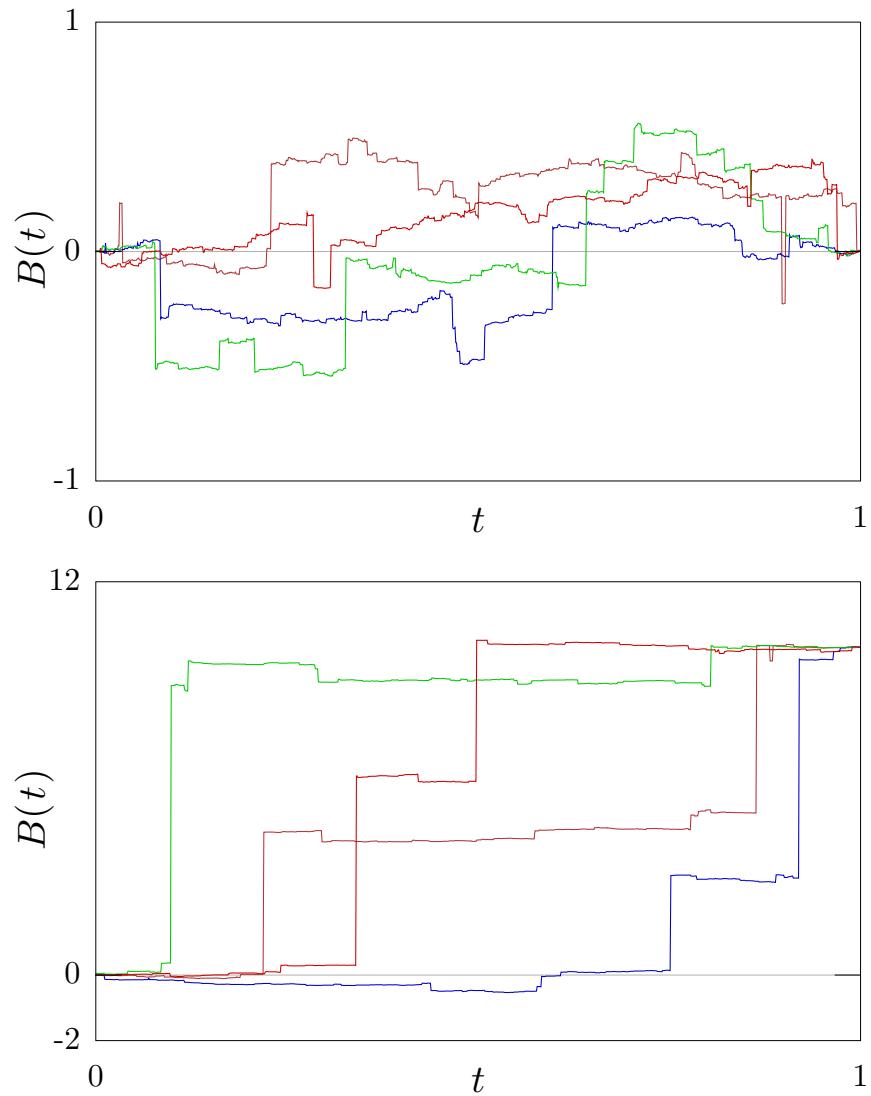
FIGURE 3.5. Simulated Lévy bridges with $\alpha = 1$ and for $L = 0$ (top) and $L = 10$ (bottom). Each bridge is generated through 10 iterations of Eq. (3.12) and has 2^{10} steps.
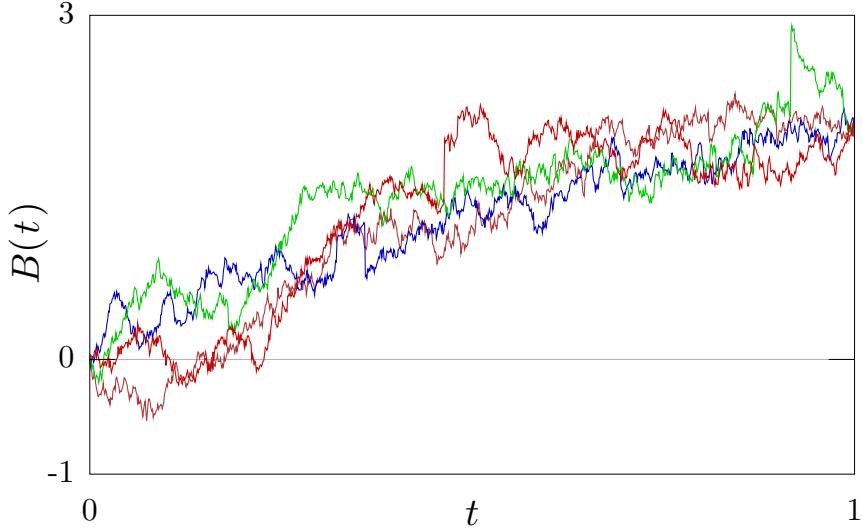
FIGURE 3.6. Simulated Lévy bridges with $\alpha = 1.9$ and for $L = 2$. Each bridge is generated through 10 iterations of Eq. (3.12) and has 2^{10} steps.

3.2.3.1. Cauchy Case

While the midstep distribution cannot be written in closed form for all α, the expression does simplify nicely for the Cauchy limit, making it a convenient example with analytically solvable results. For $\alpha = 1$ the density f_1 is a Cauchy distribution, given by

$$f_1(x;t) = \frac{1}{\pi t \sigma \left(\left(\frac{x}{t\sigma}\right)^2 + 1 \right)}, \tag{3.13}$$

and so we can write Eq. (3.12) as

$$f_1(x_{1/2};T/2|x=L;T) = \frac{f_1(x_{1/2};T/2)\, f_1(L-x_{1/2};T/2)}{f_1(L;T)}. \tag{3.14}$$

A brief look at this distribution shows that it cannot be factored into a shifted and scaled f_1 as was possible for the Gaussian case. Furthermore, while the Gaussian case had a spatial shape that was *independent* of L, the conditional density for $\alpha = 1$ has a shape that is strongly L-dependent. To show this we

69
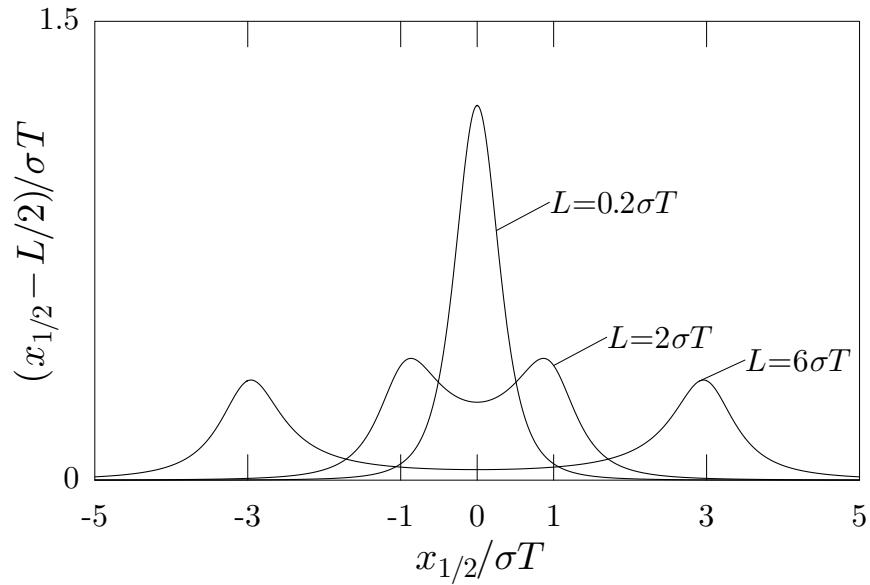
FIGURE 3.7. Plot of the midstep distribution for $\alpha = 1$ for various L.

have plotted Eq. (3.14) for various L in Fig. 3.7. For $L = 0.2$ we can see that the distribution has a single peak at $x_{1/2} = L/2$; this has a simple intuitive interpretation: if a particle travels from $x = 0$ to L in time T, then the most probable intermediate position at $T/2$ is $L/2$. This interpretation was completely correct for the Gaussian case, and we see that it also holds true in the Cauchy case for sufficiently small L. However, we also see the clear dependence of the conditional density shape on L and the splitting of the single peak into two peaks for larger values of L. This demonstrates that there is a bifurcation point L_b in the Cauchy limit such that the midstep distribution for $L > L_b$ becomes bimodal. The bifurcation point L_b can be found by noting that the curvature of the midpoint distribution evaluated at $x_{1/2} = L/2$ must change signs when the maximum becomes a local minimum. This gives us the condition

$$\left(\frac{\partial^2}{\left(\partial x_{1/2}\right)^2} f_1(x_{1/2}; T/2 | x = L_b; T) \right) \Bigg|_{x_{1/2}=L/2} = 0, \qquad (3.15)$$

70

which due to the symmetry of f_1 can be simplified to the condition

$$f_1''(L_\mathrm{b}/2; T/2) f_1(L_\mathrm{b}/2; T/2) - f_1'(L_\mathrm{b}/2; T/2)^2 = 0 \qquad (3.16)$$

where the primes indicate derivatives with respect to the first argument. This equation can be solved analytically, and we find that the bifurcation point for the Cauchy case is given by

$$L_\mathrm{b} = \sigma T. \qquad (3.17)$$

Another way to solve for the bifurcation length is by finding the maxima and minima of the midpoint distribution as a function of L using the condition $\frac{\partial}{\partial x_{1/2}} \left(f_1(x_{1/2}; T/2 | x = L; T) \right) = 0$ and solving for $x_{1/2}$. The 3 real solutions are

$$x_{1/2} = \begin{cases} L/2 & L < t\sigma \\ [L \pm (L^2 - \sigma^2 T^2)^{1/2}]/2 & L > \sigma T \end{cases} \qquad (3.18)$$

where there is a clear transition at the bifurcation point L_b. The solutions in Eq. (3.18) are plotted in Fig. 3.8, showing that the most probable midpoint as a function of L. For $L < L_\mathrm{b}$ we see that the most likely midpoint is $L/2$ as in the Gaussian case. This means that the most likely single step refinement would be into two steps with a size equal to $L/2$. However, for $L > L_\mathrm{b}$ the maximum splits in a pitchfork bifurcation. Further, we can see that for $L \gg L_\mathrm{b}$ the peaks are well separated, with maxima approaching asymptotes $x_{1/2} \sim 0, L$. The interpretation here is that the long jump L tends to break down into two steps: one large step of order L and one small step. This means that a bridge with sufficiently large overall transition length L will tend to maintain this as a *single* jump discontinuity.
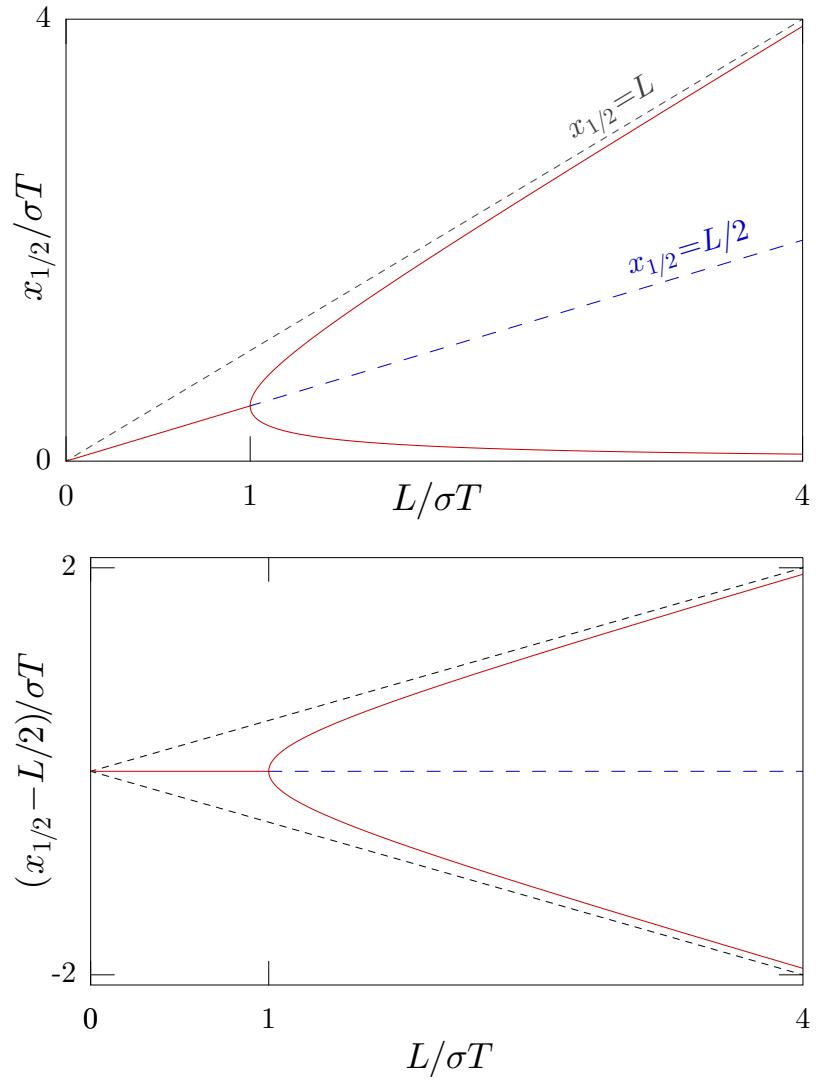
FIGURE 3.8. Plot of the extrema of the midstep distribution (Eq. (3.14)) for the $\alpha = 1$ Cauchy case where maxima are shown in red and minima in blue. For comparison, the extrema for both the unshifted form (top plot) and a symmeterized form (bottom plot) are shown. The bifurcation point is easily visible at $L = t\sigma$.
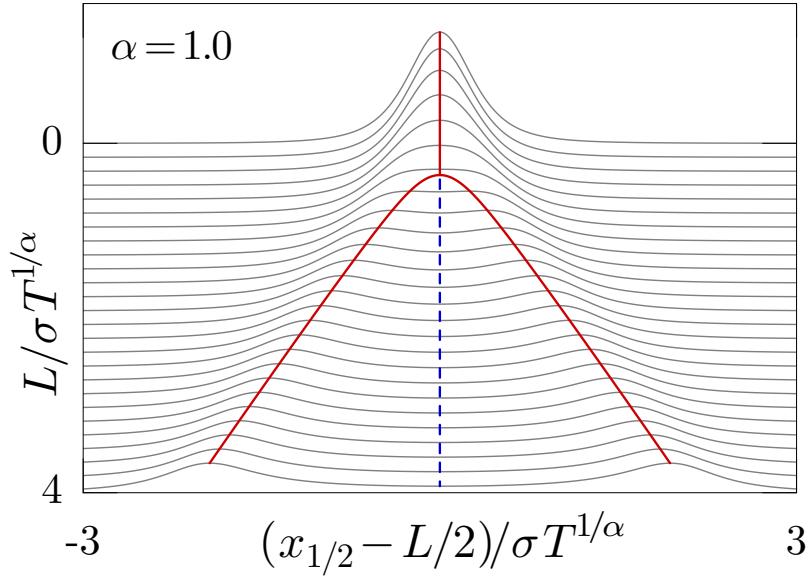
FIGURE 3.9. Bifurcation plot for $\alpha = 1.0$

Another useful way to visualize this transition is with a waterfall chart, as shown in Fig. 3.9, which allows for observation of the change in curvature and the shifts in the positions of the extrema as a function of L.

So far we have only investigated the Cauchy case with $\alpha = 1$, but since bifurcation occurs for $\alpha = 1$ but does not occur for $\alpha = 2$, we know there must be some transition as $\alpha \to 2$. To see this transition we must look at the bifurcation in the general case.

3.2.3.2. General Case

In this section we will explore the midpoint density in the general case, and we will see that similar structural changes to those observed in the Cauchy case occur for all $\alpha < 2$. While we cannot write down analytic expressions for the densities, the general character of the midpoint density can be seen in the waterfall charts in Fig. 3.10. For $\alpha = 1.5$ we see that there is a single pitchfork bifurcation,

similar to the Cauchy case, where two maxima and a minimum are born from a single maximum. We also can see that the point of bifurcation L_b is not simply σT as it was for the Cauchy case, but is somewhat larger.

However, if we take α closer to the Gaussian limit (both $\alpha = 1.99$ and $\alpha = 1.99999$ in Fig. 3.10), we begin to see additional structure beyond what was present for both the Cauchy case and for $\alpha = 1.5$. As L is increased, instead of the maximum splitting in a pitchfork bifurcation, we see that a pair of side peaks form via tangent bifurcations. As L is increased further, these side peaks begin to dominate the central peak, until finally, a central minimum forms through reverse-pitchfork bifurcation. While the transition for these examples is more complicated, we still see that for all α, a sufficiently large L results in similar midpoint behavior, where the density is asymptotically bimodal with well separated peaks.

To analyze this behavior in more detail, we will first look at the how the central peak's bifurcation length L_b generalizes for all $\alpha < 2$. To do this, we will first take a closer look at the more complex bifurcation structure seen for larger α by finding the critical value of α where the more complex structure first arises, as well as characterizing the bifurcation length for where the side peaks appear.

3.2.3.3. Characteristic Bifurcation Length

To characterize the bifurcation point for all $\alpha < 2$, we define the length L_b as the value of L for which the curvature of the midpoint density (Eq. (3.12)) at $x_{1/2} = L/2$ changes sign (Fig. 3.13). That is, we can simply generalize the condition
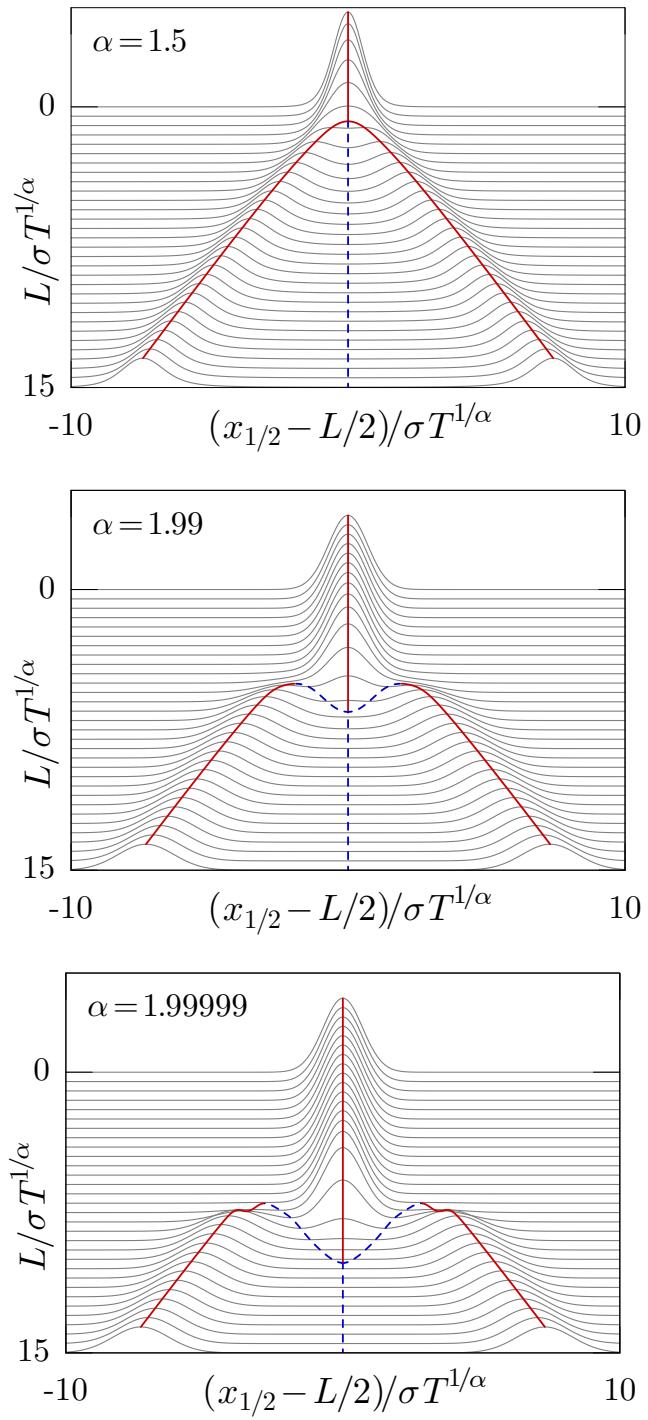
FIGURE 3.10. Variation of the midstep density (Eq. (3.12)) with Lévy index α and arrival point L. Curves highlighting maxima (solid/red) and minima (dashed/blue) are superimposed.
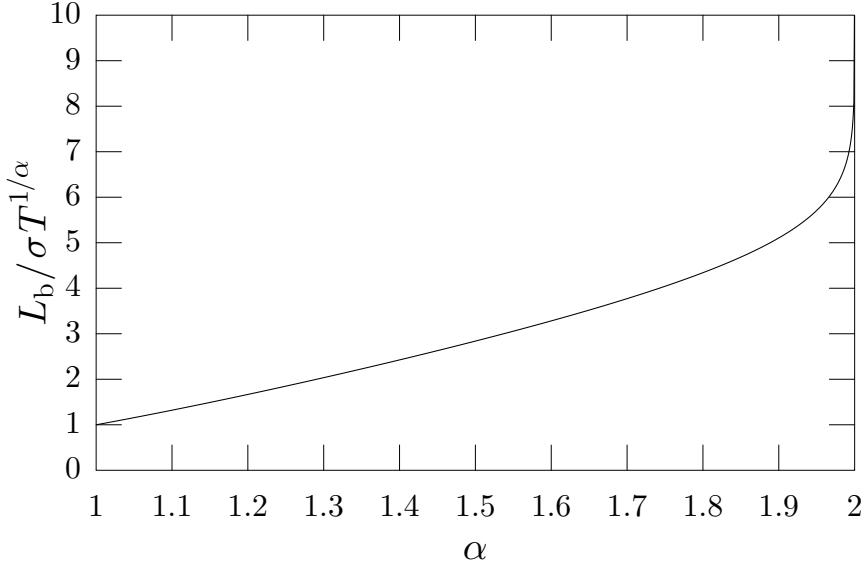
FIGURE 3.11. Bifurcation length L_b plotted as a function of α.

for the Cauchy case in Eq. (3.15) for all α, giving us the condition

$$\left(\frac{\partial^2}{\left(\partial x_{1/2}\right)^2}f_\alpha(x_{1/2};T/2|x=L_b;T)\right)\Bigg|_{x_{1/2}=L/2} = 0, \qquad (3.19)$$

which again due to the symmetry of f_α can be simplified to

$$f_\alpha''(L_b/2;T/2)f_\alpha(L_b/2;T/2) - f_\alpha'(L_b/2;T/2)^2 = 0. \qquad (3.20)$$

This more general condition can be be solved numerically, giving us the bifurcation point L_b as a function of α as shown in Fig. 3.11. The bifurcation point L_b can roughly be considered the transition between "long" and "short" jumps, where for $L \ll L_b$ the displacement is more likely to be a composite of multiple small jumps, while for $L \gg L_b$ the displacement is more likely to be due to a single large jump.

Fig. 3.11 also is important in illustrating the transition to the Gaussian limit as $\alpha \longrightarrow 2$. As we can see, the bifurcation length diverges in this limit, so

76

that for the Gaussian ($\alpha = 2$) case, *any* final step L is effectively a "short step." This matches our intuition for Wiener processes which explicitly forbid any jump discontinuities. To analyze this divergence further, we can make use of the the asymptotic density

$$f_\alpha(x; t = 1) \sim f_2(x; 1) + \delta|x|^{\delta-3}, \tag{3.21}$$

which is valid for large $|x|$ and small $\delta := 2 - \alpha$, and plug it into Eq. (3.12) [92]. After some simplifications, we find that L_b diverges as

$$L_\mathrm{b} \approx \sqrt{2t}\sigma\sqrt{-2W_{-1}\left(-\frac{\pi\delta^2}{2}\right)} \approx 2\sigma\sqrt{t}\sqrt{-\log\left(\frac{\pi\delta^2}{2}\right) + \log\left[-\log\left(\frac{\pi\delta^2}{2}\right)\right]}, \tag{3.22}$$

which can be approximated by

$$L_\mathrm{b} \sim [-4\sigma^2 T \log(\pi\delta^2/2)]^{1/2}. \tag{3.23}$$

The nature of this divergence is somewhat peculiar, as even very close to the Gaussian limit $\alpha = 2$, L_b remains relatively small. This becomes especially noticeable for large α, for instance for $\alpha = 1.99999$ we have $L_\mathrm{b} \approx 10.188473$, even though as $\alpha \to 2$ we must have that $L_\mathrm{b} \to \infty$.

3.2.3.4. Critical α, Alternative Bifurcation Lengths

As we saw in Fig. 3.10, for larger α the midpoint density does not exhibit a simple bifurcation to a bimodal density; rather, the first sign of a transition is the formation of side peaks through tangent bifurcations, then the side peaks grow until they overcome the central peak, and then finally the central peak becomes a
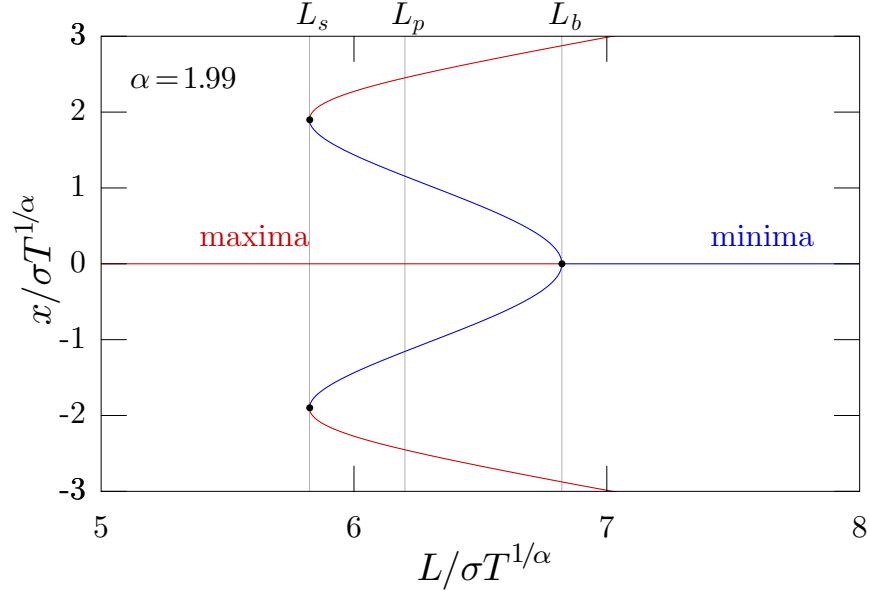
FIGURE 3.12. Bifurcation closeup for $\alpha = 1.99$

local minimum. Thus, to characterize this more complex transition there are three useful values for L along the transition:

1. The side peak bifurcation length L_s, the value of L where the side peaks first originate.

2. The equal peak weight length L_p, the value of L where the side peaks have the same peak probability as the central peak.

3. Characteristic bifurcation length L_b, the value of L where the central maximum becomes a minimum.

These are shown in Fig. 3.12 for $\alpha = 1.99$, and as a function of α in Fig. 3.13 for $\alpha > \alpha_\mathrm{c}$. Interestingly, all the important lengths diverge in a similar manner as Eq. (3.23), so we use L_b as it is the most straightforward to calculate.

To find the critical value α_c for which these transitions first appear, it is helpful to consider the diagram explicitly showing the shoulder bifurcation points
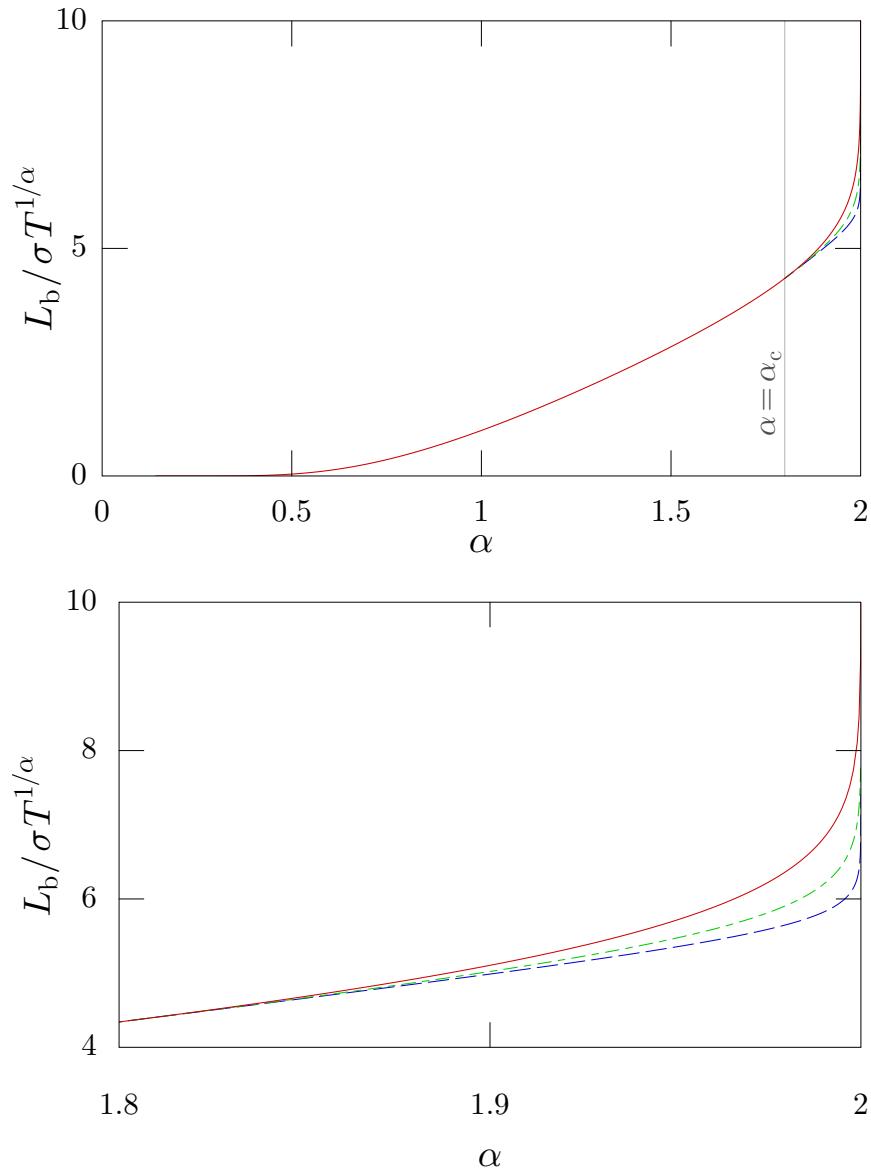
FIGURE 3.13. Variation of boundaries between "small" and "large" steps with α. Curves indicate bifurcation lengths L_b for which the center of the midstep density has vanishing curvature (red/solid), the midstep distribution develops side peaks (blue/dashed), and side peaks are equal in height to the center peak (green/dot-dashed). Inset: magnified view for $\alpha > \alpha_\mathrm{c}$.

(L_s) and the reverse pitchfork bifurcation (L_b), as shown in Fig. 3.12. As α is lowered, the difference between the L_s and L_b will decrease, until they finally meet when $\alpha = \alpha_c$ and merge together. The condition for this type of transition is that both the second and fourth derivatives vanish:

$$\left(\frac{\partial^2}{\left(\partial x_{1/2}\right)^2} f_{\alpha_c}(x_{1/2}; T/2 | x = L_\mathrm{b}; T) \right) \Bigg|_{x_{1/2} = L_\mathrm{b}/2} = 0, \tag{3.24}$$

$$\left(\frac{\partial^4}{\left(\partial x_{1/2}\right)^4} f_{\alpha_c}(x_{1/2}; T/2 | x = L_\mathrm{b}; T) \right) \Bigg|_{x_{1/2} = L_\mathrm{b}/2} = 0. \tag{3.25}$$

Numerically solving these equations for the critical value gives

$$\alpha_\mathrm{c} \approx 1.7999233 \tag{3.26}$$

which is curiously close to, but *not* equal to 1.8.

3.2.4. Conditioned Sampling

The importance of the bifurcation length can be seen by analyzing Lévy bridges of different lengths. We will see that above the bimodal transition ($L \gg L_\mathrm{b}$), the typical conditioned history is biased towards containing only a single large event, while below the transition ($L \ll L_\mathrm{b}$) the tendency is towards a composite of smaller events.In Fig. 3.14 we show conditioned Lévy processes for various L and α, with the choices of L are scaled in terms of the bifurcation length L_b to illustrate the transition.

For the $\alpha = 1.9$ the bias towards preserving long jumps when L exceeds the bifurcation length is easily evident. For $L = 1.5 L_\mathrm{b}$ each sample path has
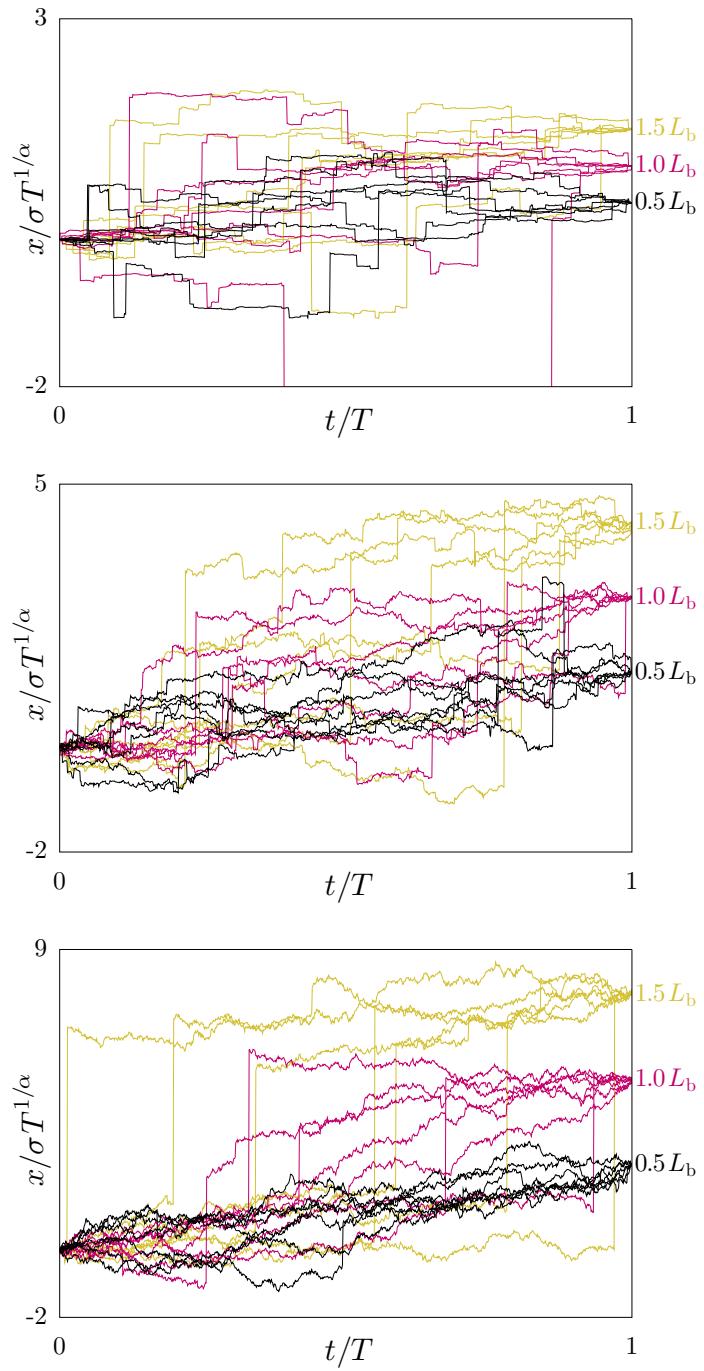
FIGURE 3.14. Typical sample paths of Lévy bridges for $\alpha = 1.9$, illustrating the qualitative transition with L. Each path was generated through 10 recursive subsamplings from the midpoint distribution (Eq. (3.12)).
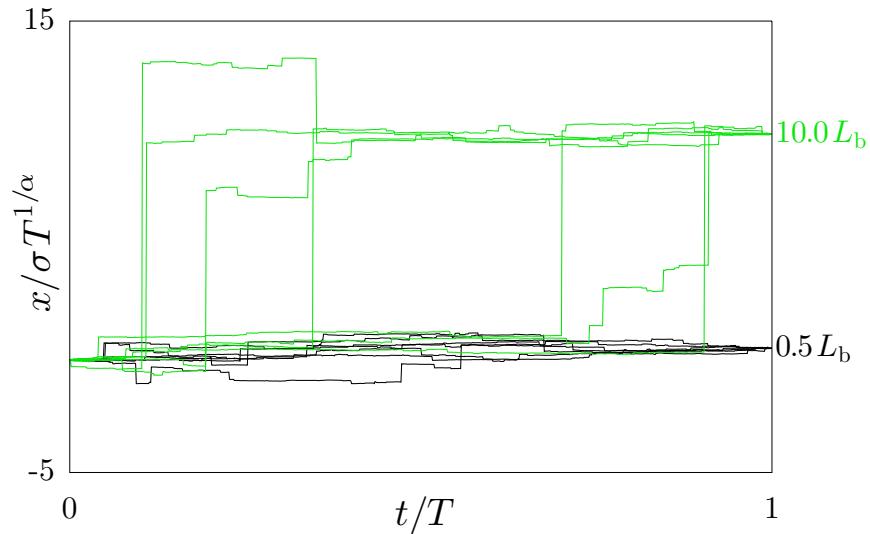
FIGURE 3.15. Typical sample paths of Lévy bridges for $\alpha = 1.0$. Each path was generated through 10 recursive subsamplings from the midpoint distribution (Eq. (3.12)).

a single long jump, while for $L = 0.5L_{\mathrm{b}}$ nearly every sample path looks akin to Brownian motion, with a just a handful of larger jumps. The reason for this behavior is that in each stage of sampling of the midstep distribution Eq. (3.12), if the displacement L exceeds L_{b} then the entire length L is likely to be preserved as a single jump in one of the two subintervals. Then, upon subsampling of the subinterval containing this jump, L_{b} is effectively smaller due to the smaller time interval. This means the jump length L tends to exceed L_{b} by an ever increasing margin, making it progressively less likely to be split into smaller jumps.

The case for $\alpha = 1$ and $\alpha = 1.5$ is not quite as clear. While $L > L_b$ indicates that refinement is likely to have a large jump of order L persist under refinement, it does not preclude the possibility that additional jumps may form, which ultimately make the qualitative difference between $L < L_{\mathrm{b}}$ and $L > L_{\mathrm{b}}$ less distinct when L is close to L_{b}. Ultimately through, the same behavior still holds as long as L is much greater than L_{b}, as can be see in Fig. 3.15.

3.2.5. Jump Detection

As we discussed previously, the conditioned subsampling shows that the bifurcation length L_b is intimately related to the long jumps of an α-stable process. We found that an observed final displacement $|x(T)| \gg L_b$ most likely corresponds to a single, similarly large jump discontinuity. Meanwhile, a smaller final displacement $|x(T)| \sim L_b$ is more likely to be a composite event comprising multiple smaller jumps. Perhaps most surprising is not this behavior itself, as Lévy processes are well known to have long jumps, but that it implies the existence of a distinct length scale L_b.

Through the Lévy–Khintchine representation of Lévy processes (Section 2.2.1), as well as the pure power law jump distribution for the stable Lévy processes (Section 2.2.2), we saw that the stable Lévy processes are inherently scale invariant. It thus seems like a contradiction for there to be an intrinsic length scale L_b. So how is this possible? The answer is subtle: in a sense, conditioning introduces a time scale T, which in turn induces a natural length scale $\sigma T^{1/\alpha}$. If this were the scale of the bifurcation length, the result would seem underwhelming as it came from natural consequence of the conditioning. However, the bifurcation length L_b *differs* from $\sigma T^{1/\alpha}$, and not by a little: L_b can be arbitrarily large as it diverges as $\alpha \to 2$. This suggests that it truly is a distinct scale that captures a different quality than the natural scale $\sigma T^{1/\alpha}$, namely the scale of the large visible jumps present in the Lévy process.

To show this, we can consider comparing L_b to the individual displacements of (finite) sample paths. As any given sample path is effectively a collection of bridges between a set of known values, the midpoint density gives insight into

the likely behavior of the process in between sampled values, and the bifurcation length L_b becomes useful criteria for detecting "long" jumps.

To demonstrate the usefulness of this L_b criteria, it's useful to consider the other natural alternative criteria. While the typical natural criteria for detecting "long" jump outliers is the standard deviation, as this diverges for Lévy processes we instead use the natural scale $\sigma T^{1/\alpha}$. In Fig. 3.16 we show a Lévy process with with highlights on all increments that have $\Delta x > \sigma \Delta t^{1/\alpha}$. Interestingly, the majority of steps are highlighted, and it does not seem particularly useful for identifying extreme events.

On the other hand, if we highlight increments that exceed the bifurcation length L_b as in Fig. 3.16, we find that the intuitive "jumps" are clearly highlighted, while the other motion appears almost Brownian in character. This is somewhat peculiar, as the Lévy–Itō decomposition (Section 2.2.1) shows that stable Lévy processes are distinct from the Wiener process. However, while the processes are distinct, analysis using the bifurcation length has led to more grounded reasons for the strong similarities that arise between Wiener processes and Lévy processes.

3.2.6. Corrected Stretched Lévy Bridges

In Section 3.2.1 we saw that stretched Lévy Bridges did not satisfy the properties we want in a Lévy bridge, and generated incorrect statistics. However, since we have a criteria that can detect if a jump is "long", it becomes possible to deal with excessive stretches through a rejection algorithm. The basic idea is to generate unconditioned Lévy bridges until one of them is sufficiently close to the intended final position, and then stretch it as before. The key is that the
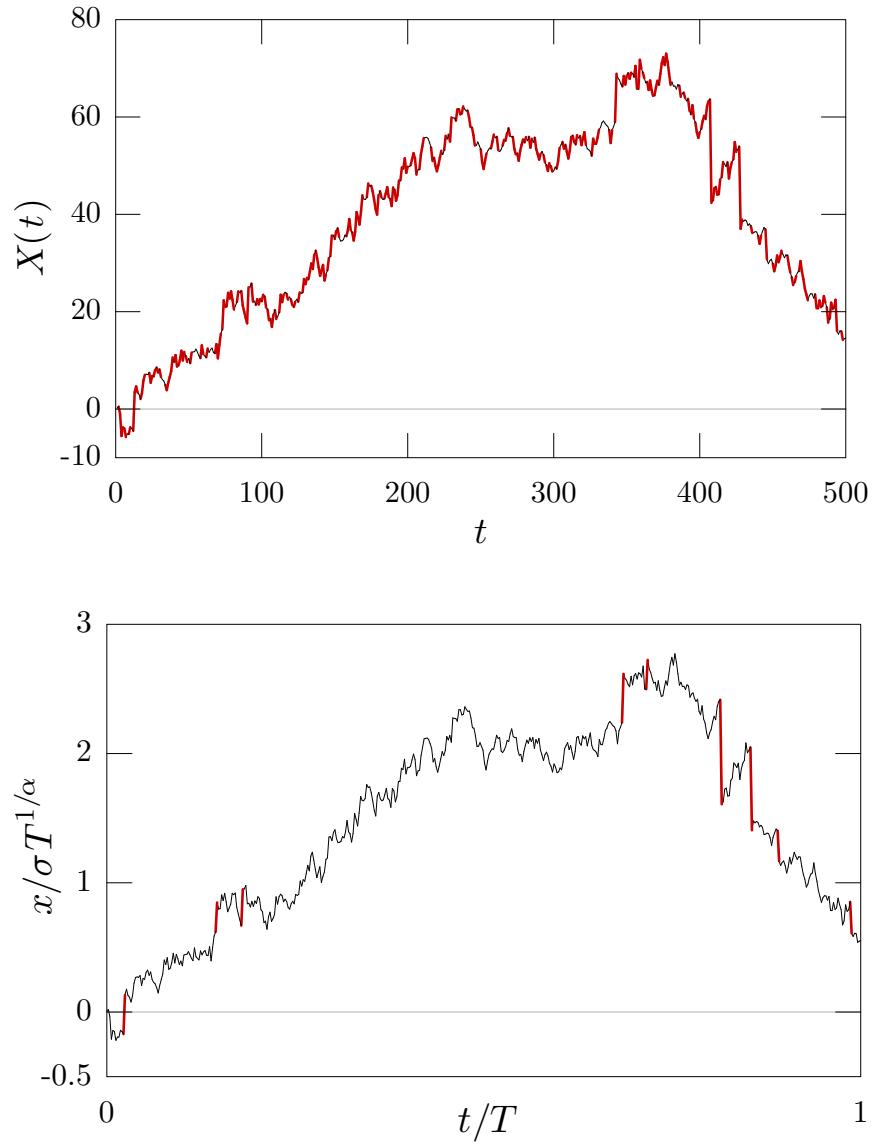
FIGURE 3.16. Sample path of an unconditioned α-stable Lévy process, $\alpha = 1.9$. Simulated path has time steps $\Delta t = T/500$ For the top plot, increments $\Delta x > \sigma(\Delta t/T)^{1/\alpha}$ emphasized (bold/red). For the bottom plot, increments $\Delta x > L_{\mathrm{b}}(\Delta t/T)^{1/\alpha}$ emphasized (bold/red).

bifurcation length can serve as the criteria for "sufficiently close", allowing for more efficient stretched bridge generation.

The algorithm is as follows: take an unconditioned Lévy sample path L', and choose some threshold L_{thresh}. If the difference between the final position of the generated path and the intended final position is smaller than the threshold, that is $|L'(T) - L| < L_{\text{thresh}}$, then the bridge is accepted; otherwise it is rejected and a new candidate bridge is generated. In the limit that $L_{\text{thresh}} \to 0$ the bridge is exact, though such a bridge would take arbitrarily long to generate. A naïve approach for this rejection algorithm would use $\sigma\Delta t^{1/\alpha}$ as the scale for L_{thresh}, but using the α-dependent bifurcation length L_{b} as the scale for L_{thresh} is much more efficient and should still produce accurate Lévy bridges. This method has been confirmed to work through numerical simulations [93].

3.2.7. Conditioned First Passage

In this section we analyze the first passage distributions for conditioned Lévy processes. As we have seen, conditioned Lévy bridges have a particularly sensitive transition as $\alpha \longrightarrow 2$, which we will see is especially true for first-passage times.

An intuitive picture of the conditioned first-passage time follows from the qualitative appearance of the sample paths for $L = 1.5L_{\text{b}}$ in Fig. 3.14. In this figure a long jump is consistently present among the paths, but not at any particular time. This can be regarded as an outcome of recursively sampling the midpoint density [Eq. (3.12)]. For $L \gg L_{\text{b}}$, a large jump likely persists under sampling iterations, but due to the symmetry of the midpoint distribution, the jump is equally likely to be associated with any time subinterval. Since the first-

passage time is likely due to the long jump, the first-passage time should be uniformly distributed.

Fig. 3.17 confirms this intuition with simulations of the first passage density. For $L = 2L_{\mathrm{b}}$ the first passage density is indeed uniform. A small change from $\alpha = 1.99999$ to the Gaussian case yields a remarkably different distribution, one that is approximately Gaussian and centered at $t \approx T/2$. The Gaussian result follows intuitively from since the most likely bridges in this regime are concentrated around the ballistic path to the endpoint.

For a smaller overall jump ($L = 0.1L_{\mathrm{b}}$), the first-passage-time densities in the $\alpha = 1.99999$ and Gaussian cases match closely. This is consistent with the observation that for $L \ll L_{\mathrm{b}}$, the conditioned Lévy bridges are qualitatively similar to Brownian bridges. Nevertheless, the rare but important long jumps generate remarkably non-Gaussian behavior, even close to the Gaussian limit. This result also suggests that boundary-crossing statistics can be a particularly sensitive probe for the presence of anomalous diffusion, which is investigated more in Chapter IV.

3.2.8. Inferring α from a Sample Path

Calculating the value of α for a sample path is a useful way to characterize heavy tailed random walks or anomalous diffusive motion. For collections of many sample paths, there are two common approaches to calculating the exponent from the distributions of particle positions. First is the dynamic approach, where the width of the distribution will grow as $\sim t^{1/\alpha}$, which can be measured by fitting a distribution, calculating the FWHM, or calculating measures of self-similarity. The other approach is by analyzing the shape of the distribution, as the tails of the distribution scale as $\sim x^{-\alpha-1}$.
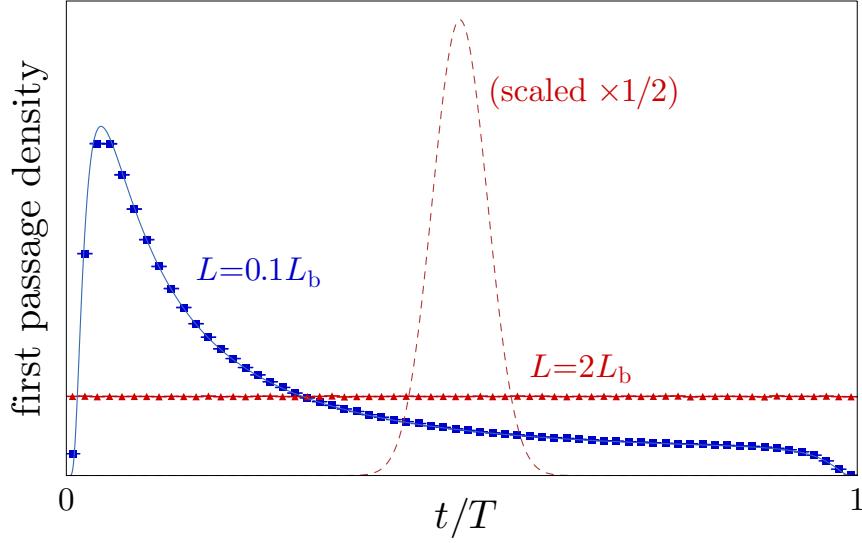
FIGURE 3.17. Simulated conditioned first passage time distributions for $\alpha = 1.99999$ and $d = L/2$ are shown for $L = 0.1L_\mathrm{b}$ (blue/squares) and $L = 2L_\mathrm{b}$ (red/triangles). Exact densities for $\alpha = 2$ [1] for the same L values are shown for comparison in each case (blue/solid and red/dashed, respectively). Simulations averaged 10^7 paths, with $\Delta t = 2^{-14}T$.

On the other hand, it is also possible to analyze individual trajectories. For a given trajectory with N steps Δx_i occurring every Δt, the most straightforward method to infer the value of α is by using maximum-likelihood estimation. For a uniform prior $\alpha \in (1, 2)$ the maximum likelihood is found by maximizing the function [94]

$$P(\alpha|\Delta x_i \Delta x_N) \propto P(\Delta x_1|\alpha) \ldots P(\Delta x_N|\alpha) \tag{3.27}$$

which can be written as

$$P(\alpha|\Delta x_i) \propto f_\alpha(\Delta x_1; \Delta t) \ldots f_\alpha(\Delta x_N; \Delta t). \tag{3.28}$$

Typically this is calculated via the log likelihood

$$\log(P(\alpha|\Delta x_i)) = \log(f_\alpha(\Delta x_1; \Delta t)) + \ldots + \log(f_\alpha(\Delta x_N; \Delta t)) + C, \tag{3.29}$$
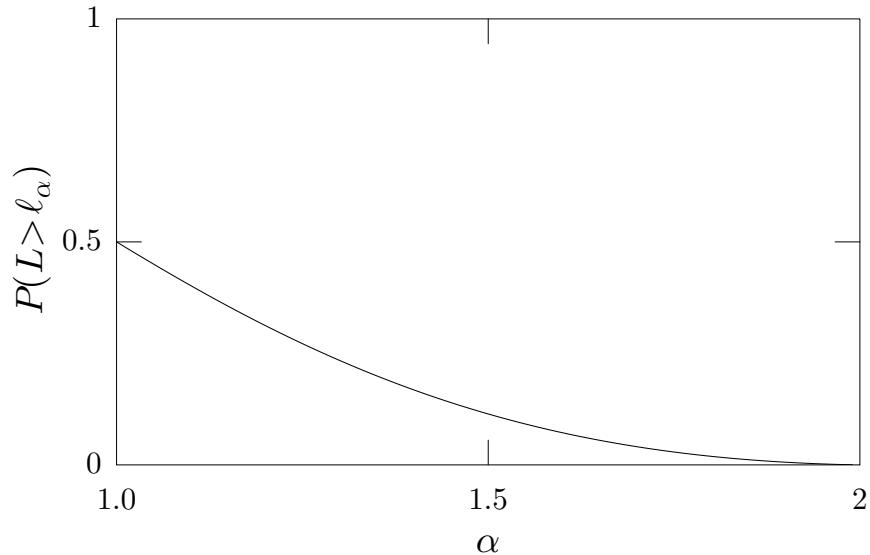
FIGURE 3.18. Long jump probability as a function of α.

where C is the log of the proportionality constant which is not relevant for the maximization. While this approach is straightforward and well grounded in theory, the downside is that it requires many evaluations of f_α which can be computationally costly.

In this section we infer α through an alternative method by analyzing the sample path using the concept of "jumps". We look at each step Δx and consider it a jump if $\Delta x > L_{\mathrm{b}}$, which can be used to create the jump rate statistic j_α defined as the number of jumps in the sample path over the total number of steps.

This proposed jump rate can be compared to the theoretical jump rate J_α, which is defined as

$$J_\alpha := \int_{L_{\mathrm{b}}}^{\infty} dx \, f_\alpha(x; t), \tag{3.30}$$

and this function is plotted for reference in Fig. 3.18.

We must now numerically solve for the α such that

$$J_\alpha = j_\alpha, \tag{3.31}$$

89
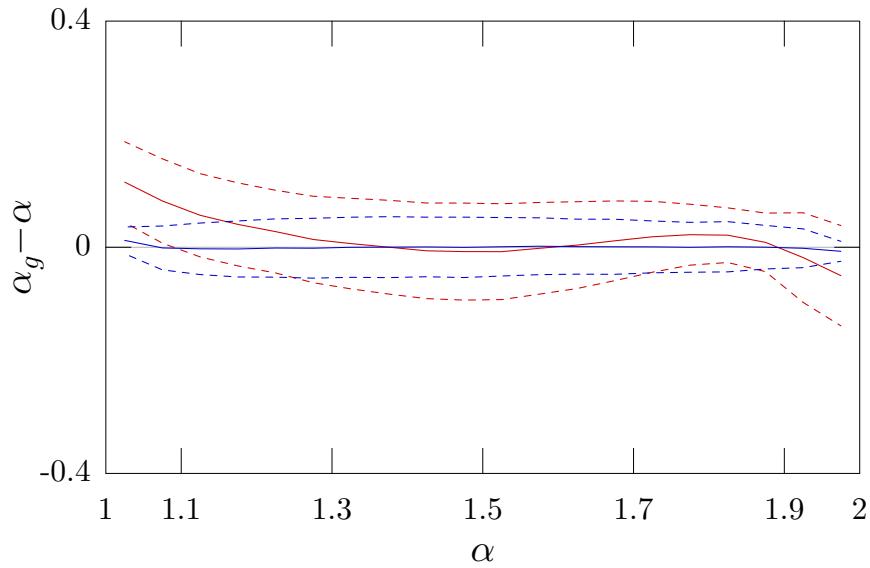
FIGURE 3.19. Effectiveness of the jump rate method (red) of alpha estimation when compared to the maximum-likelihood method (blue). Mean differences between the estimated and true parameter are shown with solid lines, while dashed lines indicate the standard deviation.

which should hold for the correct choice of α when $N \to \infty$. This approach is less computationally intensive since there is no need to evaluate f_α at n points for each value of α considered. Instead, one only needs to evaluate the inequality $\Delta x_i > L_{\text{b}}$ at n points and compare it to J_α for each α considered. For this approach, only two lookup tables are needed, one for L_{b} and one for J_α, compared to a lookup table for every f_α.

Results from a test of the jump based inference method compared to the maximum-likelihood method are shown in Fig. 3.19; while the maximum-likelihood method still has superior accuracy, the jump method still has reasonable results.

CHAPTER IV

ANOMALOUS DIFFUSION IN SISYPHUS COOLING

Optical molasses refers to a laser cooling technique for atoms isolated in vacuum, where counter-propagating laser fields create a region of space where the motion of atoms is strongly damped, as if the atoms were embedded in viscous molasses. Laser cooling of atoms in optical molasses is a complex process that can involve multiple cooling mechanisms, including Doppler cooling, Sisyphus cooling, and polarization-gradient cooling. Also present are lattice effects of the cooling beams, as well as spontaneous emission, where photons are released in a random direction, causing a heating effect. Atom losses, caused by molecule formation, spontaneous emission to non-resonant energy levels, or simply diffusion out of the cooling region, can also play a significant role. Altogether these effects drive a complex diffusion process.

Despite the complexity, under some reasonable choices of parameters it is possible to tune the optical molasses into a regime where the Sisyphus force dominates and can be approximated as a simple stochastic differential equation (Section 4.1). In this approximation, the system has a steady-state momentum distribution with heavy tails known as a Tsallis distribution (Section 4.2), indicating the possibility of anomalous diffusion. When the optical lattice depth is sufficiently deep, the power laws tails of the momentum distribution have only a minor impact, and the particle position can asymptotically be represented by a simple diffusion equation (Section 4.3). But, for sufficiently shallow optical lattices, the diffusion constant diverges, which leads to Lévy walks behavior and the onset of anomalous diffusion (Section 4.4).

To detect the transition between these regimes, previous experiments have made use of cloud-expansion experiments (for instance, Refs. [21, 30, 38]). However, as mentioned in the Introduction, this led to some conflict between experimental and theory thought to be in part due to atom losses [39]. To mitigate the effect of losses, single-atom experiments are particularly useful, and have also been demonstrated to be effective at detecting the onset of anomalous diffusion [37].

This chapter begins by introducing the semiclassical Sisyphus cooling force in Section 4.1 as well as some of its known properties, including the steady state momentum distribution (Section 4.2) and diffusive behavior in both the Brownian regime (Section 4.3) and the Lévy regime (Section 4.4). Original work begins in Section 4.5 with a formula that characterizes the extent of power-law scaling in the Lévy regime. Next, in Section 4.6, we propose a scheme for detecting the onset of anomalous diffusion through the use of boundary-crossing statistics. This leads to one of our primary results, where through simulations we demonstrate the existence of a peak in boundary-crossing statistics that is connected to the Lévy behavior of the underlying process. To show this connection, we derive scaling relations that characterizing the peak. These scaling relations also serve as a useful starting point for future single-atom experiments.

4.1. Sisyphus Cooling

Sisyphus cooling is a laser cooling mechanism that arises due to coherent pumping between ground-state Zeeman sublevels. The details of Sisyphus cooling are complex, but we will summarize the basic mechanism (for more details, see Refs. [27, 95]). For cooling in one dimension, Sisyphus cooling can be

92

created by setting up two counter-propagating beams with perpendicular linear polarizations (lin⊥lin configuration). The total electric field created by these counter-propagating beams has an elliptical polarization that varies in space. This polarization gradient in turn causes oscillations in the ground state energies. For laser frequencies ω detuned below the atomic frequency ω_0 by an amount Δ, optical pumping will occur from the upper ground state to the lower ground state. This alone would reduce the energy of the atom by an amount $\sim \hbar\Delta$. However, this cycle can be repeated: for an atom in motion along the polarization gradient, if it has already been pumped into the lower ground state, it will effectively be moving up a potential gradient due to the oscillations in the ground state energies. Moving up the potential gradient has a corresponding loss in kinetic energy. Finally, at some point the atom will have moved far enough such that the lower ground state becomes the upper ground state, and the atom will again be optically pumped from the (now) upper ground state into the (now) lower ground state. Roughly speaking, kinetic energy of order $\hbar\Delta$ will be dissipated each cycle, with the cycle period limited by both the optical pumping rate and the time to climb the potential well (related to laser wavelength and the velocity of the atom).

While we have only depicted a simple model for Sisyphus cooling, the full treatment is an inherently quantum-mechanical process. However, it has been shown in Ref. [32] that fully quantum treatments are well approximated by the semiclassical models discussed in Ref. [27], which has also been confirmed with comparisons to fully quantum models [31].

The formalism we use to describe Sisyphus cooling follows closely from Ref. [96], which gives the semiclassical momentum evolution by the pair of SDEs

$$dp = -\frac{\bar{\alpha} p}{1 + (p/p_c)^2} + \sqrt{2D(p)}\, dW(t), \tag{4.1}$$

$$dx = \frac{p}{m}\, dt. \tag{4.2}$$

The parameters $\bar{\alpha}$ is the Sisyphus friction coefficient, p_c is a characteristic momentum for the system, and

$$D(p) = D_1 + \frac{D_2}{1 + (p/p_c)} \tag{4.3}$$

is the momentum dependent diffusion parameter. The parameters $\bar{\alpha}$, p_c, D_1, and D_2 can be directly related to physical parameters as in Ref. [31], but there is a more convenient form which we will derive momentarily. First, it turns out that the asymptotic behavior for this system does not depend on D_2, so for simplicity we can set $D_2 = 0$ [31]. We will now transform into a non-dimensionalized form by defining new coordinates via

$$p = p'\, p_c \tag{4.4}$$

$$t = t'/\alpha \tag{4.5}$$

$$x = x'p_c/m\bar{\alpha}, \tag{4.6}$$

where the primes denote dimensionless variables. With these substitutions and dropping the primes for simplicity, the equations of motion become

$$dp = -\frac{p}{1+p^2} + \sqrt{2D}\,dW(t) \tag{4.7}$$

$$dx = p\,dt, \tag{4.8}$$

where we have defined the parameter D by

$$D = \frac{D_1}{\bar{\alpha}p_c^2}. \tag{4.9}$$

The relation of D to physical parameters is given by

$$D = C\frac{E_R}{U_0} \tag{4.10}$$

where the recoil energy $E_R = \hbar^2 k^2/2m$, U_0 is the optical potential depth, and the dimensionless constant C depends on the particular[1] atomic transition [39].

It is important to note that this model has several implicit approximations. First, we are ignoring the effect of the potential lattice induced by the counter propagating beams. This is valid as long as the atom's kinetic energy is large compared to the depth of the lattice wells. Second, we have neglected the Doppler cooling force. This is valid for large detunings and slower atom velocities. However, it is important to note that Doppler cooling will become relevant for sufficiently large velocities. This sets an upper limit on the velocity, beyond which the model in Eq. (4.7) begins to break down [31].

---

[1]Values for the constant C vary, but tend to be of of order 10. For a $J_g = 1/2 \rightarrow J_e = 3/2$ transition $C = 12.3$. [31, 39].

Simulations of Eq. (4.7) are shown in Fig. 4.1 for various values of D. Qualitatively D has a significant effect: for $D = 0.01$ we can see that the trajectories are akin to Brownian motion, while for $D = 0.19$ the motion is mostly Brownian in appearance, though there are occasionally some longer jumps present. For $D = 0.3$ and $D = 0.5$ the jumps are obvious and tend to dominate the motion.

4.2. Steady-State Momentum Distribution

A straightforward statistic that characterizes the Sisyphus force is the steady-state momentum distribution. Following Refs. [28, 30], a convenient way to find the steady state momentum distribution is by transforming Eq. (4.7) to the Fokker–Plank equation. Recalling from Section 2.1.6 that any stochastic differential of the form

$$dx = g(x,t)dt + h(x,t)dW, \tag{4.11}$$

has a corresponding Fokker-Plank equation given by

$$\frac{\partial}{\partial t} f(x,t) = -\frac{\partial}{\partial x}\left(g(x,t)f(x,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(h(x,t)^2 f(x,t)\right). \tag{4.12}$$

To find the steady state momentum distribution, we set $\frac{\partial}{\partial t} f(p,t) = 0$, $g(p,t) = -\frac{p}{1+p^2}$, and $h(p,t) = \sqrt{2D}$. Rewriting Eq. (4.12) we have

$$\frac{\partial^2}{\partial p^2} f(p) = \frac{2}{2D}\frac{\partial}{\partial p}\left(g(p)f(p)\right), \tag{4.13}$$

where we have dropped the t dependence of both $f(p,t)$ and $g(p,t)$. Integrating by p gives

$$\frac{\partial}{\partial p} f(p) = \frac{2}{2D}\left(g(p)f(p)\right) + C_0 \tag{4.14}$$
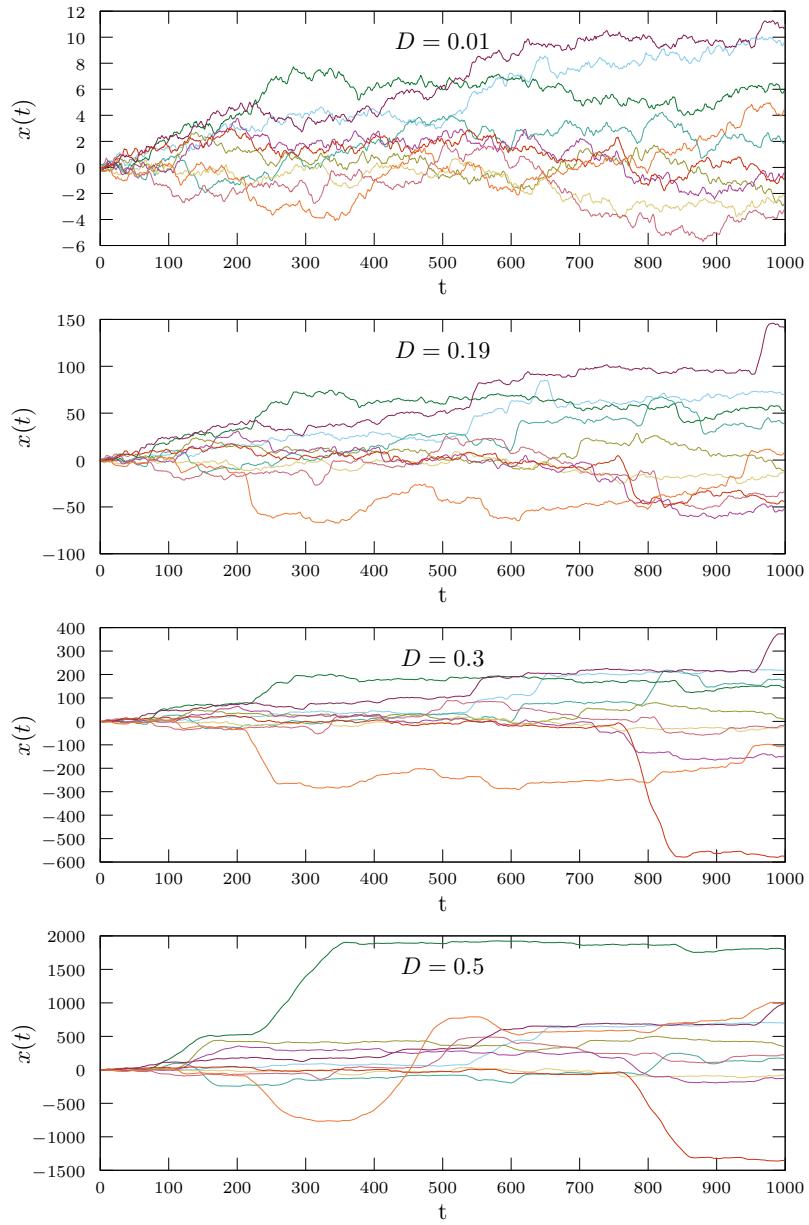
FIGURE 4.1. Simulated trajectories of Eq. (4.7) for various values of the diffusion constant D.

for some integration constant C_0. The force $g(p)$ vanishes at $p = 0$, and since the distribution $f(p)$ is an even function, $\partial_p f(p)$ must also vanish at $p = 0$, thus forcing $C_0 = 0$. Integrating once more, we have

$$\int_{f(0)}^{f(p)} df \frac{1}{f} = \frac{2}{2D} \int_0^p dp' g(p'), \qquad (4.15)$$

which has the solution

$$f(p) = f(0) \exp\left(\frac{2}{2D} \int_0^p dp' g(p')\right). \qquad (4.16)$$

Performing the final integral

$$\int_0^p dp' g(p') = -\int_0^p dp' \frac{p}{1+p^2} = -\frac{1}{2} \ln\left(1 + p^2\right), \qquad (4.17)$$

we have

$$f(p) = f(0) \exp\left(-\frac{1}{2D} \ln\left(1 + p^2\right)\right) = f(0) \left(1 + p^2\right)^{-1/2D}. \qquad (4.18)$$

Finally, $f(0)$ is just a normalization constant, which is given by

$$f(0) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{1}{2D})}{\Gamma(\frac{1}{2D} - \frac{1}{2})} \qquad (4.19)$$

Putting these together, we have the steady state momentum distribution

$$f(p) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{1}{2D})}{\Gamma(\frac{1}{2D} - \frac{1}{2})} \left(1 + p^2\right)^{-1/2D}. \qquad (4.20)$$
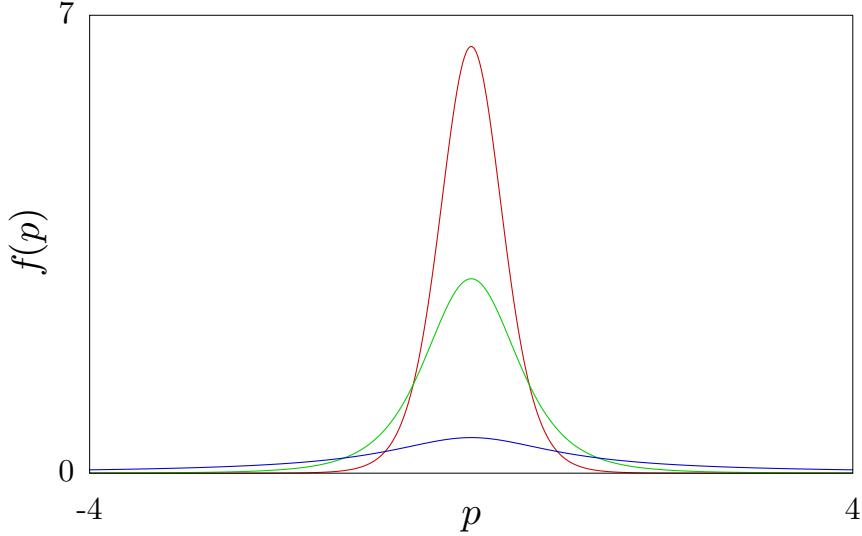
FIGURE 4.2. Plot of Tsallis distribution in Eq. (4.20) with $D = 0.1$ (red), $D = 0.2$ (blue), and $D = 0.6$ (green).

This is known as a Tsallis distribution. This distribution is normalizable for $0 < D < 1$ and is plotted in Fig. 4.2 for various values of D. We can see that for larger values of D the distribution has particularly heavy tails.

The momentum distribution can also be helpful for interpreting the momentum diffusion parameter D, which relates the relative magnitude of the damping and heating effects. In Fig. 4.3 we plot the fraction of particles that have momentum $|p| > 1$, which shows that D roughly tracks the fraction of particles in the steady state momentum distribution with $|p| > 1$. This is useful as the threshold $p \sim 1$ is approximately where the Sisyphus force becomes nonlinear. For smaller D, linear damping effects dominate, causing a larger fraction of particles to be in the linear $F(p) \sim -p$ regime of the force. Conversely, larger D is dominated by the heating effects, causing more particles to be affected by the $F(p) \sim -1/p$ regime of the force.

While we derived the Tsallis distribution using the Fokker–Planck equation, it is also possible to do so via the Itō SDE. This is done through calculating all

99
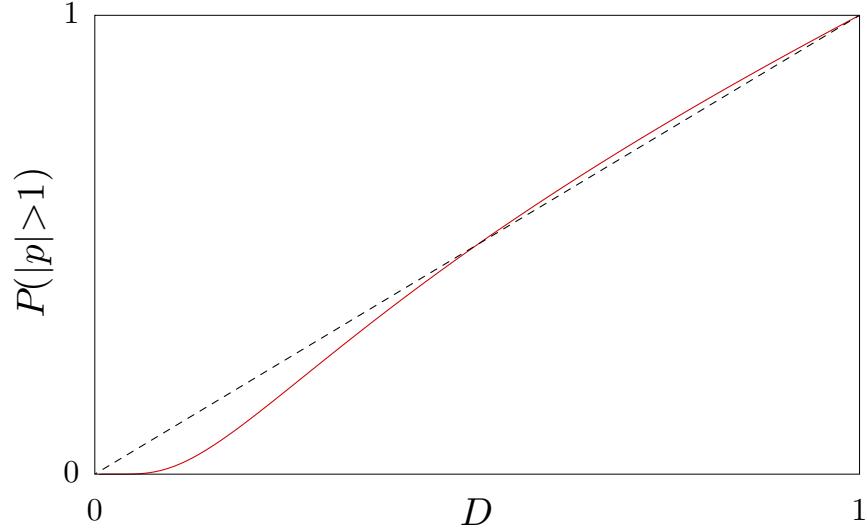
FIGURE 4.3. The solid red line shows the fraction of particles with $|p| > 1$ in the steady state distribution as a function of the diffusion parameter D. The dashed black line is a guide line with slope 1.

of the moments of the distribution as shown in Appendix A.1, and then using an appropriate inversion as shown in Appendix A.2. Interestingly, having both derivations yields an inversion formula for a particular hypergeometric function.

4.3. Brownian Regime

For certain values of the momentum diffusion parameter D, through a process of adiabatic elimination of the momentum variable, it is possible to show that position distribution is described asymptotically by a diffusion equation, with a diffusion constant given by [30, 39]

$$K_2 = \frac{1}{D\mathcal{Z}} \int_{-\infty}^{\infty} dp \, e^{V(p)/D} \left[\int_p^{\infty} dp' \, e^{-V(p)/D} p' \right]^2 , \tag{4.21}$$

where $\mathcal{Z}$ is given by

$$\mathcal{Z} = \sqrt{\pi} \, \Gamma\left(\frac{1-D}{2D}\right) \Big/ \Gamma\left(\frac{1}{2D}\right) , \tag{4.22}$$

and $V(p)$ is an effective potential for the Sisyphus force in momentum space, given by

$$V(p) = \int dp F(p) = (1/2)\ln(1 + p^2).\qquad(4.23)$$

This integral is finite for $0 < D < 1/5$, and in this case K_2 has the simplified expression

$$K_2 = \frac{D(4D - 1)}{(2D - 1)(3D - 1)(5D - 1)}.\qquad(4.24)$$

We can see that this diverges as $D \to 1/5$, indicating the onset of the Lévy regime and the failure of the standard diffusion equation.

In Fig. 4.4 we show simulations of atoms cooled via the Sisyphus force compared to Gaussian diffusion simulation. For $D = 0.01$ the position distributions for both simulations are nearly identical. However, for $D = 0.19$ we see the presence of power laws in the tails of the position distribution for Sisyphus cooling. While this is in the regime of Gaussian asymptotics, we see a slow convergence to a Gaussian distribution similar to that seen in the truncated Cauchy case (Section 2.2.3.1). This also seems consistent with the "jumps" visible in Fig. 4.1 for $D = 0.19$.

4.4. Lévy Regime

When the diffusion constant diverges for $D > 1/5$, the Sisyphus cooling behavior is remarkably different from the Brownian regime and leads to long correlated motion that prevents the development of asymptotic Gaussian behavior. A qualitative explanation is that when D is large enough, the atom is more likely to escape out of the linear regime of the Sisyphus force. For large enough p, the asymptotic force goes as $-1/|p|$, with the magnitude of the damping diminishing
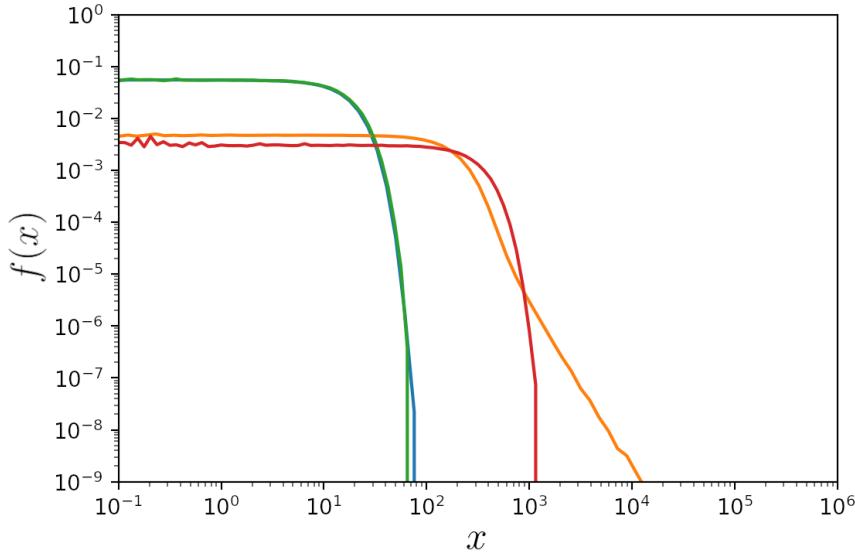
FIGURE 4.4. Comparison of position distributions for Sisyphus simulations using Eq. (4.7) and Gaussian diffusion with a diffusion constant given by Eq. (4.24). For $D = 0.01$, a Gaussian simulation (green) and Sisyphus simulation (blue) are shown. For $D = 0.19$, a Gaussian simulation (red) and a Sisyphus simulation (yellow) are shown. Simulation parameters are $\Delta t = 0.1$, $T = 10^4$, $N = 2^{20}$.

for larger p. This gives rise to long, correlated motions that can be modeled as Lévy flights.

The derivation to show that this behavior leads to Lévy flights nontrivial and is outlined in Refs. [39, 64]. The general idea is that the microscopic behavior can be split the motion into a series of correlated "jumps" χ and jump durations τ. The jumps and jump durations (also called waiting times) are defined to be the displacement and time elapsed between momentum zero-crossings, respectively.

It was shown [31] that χ and τ have distributions with asymptotic behavior given by

$$f_\chi(x) \sim |x|^{-(4/3+1/3D)} \tag{4.25}$$

$$f_\tau(t) \sim |t|^{-(3/2+1/2D)}, \tag{4.26}$$

102

as well as correlations that scale as

$$\chi \sim \tau^{3/2}. \tag{4.27}$$

While the jumps and waiting times are correlated, the asymptotic behavior can be found by approximating the joint jump-wait distribution as the uncorrelated product

$$g_{\chi,\tau}(x,t) \sim f_\chi(x) f_\tau(t). \tag{4.28}$$

While this ignores the correlations, ultimately the correlations are not relevant in the asymptotic limit.

The Lévy behavior follows from this assumption after applying a specialized Montroll–Weiss equation, which relates continuous time step distributions to the position probability density [39]. At this point the process can be asymptotically described by an α-stable Lévy process with

$$\alpha = \frac{1+D}{3D}, \tag{4.29}$$

and an anomalous diffusion coefficient given by

$$K_\alpha = \frac{1}{\mathcal{Z}} \frac{\pi(3\alpha-1)^{\alpha-1}}{\sin\left(\frac{\pi\alpha}{2}\right) 3^{2\alpha-1} \left|\Gamma(\alpha)^2\right|}, \tag{4.30}$$

where again $\mathcal{Z}$ is given by

$$\mathcal{Z} = \sqrt{\pi}\,\Gamma\left(\frac{1-D}{2D}\right) \Big/ \Gamma\left(\frac{1}{2D}\right). \tag{4.31}$$

103

This is the connection that justifies the modeling of diffusion in Sisyphus cooling with stable Lévy processes.

4.5. Power Law Extent

Interestingly, as D increases the position distribution goes through several phases. Close to $D = 0$ the position distribution is Gaussian, but then it begins to develop heavy tails for finite simulation time. For $D < 1/5$ these tails eventually decay at long times, approaching asymptotically Gaussian behavior. However, for $D > 1/5$ the tails persist and grow in extent as the simulation time increases. To investigate this behavior, we want to estimate the start and end of the power law behavior.

The onset of the power laws for $D > 1/5$ is given by

$$x_0 := (K_\alpha T)^{-\frac{1}{\alpha}}, \tag{4.32}$$

which is the effective width of the Gaussian-like portion of the distribution. Meanwhile, the cutoff of the power law for $D > 1/5$ is given by

$$x_1 := \sqrt{D} T^{-3/2}, \tag{4.33}$$

which is when correlations between χ and τ become important (for a finite simulation time T, there is an effective maximum jump length).

If we assume a uniform position distribution with a magnitude a, followed by the power law decay

$$f(x) = bx^{-(\alpha+1)}, \tag{4.34}$$

extending from x_0 to x_1, we have the normalization condition

$$1 = ax_0 + b \int_{x_0}^{x_1} x^{-(\alpha+1)} dx. \tag{4.35}$$

This leads to

$$1 = ax_0 - \frac{b}{\alpha} \left(x_1^\alpha - x_0^\alpha \right), \tag{4.36}$$

and for continuity we require

$$bx_0^{-\alpha-1} = a. \tag{4.37}$$

This in turn gives us the relation

$$1 = bx_0^{-\alpha} - \frac{b}{\alpha} \left(x_1^\alpha - x_0^\alpha \right), \tag{4.38}$$

allowing us to solve for b

$$b = \frac{\alpha x_0^\alpha x_1^\alpha}{\alpha x_1^\alpha + x_1^\alpha - x_0^\alpha}. \tag{4.39}$$

This can be used to plot the power laws as shown in Fig. 4.5.

We can see from Fig. 4.4 and Fig. 4.5, that as D is increased, the position distribution begins Gaussian ($D = 0.01$), develops tails ($D = 1.9$), has long tails ($D = 3.0$), and then the tails begin to shorten ($D = 0.5$, $D = 0.7$), in terms of orders of magnitude spanned. To illustrate this, we look at the number of order of magnitudes the power law extends over, specifically a plot of $\log_{10} \left(\frac{x_1}{x_0} \right)$, as shown in Fig. 4.6. Looking at these plots, we find an interesting dependence: For large enough T, a peak forms for a smaller value of D. This occurs because the central portion of the distribution grows faster for larger D than tail of the distribution which grows with $\sqrt{D}$. This implies that if *visible* power laws are desired, D should be tuned to be slightly beyond the $D = 0.2$ threshold. This
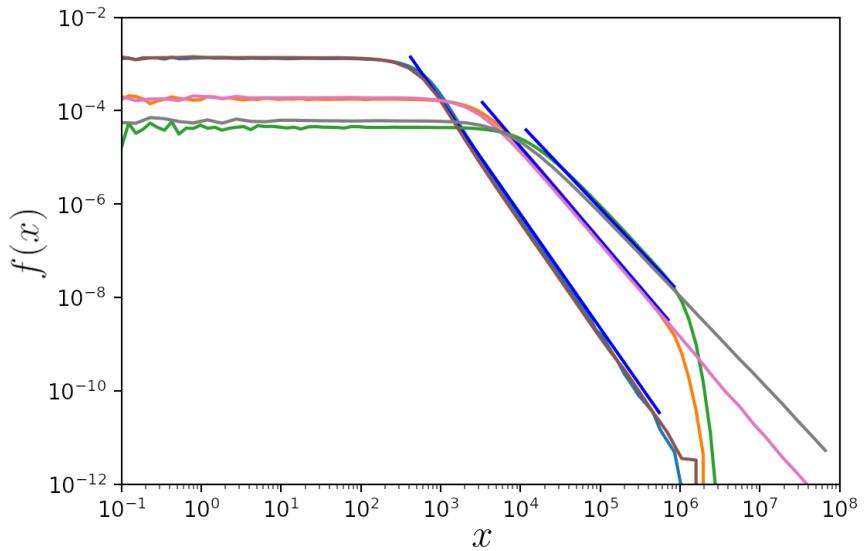
FIGURE 4.5. Comparison of position distributions for Sisyphus simulations using Eq. (4.7) and anomalous diffusion with a diffusion constant given by Eq. (4.30). For $D = 0.3$, an α-stable Lévy simulation (brown) and Sisyphus simulation (light blue) are shown. For $D = 0.5$, an α-stable Lévy simulation (pink) and Sisyphus simulation (orange) are shown. For $D = 0.7$, an α-stable Lévy simulation (gray) and Sisyphus simulation (green) are shown. Also shown are associated power laws given by Eq. (4.34). Simulation parameters are $\Delta t = 0.1$, $T = 10^4$, $N = 2^{20}$.
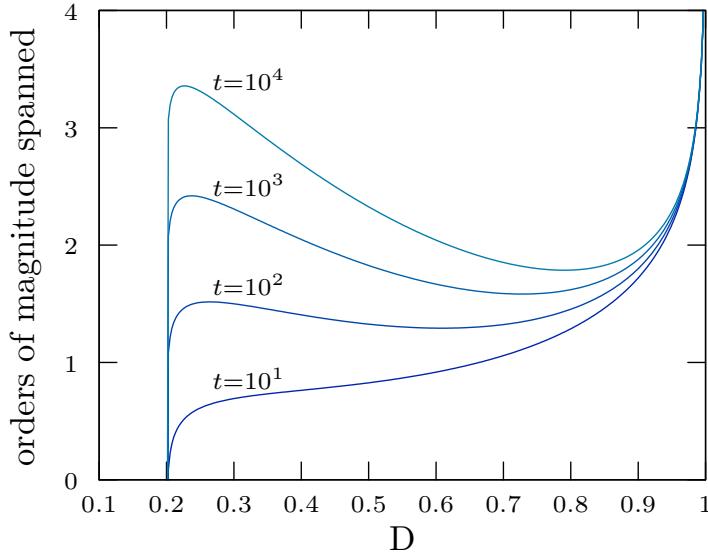
FIGURE 4.6. Plot of $\log_{10}\left(\frac{x_1}{x_0}\right)$ as a function of D for various values of T.

choice has two benefits: First, at any reasonable "long" time we expect further extent of power-law behavior, and as we increase the time further, the power laws grow in extent more quickly.

4.6. Boundary Crossing Statistics

The beginning of Chapter III considered a simple experimental setup with a single atom and a single photodetector. As a signal on the photodetector indicates the presence of the atom in the imaging region, changes in the signal correspond to the atom entering or leaving the imaging region. By measuring the times at which the atom enters or leaves the imaging region, we can therefore measure boundary-crossing statistics.

In this section we investigate the measurement of boundary-crossing statistics for atoms undergoing Sisyphus cooling. In Section 4.6.1 we look at simulations

107
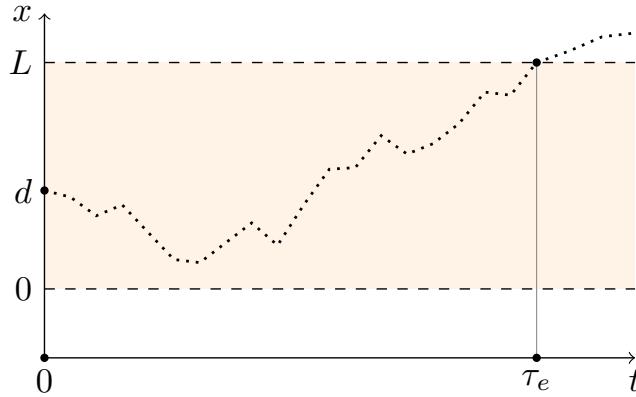
FIGURE 4.7. Schematic for escape time simulations.

of escape times (Section 2.1.8) for Sisyphus cooling in the Brownian regime, which allows for comparison to exact formula. Next in Section 4.6.2 we look at simulations in the Lévy regime for *transit* times, a statistic that can be more easily translated into an experimental system.

4.6.1. Brownian Regime

Simulations of most boundary-crossing statistics in the Brownian regime behave similarly to theoretical predictions for Brownian motion. As an example we show the escape time distributions (see Section 2.1.8 and Fig. 4.7) for $D = 0.1$ for various region sizes in Fig. 4.8.

4.6.2. Lévy Regime

For the Lévy regime we look at simulations of what we call *transit* time, τ_{tr}. Transit times are the time it takes for a particle to cross a particular region, as shown in Fig. 4.9. For these simulations the particle is started at the center of the region, and is allowed to diffuse until it exits the region. Eventually, the particle

108
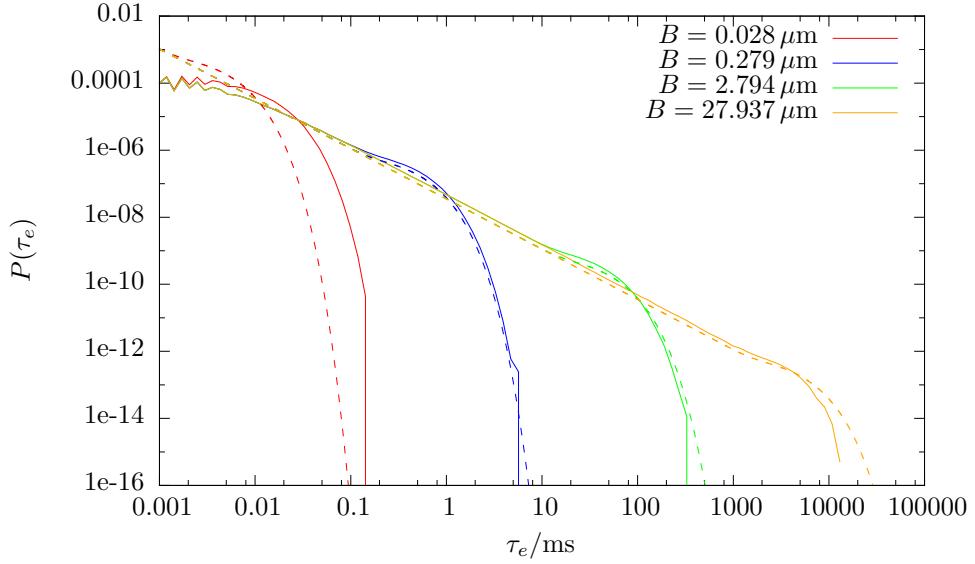
FIGURE 4.8. Escape time distributions for $D = 0.1, T = 10^4$. Sisyphus simulations using Eq. (4.7) (solid) and theoretical curves using Eq. (2.66) (dashed) are shown.

will reenter the region, and the transit time is then measured as the elapsed time between when the particle enters the region and when it leaves.

Fig. 4.10 shows simulated transit time distributions for $D = 0.7$ for various region sizes. In any particular distribution as in Fig. 4.11, there is a broad background until it is cutoff at some large time τ_{cutoff}. These distributions appear similar to the Brownian escape times, except that there is a peak present that extends from some time τ_{min} to some time τ_{max}. This peak is somewhat peculiar since there is no obvious velocity scale for the system due to the power-law tail in the Tsallis distribution.

To understand this peak, it is helpful to look at sample paths with particular transit times, shown on a semi-log scale in Fig. 4.12. Since this shows trajectories for all transit times, it gives a clearer understanding of what typical trajectories look like for a given transit time. For very short times, the typical particle

109
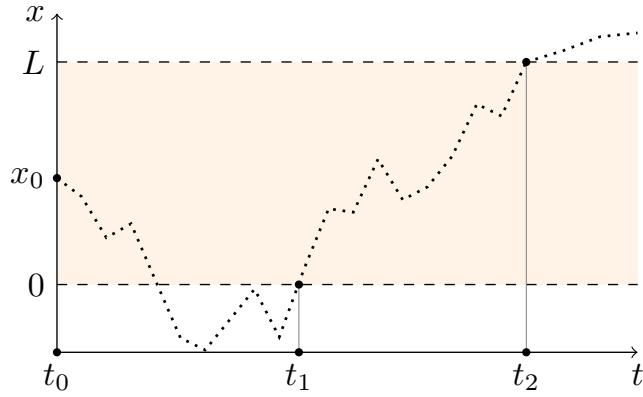
FIGURE 4.9. Schematic for transit time distributions. The transit time is given by $\tau_{\mathrm{tr}} = t_2 - t_1$.
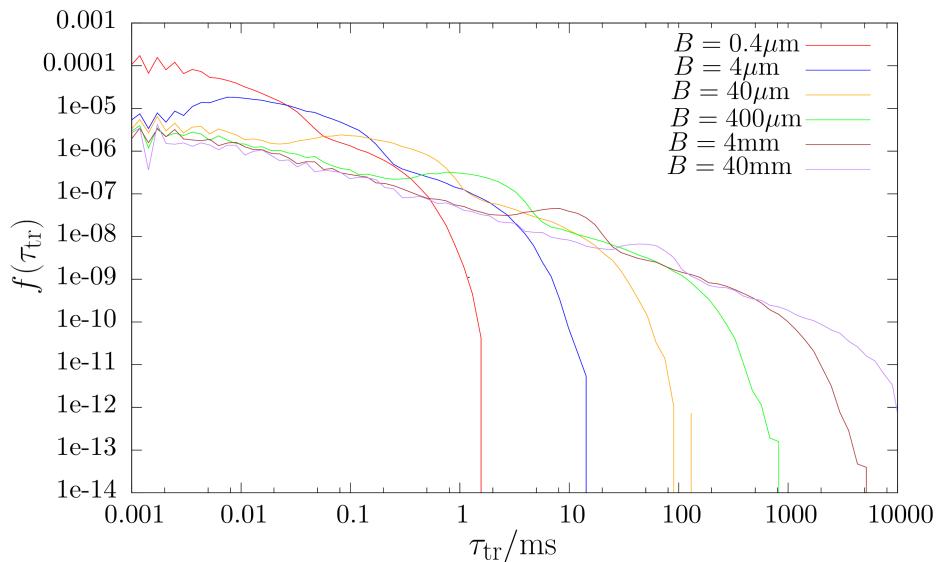
FIGURE 4.10. Transit time distributions for $D = 0.7$, $T = 10^4$.

110
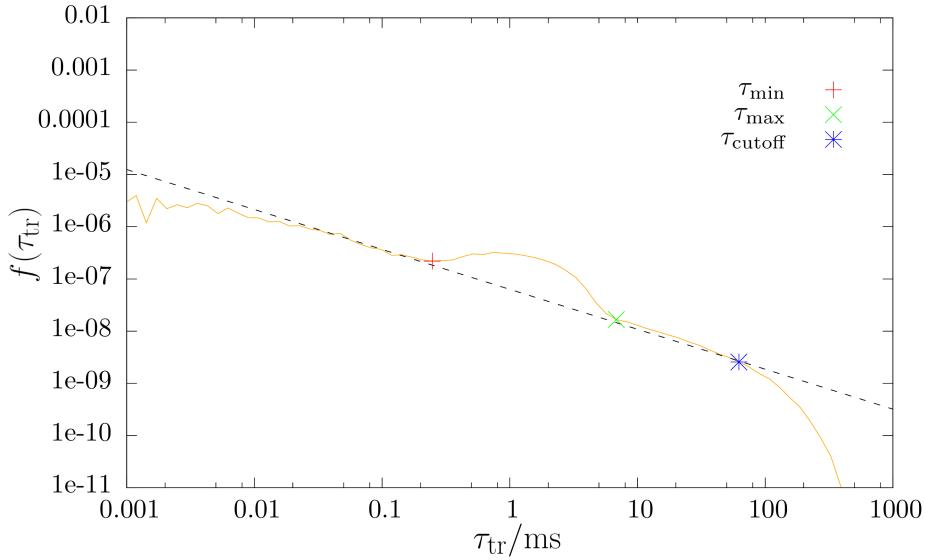
FIGURE 4.11. Escape-time anatomy for $D = 0.7$, $T = 10^4$.

trajectory enters the region, turns around, and leaves the region. However, for longer times the trajectories seem to traverse across the region in a single jump (i.e., without getting "stuck" and changing directions). Finally, for long times, the particles occasionally become trapped, and are more or less equally likely to leave the region at either boundary.

The transit times corresponding to particles that take the longest to cross (τ_{max}) are the slowest particles that cross over the entire region in a single jump. Due to the correlations between jump length χ and jump duration τ given by

$$\chi \sim \tau^{3/2}, \tag{4.40}$$

the slowest particles have shorter jumps, and so the slowest particles that cross the region in a single jump must have $\chi \sim B$. Therefore,

$$\tau_{\mathrm{max}} \sim B^{2/3}. \tag{4.41}$$

111
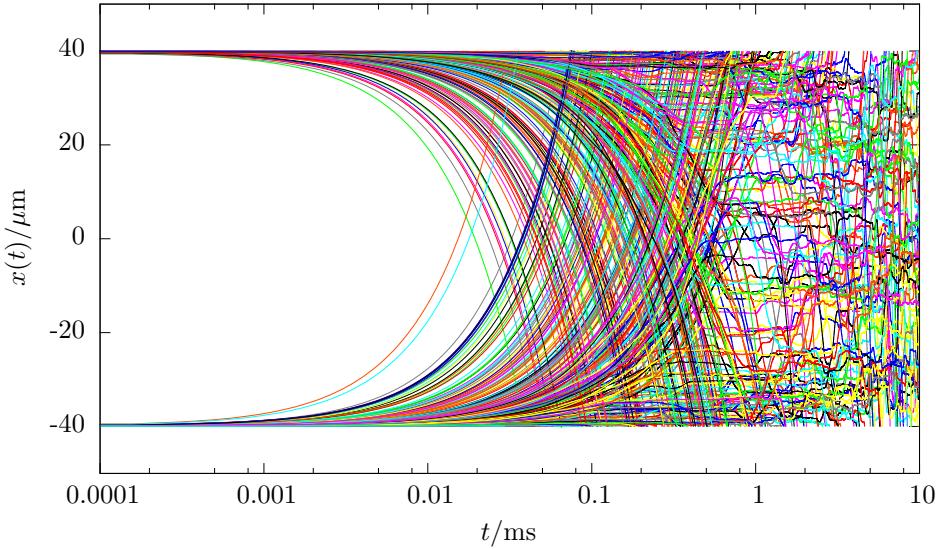
FIGURE 4.12. Selected trajectories for $D = 0.7$, $T = 10^4$, with region size $B = 39.909$. The appearance of curved trajectories is due to the semi-log scaling; unscaled, these curved trajectories appear relatively straight.

On the other hand, the transit times corresponding to particles that take the shortest time to cross ($\tau_{\min}$), must be due to the fastest particles that cross over the entire region in a single jump. Since the velocity of the fastest particles does not depend on the region size we have the relation

$$\tau_{\min} \sim B/\langle v_{\max}\rangle. \tag{4.42}$$

These scaling arguments for $\tau_{\max}$ and $\tau_{\min}$ are confirmed through simulations in Fig. 4.13 and Fig. 4.14 respectively. Further, with the observations of long, single-direction trajectories for transits occurring between these bounds in Fig. 4.12, we can conclude that the presence of these peaks can be used in the direct detection of jumps due to Lévy flight behavior.
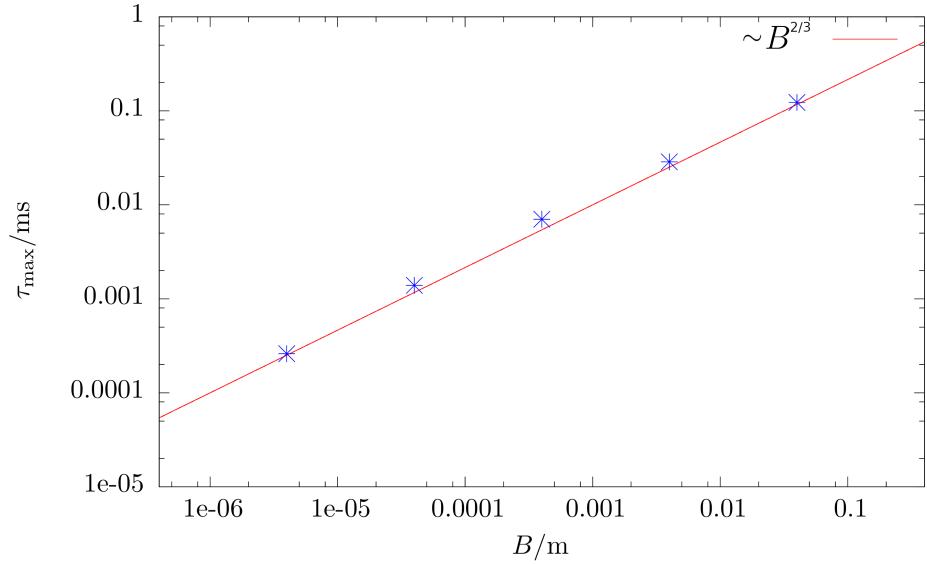
112

FIGURE 4.13. Plot of $\tau_{\max}$ extracted from transit-time distributions as a function of the region size. Parameters used: $D = 0.7$, $T = 10^4$.
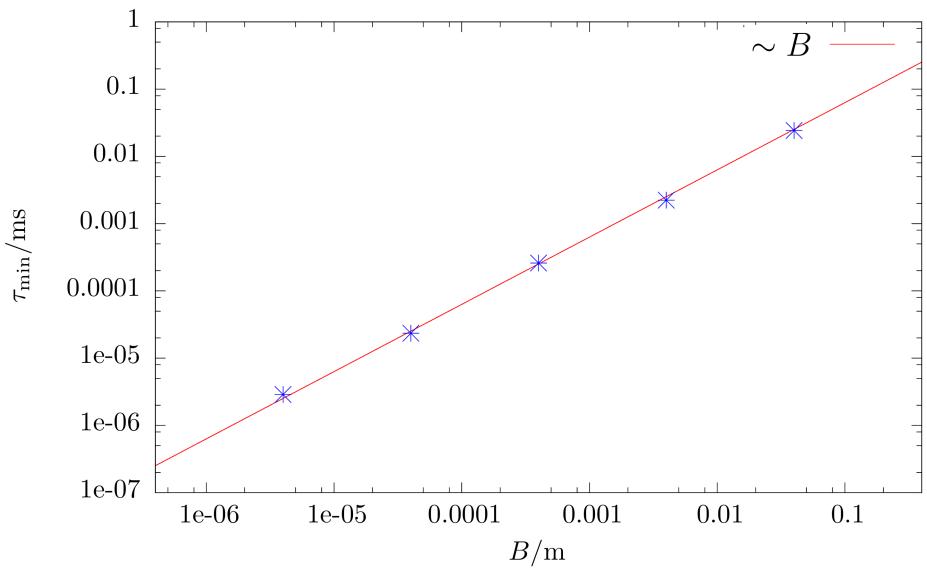
FIGURE 4.14. Plot of $\tau_{\min}$ extracted from transit-time distributions as a function of the region size. Parameters used: $D = 0.7$, $T = 10^4$.

CHAPTER V

CONCLUSION

This dissertation began with a discussion of universality and the idea that distinctly different systems often approach the same behavior. In some sense, the hope was that by understanding the universal limits, we can ignore the complexities of the individual systems. However, much discussion was devoted to the subtleties of applying these tools, and that there is nuanced process of building connections between abstract universal rules and concrete examples.

For example, from the development of fundamental stochastic processes, we saw that the asymptotic behavior guaranteed through the CLT may turn on slowly, such that the asymptotic behavior suggested by the gCLT could be employed even though the preconditions were *not* satisfied. This led to a sort of phase transition between Lévy flights and Brownian motion where both universal behaviors were present, but at different scales. We also saw that common techniques in the Brownian regime, such as the method of images or the generation of stretched bridge, break down for heavy tailed processes; and yet, the seemingly related Sparre Anderson scaling somehow still holds for both Brownian and Lévy regimes.

Developing intuition is particularly important for heavy tailed processes, since many abstract results are prone to error or controversy. We continue to see the application of ill-advised Gaussian statistics in finance, even when more general models are available [13]. As pointed out in Ref. [97], there are several examples of erroneous results in the Lévy flight literature due to poor assumptions involving absorbing boundary conditions for Lévy flights. Even whether animals truly

114

follow Lévy flight patterns, or if such search strategies are actually optimal is not completely clear [19, 59, 60]. Altogether, these examples suggest it is important to develop intuitive ways of approaching Lévy processes.

One of the more significant advances on this front was further development of conditioned Lévy processes and Lévy bridges. In particular, the observation that the Lévy process has bifurcations in the midstep distribution, unlike the Brownian case, and that these bifurcations lead to a new length scale that allows for discrimination between singularly large events and composites of many smaller events. This interpretation gives a better intuition for why sample paths appear by eye to have discrete "jumps", an observation unexplained by the more formal Lévy–Khintchine representation. Since this scale naturally is visible by eye, it seems likely that it would also arise naturally in some physical systems.

Finally, we investigated Sisyphus cooling, which represented many of the themes discussed throughout the prior chapters. One of the more interesting outcomes of studying boundary-crossing statistics was a surprising peak in the transit time distribution, which corresponds to the presence of Lévy flights. Furthermore, the scaling relationships developed to explain this peak have relations both to the underlying physical process as well as to the asymptotic behavior. These statistics and relations are useful for the design of future single-atom experiments.

APPENDIX

A.1. Moments of Momentum Distribution

In the main text (Section 4.2) we calculated the moments of p for Sisyphus cooling (described by Eq. (4.7)) through the typical approach of computing

$$\langle p^m \rangle = \int_{-\infty}^{\infty} dp\, \pi(p) p^m. \tag{A.1}$$

However, another way of analyzing this system is via the Langevin approach. We follow the method described in Jacobs' *Stochastic Processes for Physicists* pg. 41. As we have we have a nonlinear term in our force, we expect that our solution for $\langle p^2 \rangle$ will depend on all higher moments of p; however, we will see that we will still be able to solve for the moments exactly.

First note that the force is anti-symmetric in p. This means that for symmetric initial conditions we expect that $\langle p^n \rangle = 0$ for all odd moments. To find the even moments, we will need to find the increment for $d(p^n)$ which can be done using Ito's formula,

$$dy = \left(\frac{\partial f}{\partial x} \right) dx + \left(\frac{\partial f}{\partial t} \right) dt + \frac{1}{2} \left(\frac{\partial^2 f}{\partial x^2} \right) dx^2. \tag{A.2}$$

This gives us

$$d(p^n) = (np^{n-1})dp + \frac{n(n-1)}{2} p^{n-2} dp^2 \tag{A.3}$$

into which we can insert dp and dp^2 to find

$$d(p^n) = (np^{n-1}) \left(-\frac{p}{1+p^2} dt + \sigma dW \right) + \left(\frac{n(n-1)}{2} p^{n-2} \right) \sigma^2 dt, \tag{A.4}$$

117

where we have made use of $dW^2 = dt$. Taking the expectation gives us

$$d(\langle p^n \rangle) = -n \left\langle \frac{p^n}{1 + p^2} \right\rangle dt + \frac{n(n-1)\sigma^2}{2} \langle p^{n-2} \rangle dt, \qquad (A.5)$$

which can be divided through by dt and reordered into

$$\frac{d(\langle p^n \rangle)}{dt} = \frac{n(n-1)\sigma^2}{2} \langle p^{n-2} \rangle - n \left\langle \frac{p^n}{1 + p^2} \right\rangle. \qquad (A.6)$$

Setting $\frac{d\langle p^n \rangle}{dt} = 0$ to find the steady state, our equation for the moments becomes

$$\langle p^{n-2} \rangle = A_n \left\langle \frac{p^n}{1 + p^2} \right\rangle \qquad (A.7)$$

were we have defined the coefficient $A_n = \frac{2}{(n-1)\sigma^2}$. It is tempting to distribute the expectation brackets, but

$$\left\langle \frac{p^2}{1 + p^2} \right\rangle \neq \frac{\langle p^2 \rangle}{1 + \langle p^2 \rangle} \qquad (A.8)$$

since p is not a Gaussian random variable. However, we can expand this term as a series:

$$\left\langle \frac{p^n}{1 + p^2} \right\rangle = \langle p^2 \rangle - \langle p^4 \rangle + \langle p^6 \rangle + \dots \approx \sum_{k=n/2}^{N} (-1)^{k+1} \langle p^{2k} \rangle. \qquad (A.9)$$

In doing this, we are assuming the existance and finiteness of the higher moments, and that moments greater than $\langle p^{2N} \rangle$ are negligible. This truncation has also made this problem tractable, as we have the relation

$$\langle p^{n-2} \rangle \approx A_n \sum_{k=n/2}^{N} (-1)^{k+1} \langle p^{2k} \rangle \qquad (A.10)$$

which holds for $n = 2, 4, \ldots, 2(N-1), 2N$. Thus there are N equations and N unknowns (all even moments from $\langle p^2 \rangle$ to $\langle p^{2N} \rangle$), so this system can be solved. To find a recursive formula, we write a version of Eq. A.10 shifted by $n \rightarrow n+2$:

$$\langle p^n \rangle = A_{n+2} \sum_{k=n/2+1}^{N} (-1)^{k+1} \langle p^{2k} \rangle. \qquad (A.11)$$

Now, we we write another version of Eq. A.10, pulling out one term from the sum:

$$\langle p^{n-2} \rangle = A_n \left(\langle p^n \rangle - \sum_{k=n/2+1}^{N} (-1)^{k+1} \langle p^{2k} \rangle \right). \qquad (A.12)$$

Solving for the sum, we have

$$\sum_{k=n/2+1}^{N} (-1)^{k+1} \langle p^{2k} \rangle = \langle p^n \rangle - \frac{1}{A_n} \langle p^{n-2} \rangle, \qquad (A.13)$$

and substituting this into for Eq. A.11 we have

$$\langle p^n \rangle = A_{n+2} \left(\langle p^n \rangle - \frac{1}{A_n} \langle p^{n-2} \rangle \right). \qquad (A.14)$$

Now we can solve for $\langle p^n \rangle$:

$$\langle p^n \rangle = \frac{A_{n+2}}{A_n(A_{n+2} - 1)} \langle p^{n-2} \rangle \qquad (A.15)$$

Noting that $\langle p^0 \rangle = 1$, this is a general formula for the moments

$$\langle p^n \rangle = \prod_{k=1}^{n/2} \frac{A_{2k+2}}{A_{2k}(A_{2k+2} - 1)}, \qquad (A.16)$$

119

and substituting for A_n gives

$$\langle p^n \rangle = \prod_{k=1}^{n/2} \frac{(1 - 2k)\sigma^2}{(1 + 2k)\sigma^2 - 2}. \tag{A.17}$$

Finally, this product can be rewritten as

$$\langle p^n \rangle = \frac{i^n}{\sqrt{\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right) \Gamma\left(\frac{3}{2} - \frac{1}{\sigma^2}\right)}{\Gamma\left(\frac{n}{2} + \frac{3}{2} - \frac{1}{\sigma^2}\right)}. \tag{A.18}$$

The first 4 even moments are

$$\langle p^2 \rangle = \frac{(1!!)\sigma^2}{(2 - 3\sigma^2)}$$

$$\langle p^4 \rangle = \frac{(3!!)\sigma^4}{(2 - 3\sigma^2)(2 - 5\sigma^2)}$$

$$\langle p^6 \rangle = \frac{(5!!)\sigma^6}{(2 - 3\sigma^2)(2 - 5\sigma^2)(2 - 7\sigma^2)}$$

$$\langle p^8 \rangle = \frac{(7!!)\sigma^6}{(2 - 3\sigma^2)(2 - 5\sigma^2)(2 - 7\sigma^2)(2 - 9\sigma^2)},$$

where the double factorial (!!) is defined for odd integers N by $N!! = N \cdot (N - 2) \cdot (N - 4)\ldots 5 \cdot 3 \cdot 1$.

For $\sigma = \sqrt{2D}$ we have

$$\langle p^2 \rangle = \frac{(1!!)D}{(1 - 3D)}$$

$$\langle p^4 \rangle = \frac{(3!!)D^2}{(1 - 3D)(1 - 5D)}$$

$$\langle p^6 \rangle = \frac{(5!!)D^4}{(1 - 3D)(1 - 5D)(1 - 7D)}$$

$$\langle p^8 \rangle = \frac{(7!!)D^6}{(1 - 3D)(1 - 5D)(1 - 7D)(1 - 9D)}.$$

A.2. Hypergeometric Inversion Formula

As an aside, we also can use the moments derived in Appendix A.1 to find an alternative form for the momentum distribution. As we already solved for the momentum distribution in Section 4.2, this amounts to deriving a formula for an inverse Fourier transform of a particular hypergeometric function.

We first construct the moment generating function, which is just a power series with the general form

$$M_p(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \langle p^n \rangle x^n, \tag{A.19}$$

where the coefficients are the moments of $f(p)$. Thus, the moments can be generated from $M_p(x)$ by the relation

$$\langle p^n \rangle = (\partial_x^n M_p(x)) \big|_{x=0}. \tag{A.20}$$

Substituting in the moments for $f(p)$ from Eq. (A.18) and noting that the odd moments vanish), the moment generating function is

$$M_p(t) = 1 + \sum_{n=1}^{\infty} \frac{1}{(2n)!} \langle p^{2n} \rangle t^{2n}. \tag{A.21}$$

Substituting in for $\langle p^{2n} \rangle$ we get

$$M_p(x) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{\pi}} \frac{\Gamma\left(n + \frac{1}{2}\right) \Gamma\left(\frac{3}{2} - \frac{1}{\sigma^2}\right)}{\Gamma\left(n + \frac{3}{2} - \frac{1}{\sigma^2} \Gamma(2n + 1)\right)} x^{2n}. \tag{A.22}$$

It turns out that this series is equivalent to the confluent hypergeometric function given by

$$M_p(x) = {}_0F_1\left(\frac{3}{2} - \frac{1}{\sigma^2}, -\frac{x^2}{4}\right).$$
(A.23)

To find the distribution $f(p)$ from the moment generating function, we can first perform a Wick rotation $(x \to ix)$, which provides the characteristic function for $f(p)$. Finally, we take the inverse Fourier transform to recover $f(p)$:

$$f(p) = \mathscr{F}^{-1}[M_p(ix)](p) = \mathscr{F}^{-1}\left[{}_0F_1\left(\frac{3}{2} - \frac{1}{\sigma^2}, \frac{x^2}{4}\right)\right].$$
(A.24)

Combining this result with Eq. (4.20), we have the following inversion formula

$$\mathscr{F}^{-1}\left[{}_0F_1\left(\frac{3}{2} - \frac{1}{\sigma^2}, -\frac{x^2}{4}\right)\right](p) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{1}{\sigma^2})}{\Gamma(\frac{1}{\sigma^2} - \frac{1}{2})} \left(1 + p^2\right)^{-1/\sigma^2},$$
(A.25)

which can be verified numerically.

A.3. Equivalence of Discrete and Langevin Methods

The discrete random walk takes a step Δp in momentum every time step Δt. The direction of the step is determined by

$$P_{\pm} = \frac{1}{2}\left(1 \pm \frac{1}{2D}F(p)\Delta p\right), \tag{A.26}$$

where P_+ is the probability of a positive step and P_- is the probability of a negative step, and $F(p) = -p/(1 + p^2)$ is the Sisyphus cooling force. Notice that the size of the time step Δt must be related to the simulation parameters as it determins the scattering rate Γ. The correct relation is $\Gamma = 1/\Delta t = \Delta p^2/2D$. Meanwhile the size of the momentum step is what determines the accuracy of the simulation, with a smaller Δp giving better consistency with the Langevin equation[1].

To show this is equivalent to the Langevin equation, we can write this discrete step process as a sum of two Poisson processes

$$dp = \Delta p \, dN_+ - \Delta p \, dN_-, \tag{A.27}$$

with the ensemble average of $dN_{\pm}$ given by $\langle\langle N_{\pm}\rangle\rangle = \Gamma_{\pm} = P_{\pm}/\Delta t$. This approximation is valid as long as $t \gg \Delta t$, the regime where the discrete momentum kicks are not resolvable. We can now approximate our possion process as a drift term and a Wiener process by making the substituion $dN_{\pm} \to$

---

[1]Since photon emission and absorption events are discrete, Δp is actually physically determined, and the discrete method has some grounds to be a more accurate model.

$\Gamma_\pm\,dt + \sqrt{\Gamma_\pm}\,dW$ and then simplifying:

$$dp = \Delta p\,(\Gamma_+ - \Gamma_-)\,dt + \Delta p\sqrt{\Gamma_+}\,dW_1 + \Delta p\sqrt{\Gamma_-}\,dW_2 \tag{A.28}$$

$$dp = \frac{\Delta p}{\Delta t}\,(P_+ - P_-)\,dt + \Delta p\sqrt{|\Gamma_+| + |\Gamma_-|}\,dW \tag{A.29}$$

$$dp = \frac{\Delta p}{\Delta t}\left(\frac{1}{2D}F(p)\Delta p\right)dt + \frac{\Delta p}{\sqrt{\Delta t}}\,dW \tag{A.30}$$

$$dp = \frac{\Delta p^2}{2D\Delta t}F(p)\,dt + \sqrt{2D}\,dW \tag{A.31}$$

$$dp = F(p)\,dt + \sqrt{2D}\,dW \tag{A.32}$$

This is the same as Eq. (4.7).

REFERENCES CITED

[1] Borodin, *Handbook of Brownian Motion 2nd Ed* (Birkhäuser, 2002).

[2] A. Einstein, Ann. Phys **19**, 11 (1905).

[3] A. Shukurov, *Dynamics and Evolution of Galactic Nuclei* (Taylor & Francis, 2014).

[4] L. Bachelier, Annales scientifiques de l'École normale supérieure **17**, 21 (1900).

[5] M. Sewell, UCL Research **11**, 4 (2011).

[6] B. Hughes, *Random Walks and Random Environments I* (Oxford University Press, 1995).

[7] R. Brown, *The Miscellaneous Botanical Works of Robert Brown* (The Ray Society, 1866).

[8] J. Perrin, *Brownian Movement and Molecular Reality* (Taylor and Francis, 1910).

[9] F. Black and M. Scholes, The Journal of Political Economy **81**, 637 (1973).

[10] K. Kiyono, Z. R. Struzik, and Y. Yamamoto, Physical Review Letters **96**, 068701 (2006).

[11] R. Rebonato, *Volatility and Correlation: In the Pricing of Equity, FX and Interest-Rate Options* (John Wiley & Sons, 1999).

[12] N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable* (Random House, 2007).

[13] R. Cont and P. Tankov, *Financial Modelling with Jump Processes* (Chapman & Hall, 2004).

[14] C. Gardiner, *Stochastic Methods. A Handbook for the Natural and Social Sciences*, 4th ed. (Springer, 2009).

[15] K. Jacobs, *Stochastic Processes for Physicists: Understanding Noisy Systems* (Cambridge University Press, 2010).

[16] M. F. Shlesinger, G. M. Zaslavsky, and U. Frisch, eds., *Lévy Flights and Related Topics in Physics: Proceedings of the International Workshop Held at Nice, France, 27-30 June 1994*, Lecture Notes in Physics No. 450 (Springer-Verlag, 1995).

[17] V. V. Uchaikin and V. M. Zolotarev, *Chance and Stability: Stable Distributions and Their Applications* (De Gruyter, 1999).

[18] R. Metzler, T. Koren, B. van den Broek, G. J. L. Wuite, and M. A. Lomholt, Journal of Physics A: Mathematical and Theoretical **42**, 434005 (2009).

[19] G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, and H. E. Stanley, Nature **381**, 413 (1996).

[20] T. W. Hänsch and A. L. Schawlow, Optics Communications **13**, 68 (1975).

[21] S. Chu, L. Hollberg, J. E. Bjorkholm, A. Cable, and A. Ashkin, Physical Review Letters **55**, 48 (1985).

[22] A. Ashkin, Physical Review Letters **24**, 156 (1970).

[23] V. S. Letokhov and B. D. Pavlik, Soviet Journal of Experimental and Theoretical Physics **45**, 698 (1977).

[24] A. Aspect, E. Arimondo, R. e a1 Kaiser, N. Vansteenkiste, and C. Cohen-Tannoudji, Physical Review Letters **61**, 826 (1988).

[25] F. Bardou, J. P. Bouchaud, O. Emile, A. Aspect, and C. Cohen-Tannoudji, Physical review letters **72**, 203 (1994).

[26] F. Bardou, A. Aspect, J.-P. Bouchaud, and C. Cohen-Tannoudji, *Lévy Statistics and Laser Cooling: How Rare Events Bring Atoms to Rest* (Cambridge University Press, 2002).

[27] C. Cohen-Tannoudji and J. Dalibard, *Laser Cooling beyond the Doppler Limit by Polarization Gradients: Simple Theoretical Models* (World Scientific, 1989).

[28] C. Cohen-Tannoudji, in *Light Induced Kinetic Effects on Atoms, Ions and Molecules* (ETS Editrice, 1991).

[29] J. Dalibard and C. Cohen-Tannoudji, Journal of the Optical Society of America. B, Optical physics **2**, 1707 (1985).

[30] T. W. Hodapp, C. Gerz, C. Furtlehner, C. I. Westbrook, W. D. Phillips, and J. Dalibard, Applied Physics B **60**, 135 (1995).

[31] S. Marksteiner, K. Ellinger, and P. Zoller, Physical Review A **53** (1996).

[32] Dalibard and Cohen-Tannoudji, Journal of Physics B: Atomic and Molecular Physics **18**, 1661 (1985).

[33] E. Lutz, Physical Review A **67**, 10.1103/PhysRevA.67.051402 (2003).

[34] E. L. Raab, M. Prentiss, A. Cable, S. Chu, and D. E. Pritchard, Physical Review Letters **59**, 2631 (1987).

[35] W. Ketterle and N. J. Van Druten, Advances in atomic, molecular, and optical physics **37**, 181 (1996).

[36] V. Zaburdaev, S. Denisov, and J. Klafter, Reviews of Modern Physics **87**, 483 (2015).

[37] H. Katori, S. Schlipf, and H. Walther, Physical Review Letters **79**, 2221 (1997).

[38] Y. Sagi, M. Brook, I. Almog, and N. Davidson, Physical Review Letters **108**, 10.1103/PhysRevLett.108.093002 (2012).

[39] E. Barkai, E. Aghion, and D. A. Kessler, Physical Review X **4**, 10.1103/PhysRevX.4.021036 (2014).

[40] E. G. Flekkø y, Physical Review E **95**, 10.1103/PhysRevE.95.012139 (2017).

[41] D. A. Steck, Stochastic Processes, http://steck.us/teaching (2019).

[42] G. A. Pavliotis, *Stochastic Processes and Applications*, Texts in Applied Mathematics, Vol. 60 (Springer New York, New York, NY, 2014).

[43] J. Klafter and I. M. Sokolov, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, 2011).

[44] O. E. Barndorff-Nielsen, S. I. Resnick, and T. Mikosch, eds., *Lévy Processes* (Birkhäuser Boston, Boston, MA, 2001).

[45] J. Bertoin, *Lévy Processes*, Vol. 121 (Cambridge university press Cambridge, 1996).

[46] J. Bertoin, Handbook of statistics **19**, 117 (2001).

[47] D. A. Steck, Quantum and Atom Optics, http://steck.us/teaching (2019).

[48] S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, 2001).

[49] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, 1995).

[50] K. Burrage, P. Burrage, D. J. Higham, P. E. Kloeden, and E. Platen, Physical Review E **74**, 068701 (2006).

[51] K. Debrabant and A. Kvæ rnø, BIT Numerical Mathematics **57**, 153 (2017).

[52] E. S. Andersen, Mathematica Scandinavica **1**, 263 (1953).

[53] E. S. Andersen, Skand. Aktuarietikskr **36**, 123 (1953).

[54] S. Song, J. Jeong, and J. Song, Journal of the Korean Statistical Society **40**, 227 (2011).

[55] C. Li, *Option Pricing with Generalized Continuous Time Random Walk Models*, Ph.D. thesis, Queen Mary, University of London (2016).

[56] L. Wu, SSRN Electronic Journal 10.2139/ssrn.871760 (2006).

[57] T. H. Solomon, E. R. Weeks, and H. L. Swinney, Physical Review Letters **71**, 3975 (1993).

[58] M. F. Shlesinger, G. M. Zaslavsky, and J. Klafter, Nature **363**, 31 (1993).

[59] V. V. Palyulin, A. V. Chechkin, and R. Metzler, Proceedings of the National Academy of Sciences **111**, 2931 (2014).

[60] S. Benhamou, Ecology **88**, 1962 (2007).

[61] M. A. Lewis, P. K. Maini, and S. V. Petrovskii, eds., *Dispersal, Individual Movement and Spatial Ecology*, Lecture Notes in Mathematics, Vol. 2071 (Springer Berlin Heidelberg, 2013).

[62] H. V. Ribeiro, S. Mukherjee, and X. H. T. Zeng, Physical Review E **86**, 10.1103/PhysRevE.86.022102 (2012).

[63] J. L. Cabrera and J. G. Milton, Chaos: An Interdisciplinary Journal of Nonlinear Science **14**, 691 (2004).

[64] D. A. Kessler and E. Barkai, Physical Review Letters **108**, 10.1103/PhysRevLett.108.230602 (2012).

[65] G. Afek, J. Coslovsky, A. Courvoisier, O. Livneh, and N. Davidson, Physical Review Letters **119**, 10.1103/PhysRevLett.119.060602 (2017).

[66] E. Aghion, D. A. Kessler, and E. Barkai, Physical Review Letters **118**, 10.1103/PhysRevLett.118.260601 (2017).

[67] J.-i. Inoue and N. Sazuka, Physical Review E **76**, 021111 (2007).

[68] R. Mantegna, Phys. Rev. Lett. **73** (1994).

[69] J. M. Chambers, C. L. Mallows, and B. W. Stuck, J. Am. Stat. Assoc. **71**, 6 (1976).

[70] A. V. Chechkin, R. Metzler, V. Y. Gonchar, J. Klafter, and L. V. Tanatarov, Journal of Physics A: Mathematical and General **36**, L537 (2003).

[71] T. Koren, A. Chechkin, and J. Klafter, Physica A: Statistical Mechanics and its Applications **379**, 10 (2007).

[72] B. o. Dybiec, E. Gudowska-Nowak, and A. Chechkin, Journal of Physics A: Mathematical and Theoretical **49**, 504001 (2016).

[73] R. M. Blumenthal, R. K. Getoor, and D. B. Ray, Trans. Am. Math. Soc. **99**, 16 (1961).

[74] D. C. Brody, L. P. Hughston, and A. Macrina, in *Advances in Mathematical Finance*, edited by J. J. Benedetto, A. Aldroubi, I. Daubechies, C. Heil, J. McClellan, M. Unser, M. V. Wickerhauser, D. Cochran, H. G. Feichtinger, M. Kunt, W. Sweldens, M. Vetterli, M. C. Fu, R. A. Jarrow, J.-Y. J. Yen, and R. J. Elliott (Birkhäuser Boston, Boston, MA, 2007) pp. 231–257.

[75] B. Moskowitz and R. Caflisch, Mathematical and Computer Modelling **23**, 37 (1996).

[76] J. S. Horne, E. O. Garton, S. M. Krone, and J. S. Lewis, Ecology **88**, 2354 (2007).

[77] H. Gies, K. Langfeld, and L. Moyaerts, Journal of High Energy Physics **2003**, 018 (2003).

[78] J. B. Mackrory, T. Bhattacharya, and D. A. Steck, Physical Review A **94**, 042508 (2016), arXiv:1606.00150 .

[79] D. S. Dean, A. Iorio, E. Marinari, and G. Oshanin, Physical Review E **94**, 032131 (2016).

[80] P. Levitz, D. S. Grebenkov, M. Zinsmeister, K. M. Kolwankar, and B. Sapoval, Physical Review Letters **96**, 180601 (2006).

[81] F. Mori, S. N. Majumdar, and G. Schehr, Physical Review Letters **123**, 200201 (2019).

[82] A. Perret, A. Comtet, S. N. Majumdar, and G. Schehr, Physical Review Letters **111**, 240601 (2013).

[83] F. B. Knight, Astérisque **263**, 19 (1996).

[84] P. J. Fitzsimmons and R. K. Getoor, Stochastic Processes and their Applications **58**, 17 (1995).

[85] E. Hoyle, L. P. Hughston, and A. Macrina, Stochastic Processes and their Applications **121**, 856 (2011).

[86] E. Hoyle, L. P. Hughston, and A. Macrina, Stable-1/2 Bridges and Insurance: A Bayesian approach to non-life reserving (2010).

[87] L. Chaumont and G. Uribe Bravo, The Annals of Probability **39**, 609 (2011).

[88] P. Fitzsimmons, J. Pitman, and M. Yor, in *Seminar on Stochastic Processes, 1992*, edited by E. Çinlar, K. L. Chung, M. J. Sharpe, R. F. Bass, and K. Burdzy (Birkhäuser Boston, Boston, MA, 1993) pp. 101–134.

[89] N. Privault, Annales de Institut Henri Poincare (B) Probability and Statistics **40**, 599 (2004).

[90] A. Janicki and A. Weron, *Simulation and Chaotic Behavior of Alpha-Stable Stochastic Processes* (CRC Press, 1993).

[91] C. Chiarella, European Journal of Political Economy **7**, 14 (1991).

[92] V. Nagaev and M. Shkol'Nik, Theory of Probability and its Applications **33**, 7 (1989).

[93] W. W. Erickson and D. A. Steck, arXiv:2002.03849 [cond-mat, physics:nlin] (2020), arXiv:2002.03849 [cond-mat, physics:nlin] .

[94] P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg, N. Wermuth, and S. Zeger, in *Springer* (1999) p. 262.

[95] A. Aspect, J. Dalibard, A. Heidmann, C. Salomon, and C. Cohen-Tannoudji, Physical review letters **57**, 1688 (1986).

[96] D. A. Kessler and E. Barkai, Physical Review Letters **105**, 10.1103/PhysRevLett.105.120602 (2010).

[97] B. Dybiec, E. Gudowska-Nowak, and P. Hänggi, Physical Review E **73**, 10.1103/PhysRevE.73.046104 (2006)