2008

On The Stochastic Behavior of Brownian Particles in Potential Wells, as Observed with Optical Traps

Ross Paul Brody

ON THE STOCHASTIC BEHAVIOR OF BROWNIAN PARTICLES IN POTENTIAL WELLS, AS OBSERVED WITH OPTICAL TRAPS

By

Ross Paul Brody

Bachelor of Science, Oregon State University 2001

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Physics)

The Graduate School

The University of Maine

May, 2008

Advisory Committee:

R. Dean Astumian, Professor of Physics, Advisor

James P. McClymer, Associate Professor of Physics

Jayendran C. Rasaiah, Professor of Chemistry

Charles W. Smith, Professor of Physics

William N. Unertl, Professor of Physics

LIBRARY RIGHTS STATEMENT

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at The University of Maine, I agree that the Library shall make it freely available for inspection. I further agree that permission for "fair use" copying of this thesis for scholarly purposes may be granted by the Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature:

Date:

ON THE STOCHASTIC BEHAVIOR OF BROWNIAN PARTICLES IN POTENTIAL WELLS, AS OBSERVED WITH OPTICAL TRAPS

By Ross Paul Brody

Thesis Advisor: R. Dean Astumian

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Physics)
May, 2008

When the random fluctuations of a system are viewed as energetic fluctuations, many of the unique qualities of the system become irrelevant to the fundamental behavior. Consequently, many stochastic processes are fundamentally identical and are treated mathematically as such. For this reason the study of colloidal particles in aqueous solution has been invaluable to investigations of biologically relevant stochastic processes. This work addresses the motion of a Brownian particle, in a potential well, whose random fluctuations are described by a Gaussian-Markov random variable.

By performing optical trapping experiments on micron sized, non-interacting, latex spheres in aqueous solution we have shown that at equilibrium, fluctuations in a forward sense are as likely to occur as fluctuations in the reverse sense. One relationship we have investigated relates the conditional probability of a transition occurring during a specific elapsed time and its spatial inverse to the equilibrium probabilities of the initial and final states a and b, $P(b, \Delta t|a, 0)/P(a, \Delta t|b, 0) =$

$e^{-\Delta U/k_B T}$. We have shown this relationship holds for times both short and long compared to the average time to transition from a to b. A second relationship we have investigated, which may be explained by microscopic reversibility alone, equates the average time for a "direct" transition from a to b to the average time for a "direct" transition from b to a. We refer to these as last-touch-first-touch-times (LTFTT), or the average time that elapses between the last touching of the initial position and the first touching of the final position. Experimental limitations prevented direct measurement of LTFTTs, however, we did observe an equality between last-observed-first-observed times (LOFOT). Using a discretized Langevin equation to simulate our system we recovered the equality between LTFTTs and obtained numerical results for LTFTTs.

DEDICATION

For those who have prodded me along, all the while patiently watching over me as I have learned to walk.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1
INTRODUCTION

1.1 Historical background

While peering through his microscope in 1827, the botanist Robert Brown observed a stunning, unchoreographed dance! He was witness to pollen grains moving unassisted in a non-uniform, unpredictable fashion. Brown ultimately concluded this motion of pollen, the "seed" of the flower, was not a sign of life, not the "origin" of life, not even *organic* in nature. He arrived at this conclusion by observing similar motion in suspensions of decidedly inorganic powders of glass, minerals, even a fragment of the Sphinx. The nature of this Brownian motion, as well as a thorough mathematical explanation of it eluded the scientific community throughout the remainder of the nineteenth century and did not arrive until 1905.

Until the latter half of the nineteenth century scientists were generally of the opinion that deterministic explanations could completely describe any physical phenomenon. In a deterministic system, all future properties may be deduced from sufficient initial conditions. These deterministic problems involved quantities such as the trajectory of a projectile, the force exerted on a charged particle in a magnetic field, the orbit of a satellite. At the turn of the twentieth century, systems exhibiting non-deterministic properties garnered greater attention, owed in part to the burgeoning new physics of quantum mechanics. These investigations were fueled in part by the independent works of Albert Einstein and Marian Smoluchowski that dealt with Brownian motion. From their works, stochastic processes were given a mathematical foundation. Einstein's 1905 paper, whose title roughly translates to "Concerning the motion, as required by the molecular-kinetic theory of heat, of

particles suspended in liquids at rest" [31], was instrumental to the field. From Einstein's paper the fundamental properties of Brownian motion can be summed up in two assertions. First, that the motion of the pollen grain arises from the incessant collisions with the surrounding water molecules, and second, that Brownian motion can only be described probabilistically, because of the statistically independent, or uncorrelated, collisions [36]. Nearly all microscopic, biological processes are dependent upon Brownian motion. Whether directly or indirectly, cells, enzymes, DNA, motor-proteins are all undergoing Brownian motion. To better understand this behavior, scientists have developed a variety of techniques enabling these microscopic objects to be localized within a small region, the most relevant technique to the work presented here is the optical tweezer.

In 1969 Arthur Ashkin began his pioneering work using lasers to exert forces on very small objects approximately equal in size to the wavelength of the laser, $d \sim \lambda_{light}$ [1]. He hypothesized that the high intensities and gradients that could be created with coherent, continuous-wave lasers would induce radiation pressures strong enough to exert observable forces on a variety of objects, ranging from the atomic to cellular scale. His first experiments used a mildly focused laser to exert an axial force on a transparent latex bead suspended in water. Not only did he observe an axial force pushing the bead along the direction of propagation of the beam, he noted another force was present, a radial force that tended to restore the bead to the beam-axis if it strayed! Further experiments and refinements followed and in 1986 a stable, three-dimensional optical trap was demonstrated [22]. This new trap stably confined colloidal silica beads in water (in both the 1969 and 1986 experiments the beads were Brownian particles). Although this was done while Ashkin and his colleagues were struggling with their intended experiment which was to trap atoms, this method of trapping micron-sized particles has proved invaluable. Where unrestricted Brownian motion had previously hindered observation and control, nu-

merous experiments in biology, physics, and engineering were made possible with this new trapping method.

1.2 Motivation, or, the energetics

Underlying every aspect of this work is the recurring theme of energetics. Chemical reactions, biological processes, mechanical behavior/motion, these can all be studied from the perspective of energy and how it fluctuates during processes and reactions. In every chemical reaction there is an associated energy exchange, that is, the reaction will be either *exothermic*, releasing energy, or *endothermic*, absorbing energy. Similarly, during the exchange of materials across membrane channels, solute particles are able to overcome potential energy barriers by absorbing thermal energy from the surrounding medium. With this energy, the solute particles may diffuse over the energy barrier, and thus traverse the channel. For a more mechanical example we have Feynman's ratchet and pawl mechanism that relies upon Brownian motion to turn a pinwheel and induce, on average, directed motion for well-defined conditions [33]. In each of these examples the inherent fluctuations within the system are responsible for driving the reactions/processes and serving as a conduit that provides the necessary energy. By studying the random behavior of analogous systems we can develop models that utilize their statistical (probabilistic) nature to understand the underlying energetics as well as fundamental physics.

It is important to note that the processes discussed in this work are Markov processes, they are history independent which means their future depends only upon their present state. Furthermore, we are dealing with an over-damped system in which the frictional forces are strong enough that we may impose mechanical equilibrium at all times. We are working in the regime of low Reynolds number. In this regime, the ratio of the inertial forces to the viscous forces is less than one.

As we enter the world of low Reynolds number, we leave Newtonian mechanics for Aristotelian mechanics (at least insomuch as we may ignore inertial effects) [56]; the Aristotelian description is a limiting case of the Newtonian. These are not unfounded statements and assumptions, the motion of a Brownian particle *really* is Markov, and when the Reynolds number is < 1, inertia *really* is negligible! Because these two conditions are met in every microscopic, biological process, a large body of theoretical work has been dedicated to their study. Furthermore, we assume the fluctuations may be described by Gaussian random variables. If $\xi(t)$ is a Gaussian random variable it will have zero-mean, $\langle \xi(t) \rangle = 0$, and be temporally uncorrelated, $\langle \xi(t)\xi(t') \rangle = \delta_{t,t'}$. Again, we have strictly considered stochastic Gaussian-Markov processes in the over-damped limit.

1.3 The focus of this work

Our work is focused around two quantities. The first of these is the the conditional probability that the particle will be found at a final position x_f at a time Δt after being observed at an initial position x_i, $P(x_f, \Delta t | x_i, 0)$. This is compared to the spatially inverted conditional probability where x_i and x_f are exchanged, $P(x_i, \Delta t | x_f, 0)$. There is good reason to study the ratio of conditional probabilities, this quantity tells us of the relative likelihood that a transition will occur in the forward or reverse direction. Furthermore, as will be seen in Chapter 2, under certain conditions, the ratio is independent of Δt and only dependent upon the potential difference between the endpoints $\Delta U = U(x_f) - U(x_i)$. The relationship is written as

$$\frac{P(x_f, \Delta t | x_i, 0)}{P(x_f, \Delta t | x_i, 0)} = e^{-\Delta U / k_B T}. \tag{1.1}$$

This relationshiop implies that observations made at arbitrary time intervals can lead to information about the inherent energetics of the system.

The second quantity we have studied was motivated by the work of Bier et al.[16]. They derived a result that asserts that the time that elapses between the last touching of the initial position x_i and the first touching of the final position x_f, last-touch-first-touch time (LTFTT), is identical for the forward and reverse directions. This independence of the direction of travel is expressed as

$$\mathrm{LTFTT}(\mathrm{x_i} \rightarrow \mathrm{x_f}) = \mathrm{LTFTT}(\mathrm{x_f} \rightarrow \mathrm{x_i}). \tag{1.2}$$

This time represents the average time the particle spends directly traversing the space between x_i and x_f. The LTFTT does not include the time that elapses while the particle fluctuates around x_i before touching x_i for the last time, that time spent around x_i is the referred to as the mean "wiggling" time (MWT). The sum of the MWT and LTFTT is the mean first passage time MFPT. Quantifying the average time to pass between two positions (or states) is important whenever we wish to know the average duration of a transition and not necessarily its likelihood.

We intended to experimentally verify the equality between LTFTTs, however, instrumental limitations prevented this. The LTFTTs are orders of magnitude shorter than the shortest time that elapses between our measurements. Although we were able to computationally show the direction independence of LTFTTs. Perhaps a more interesting result than the equality of the LTFTTs was our finding that, on average, the last-observed-first-observed time (LOFOT) is independent of the direction of travel as well. The LOFOT represents the time that elapses, or the number of observations that must be made, between last observing the particle at x_i and first observing it at x_f, i.e.

$$\mathrm{LOFOT}(\mathrm{x_i} \rightarrow \mathrm{x_f}) = \mathrm{LOFOT}(\mathrm{x_f} \rightarrow \mathrm{x_i}). \tag{1.3}$$

Over time, this result may prove to be more universally applicable as it places no limitations upon the time interval between observations making it more relevant to counting studies.

1.4 The organization of this work

We have organized this work into five chapters. Chapter 2 introduces our system in greater detail and presents the theoretical analysis that leads to expressions for the conditional probabilities and LTFTTs we are interested in. Chapter 3 describes how data for our system was obtained and analyzed. We discuss optical trapping experiments, our specific apparatus, and the computational methods that were used in simulating our system. Chapter 4 contains in-depth discussions of our primary results dealing with the time-independence of the ratio of spatially inverted conditional probabilities and the equality of both last-touch-first-touch and last-observed-first-observed transition times. Finally, Chapter 5 summarizes the entire body of work, addresses certain shortcomings, it closes with a discussion of various works dealing with generalized fluctuation-dissipation theorems and how this work may be extended to systems not in thermodynamic equilibrium.

Chapter 2
BACKGROUND & THEORY

2.1 The system

If we place a marble in a wooden bowl, resting on a table, we know quite well what will happen, the marble will find the lowest point in the bowl to which it can roll. It will remain there indefinitely, unless disturbed. Now, if we shrink the marble by four orders of magnitude, so that it's one-thousandth of a millimeter across (it was originally ten millimeters, or 1cm), 0.001 mm, and we also shrink the bowl by the same amount, what can we say about our mini-marble in our mini-bowl? The mini-marble will still seek out that lowest spot in the bowl, the point where the potential energy is smallest, the most stable place. Care must now be taken. Since the mini-marble is so small, the surrounding molecules colliding with it can, at every instant in time, be sufficient to exert a force large enough to move the mini-marble away from the mini-bowl's potential energy minimum, away from the mini-marble's equilibrium point, i.e. uphill! Such behavior is expected in quantum mechanical systems where the system may be in a superposition of states having different energies[47, 37]. That is, for a quantum mechanical particle in a box, there is a finite probability that the particle will have multiple energies, thus there is a chance the particle will be "found" away from the lowest energy state, i.e. the potential energy minimum. One difficulty in trying to understand quantum mechanics is that our intuition is mocked, we *want* to think of the quantum mechanical particle in a box as moving in space and we *really want* to picture the particle to be further up the slope of the box/bowl when it has a higher energy, but this isn't the case at all. But for the mini-marble in its mini-bowl it is. The mini-marble really can be pushed up the

sides of the bowl. So our classical intuition hasn't completely failed us, we use the word completely because our mini-system is not deterministic, it is random. This is precisely why it's interesting, the majority of all life-processes are not deterministic. This chapter is devoted to looking at a system resembling our scaled-down marble in a bowl that is filled with water. The water supplies the mini-marble with a random energy-source as well as an energy sink, following each energy *kick* from the water-molecules the mini-marble will experience a force analogous to colliding with a brick-wall because of the viscosity of the water.

There are a large number of questions we could ask about this system. Regarding the motion of the mini-marble (hereafter, the particle) we wonder at it's curious, random motion: What is the probability that we'll find it at any given point in the bowl (hereafter, the potential)? How much force does the particle feel from this potential? How can we determine the force exerted by the potential when the particle is always being pushed around by the molecules surrounding it? But there is more we can ask. For example, if the particle starts at one point, x_i, what is the probability it will go to a different point, x_f, after a certain amount of time, Δt? And if the starting and finishing points are reversed, what will that probability be? Or, how long does it take to go from x_i to x_f? And the reverse? To answer these questions we must further define the processes we are interested in.

Our microscopic particle submerged in water, we presently ignore any energy potentials, is a Brownian particle undergoing random motion, caused by thermal fluctuations. This random motion falls into the category of a stochastic Markov process. It is stochastic because of the random forces it experiences from the water molecules. It is Markovian or *memoryless* because our particle does not experience inertia, at least not on significant timescales. We say that inertial forces are negligible in our system because they are short-lived and miniscule in comparison to the viscous force. This implies that prior movements have no bearing on future

movements because the viscous damping of the medium is always large enough to effectively stop the particle's motion instantaneously. We borrow an analogy from Ed Purcell's lovely discussion titled "Life at low Reynolds number": to experience the sensation of being 0.001mm in size and swimming in water we imagine trying to swim in a pool of molasses[56]. It is easy to picture oneself stuck in a pool of very thick molasses and not moving very fast at all, it is also easy to see that if a friend were to give a push that very little motion would follow; a 0.001 mm particle has no lingering inertia from all of those Brownian *kicks* it's receiving. Quantitatively, the Reynolds number is the ratio of the inertial forces to the viscous forces. The inertial force is $F_{inertial} = A\rho v^2$, where ρ is the fluid density, v is the velocity, and A is the surface area of the particle. The viscous force is $F_{viscous} = \nu A v/r$, where ν is the viscosity of the fluid and r is the particle radius. So the ratio of these is $R = rv/\eta$ with $\eta \equiv \nu/\rho$ being the kinematic viscosity which is roughly 10^{-2} cm^2/s for water. When someone jumps into a pool of water and starts swimming they have a Reynolds number around 10^4, our 0.001 mm Brownian particle has a Reynolds number around 1. If the inertial forces are smaller than the viscous forces, momentum will contribute very little to moving through the medium!

This is our physical system, a tiny spherical object submerged in water and trapped in a potential well. Because of the size of the particle, it is an over-damped Brownian particle undergoing random fluctuations with a future completely independent of the past. We now investigate conditional probabilities and transition times in this context.

2.2 Theory

2.2.1 Brownian motion described

As we have already stated, it was the separate and simultaneous work of Albert Einstein and Marian Smoluchowski that propelled the mathematical analysis of stochastic processes. They showed that the mean of the square of the position is proportional to the time elapsed. The constant of proportionality which relates the mean square displacement to the elapsed time is the diffusion constant, D. These are the two relationships associated with Einstein's work on Brownian motion,

$$\langle x^2 \rangle = 2Dt, \tag{2.1}$$

$$D = k_B T/\gamma, \tag{2.2}$$

where γ is the coefficient of viscous drag for the Brownian particle. For a spherical particle with radius r, we have $\gamma = 6\pi\eta r$. While Einstein was publishing his work, Marian Smoluchowski was also working independently on the same problem. Shortly after Einstein and Smoluchowski published their works, Paul Langevin derived the same results but by altogether different means. He began by writing Newton's second law with a random force $\xi(t)$ included, which in one-dimension may be written as

$$m\ddot{x} = F_{net} + \xi(t), \tag{2.3}$$

where F_{net} represents all forces other than the random force. Furthermore, Langevin showed that $\langle x^2 \rangle = [k_B T/(3\pi\eta r)]t$, a result identical to Eq.(2.1) provided $D = k_B T/(6\pi\eta r)$. The significance of Langevin's work is that he was first to solve a *stochastic differential equation*. The Langevin equation will be the starting point for all further theoretical analysis of this system.

2.2.2 Conditional probabilities: Onsager and Machlup

The ratio of two spatially inverted, time dependent conditional probabilities can be used to obtain equilibrium information about a system. By studying the inherently stochastic trajectories of a trapped Brownian particle, we further elucidate this fact. Given that the particle is contained within a potential well having the form $U(x) = kx^2/2$, where k represents the *stiffness* of the well, we ask, if the particle has position $x(t) = x^{(i)}$, what is the probability that $x(t+\Delta t) = x^{(f)}$? We write this quantity as $P(x^{(f)}, t + \Delta t|x^{(i)}, t)$. Although the quantitative analysis that follows assumes a quadratic potential, it may be generalized to an arbitrary potential/force [16].

Onsager and Machlup quantified the conditional probability of a transition (set of fluctuations) to occur for Gaussian-Markov processes in 1953 [52]. Their expression for the conditional probability can be used to show that the ratio of conditional probabilities having reversed spatial coordinates is simple, elegant, and rather curious [16]. This ratio is

$$\frac{P(x^{(f)}, t + \Delta t|x^{(i)}, t)}{P(x^{(i)}, t + \Delta t|x^{(f)}, t)} = e^{-\Delta U/k_B T}, \tag{2.4}$$

a result that remarkably tells us that the equilibrium Boltzmann distribution can be determined from a consideration of nothing more than the analysis of conditional probabilities for systems without dissipative forces. Furthermore, Bier et al. derive this result for an arbitrary force. Their analysis describes the probability that a specific trajectory is followed, this depends upon a specific sequence of fluctuations. The probability that the reverse trajectory is followed requires a time reversal in the sequence of fluctuations and the ratio of these probabilities amazingly simplifies to Eq.(2.4). We paraphrase Onsager and Machlup's derivation of $P(x^{(f)}, t + \Delta t|x^{(i)}, t)$ to obtain this result for the quadratic potential. During the derivation we will write

the action associated with the path/trajectory, this is precisely the quantity that was considered by Bier et al.

The following analysis is strictly for systems of finite variables described by a stationary Gaussian-Markov process. Again, we have assumed a time independent quadratic potential, but in the formalism of Onsager and Machlup the requirement is that, "the fluxes depend linearly on the forces that 'cause' them." That is, the flows/velocities of the random variables must be linearly related to the net forces, not the random variables themselves.

This analysis begins with the equation of motion for our system. We first write the over-damped Langevin equation, Eq.(2.3) becomes

$$\gamma\dot{x} = -kx + \xi(t), \tag{2.5}$$

where $\ddot{x} = 0$ in the over-damped limit, and $F_{net} = -\gamma\dot{x} - \nabla U$. Onsager and Machlup next write the conditional probability distribution function (p.d.f.) for a transition between two states $x^{(1)}$ and $x^{(2)}$, separated by the time interval $\Delta t = \tau$ [63],

$$f_1\left(\begin{array}{c|c} x^{(2)} & x^{(1)} \\ t+\tau & t \end{array}\right) = (2\pi(1-e^{-2k\tau/\gamma}))^{-\frac{1}{2}}\frac{k}{k_BT}\exp\left\{-\frac{k}{2k_BT}\frac{[x^{(2)}-e^{-k\tau/\gamma}x^{(1)}]^2}{(1-e^{-2k\tau/\gamma})}\right\}. \tag{2.6}$$

By separating the interval $(t, t+\tau)$ into p equal sub-intervals the conditional probability is written as

$$f_1\left(\begin{array}{c|c} x^{(p+1)} & x^{(1)} \\ t_{p+1} & t_1 \end{array}\right) = \int (p-1)\text{-fold} \int f_1\left(\begin{array}{c|c} x^{(p+1)} & x^{(1)} \\ t_{p+1} & t_1 \end{array}\right)\cdots f_1\left(\begin{array}{c|c} x^{(2)} & x^{(1)} \\ t_2 & t_1 \end{array}\right) dx^{(2)}\cdots dx^{(p)}, \tag{2.7}$$

which is the Chapman-Kolmogorov relation. The Markov nature of the process ensures that the conditional probability we are interested in depends only on the

one-step conditional probabilities in the set of events occurring during the interval $(t, t+\tau)$. As we increase p, or decrease τ, the Langevin equation can be well approximated by the difference equation,

$$x^{(j)} - \lambda x^{(j-1)} = \epsilon^{(j)}, \tag{2.8}$$

where $\lambda = 1 + k/\gamma\tau$ and $\epsilon^{(j)} = \xi^{(j)}\tau/\gamma$. Taking ϵ to be Gaussian distributed with a mean of zero and variance σ_ϵ^2 we can replace the conditional p.d.f.'s of Eq.(2.7) with exponentials,

$$f_1\left(\begin{array}{c|c} x^{(p+1)} & x^{(p)} \\ t_{p+1} & t_p \end{array} \right) \propto \int \exp\Big\{ -\frac{1}{2\sigma_\epsilon^2}[(x^{(p+1)} - \lambda x^{(p)})^2] \Big\} dx^{(p)}, \tag{2.9}$$

and Eq.(2.7) becomes,

$$f_1\left(\begin{array}{c|c} x^{(p+1)} & x^{(1)} \\ t_{p+1} & t_1 \end{array} \right) \propto \int (p-1)\text{-fold} \int \exp\Big\{ -\frac{1}{2\sigma_\epsilon^2}[(x^{(p+1)} - \lambda x^{(p)})^2 + \cdots + (x^{(2)} - \lambda x^{(1)})^2] \Big\} dx^{(p)} \cdots dx^{(1)}. \tag{2.10}$$

Finally, we argue that because of the Gaussian nature of the random process, the "average" values obtained from each integration may be replaced by the "most probable" (minimum) value (they are the same),

$$f_1\left(\begin{array}{c|c} x^{(p+1)} & x^{(1)} \\ t_{p+1} & t_1 \end{array} \right) \propto \exp\Big\{ -\frac{1}{2\sigma_\epsilon^2}[(x^{(p+1)} - \lambda x^{(p)})^2 + \cdots + (x^{(2)} - \lambda x^{(1)})^2] \Big\}_{\min}; \tag{2.11}$$

with respect to $x^{(2)}, \cdots, x^{(p)}$.

We exchange the $p-1$ integrals for a single integral by insisting the p intervals are sufficiently many so that the sum in the exponent of Eq.(2.11) may be written as an integral to be minimized with the specified end points.

If we now return to differential form, with $\sigma_\epsilon^2 = 2k_BT\tau^2/\gamma^2$, the sum may be written as

$$-\frac{1}{2\sigma_\epsilon^2}\left[\sum_{k=1}^{p}\left(x^{(k+1)}-\lambda x^{(k)}\right)^2\right]_{\min} = -\frac{1}{4k_BT}\left(\int_{t_1}^{t_{p+1}}\gamma[\dot{x}(t)+\frac{k}{\gamma}x(t)]^2dt\right)_{\min}, \quad (2.12)$$

and we may now write Eq.(2.11) as

$$f_1\left(\begin{array}{c|c} x^{(p+1)} & x^{(1)} \\ t_{p+1} & t_1 \end{array}\right) \propto \exp\left\{-\frac{1}{4k_BT}\left(\int_{t_1}^{t_{p+1}}\gamma[\dot{x}(t)+\frac{k}{\gamma}x(t)]^2dt\right)_{\min}\right\}, \quad (2.13)$$

with the constraints $x(t_1) = x^{(1)}; \cdots ; x^{(t_{p+1})} = x^{(p+1)}$. Because it was assumed that our system is Markov, it is necessary to show that Eq.(2.13) is independent of t. To prove this, Onsager and Machlup insist the necessary condition to be satisfied is that as the system regresses (moves) from state $x^{(1)}$ to $x^{(2)}$, the conditional p.d.f. is a maximum,

$$f_1\left(\begin{array}{c|c} x^{(2)} & x^{(1)} \\ t_2 & t_1 \end{array}\right) = \max. \quad (2.14)$$

They also assert that the integrand in Eq.(2.13) must be a minimum for this to be true,

$$\dot{x} + \frac{k}{\gamma}x = 0, \quad (2.15)$$

In essence, this is a proof that the velocity ("flux" in general) is linearly related to the force. Finally, to show that the conditional p.d.f. does obey this linearity (as well as the Boltzmann distribution), we expect that

$$f_1\left(\begin{array}{c|c} x^{(2)} & 0 \\ t_2 & -\infty \end{array}\right) \propto \exp\left[-\frac{1}{2}\frac{k}{k_BT}(x^{(2)})^2\right]. \quad (2.16)$$

Here the initial conditions are taken to be $t_1 = -\infty$ and $x^{(1)} = 0$, implying that we wish to consider the conditional probability the system is equilibrated at t_1

To prove this, we first write the Lagrangian, $\mathscr{L}(\dot{x}, x) = \gamma(\dot{x} - kx/\gamma)^2$, and use the Euler-Lagrange equation to find the extreme value of the integral,

$$\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{x}} - \frac{\partial \mathscr{L}}{\partial x} = 0, \tag{2.17}$$

so that

$$\ddot{x} - (\frac{k}{\gamma})^2 x = 0. \tag{2.18}$$

By inspection, the solutions to this are $e^{\pm kt/\gamma}$. The exponentially decreasing term vanishes because of the initial condition that $x(-\infty) = 0$, the condition that $x(t_2) = x^{(2)}$ means the solution is

$$x(t) = (x^{(2)})^2 e^{-k/\gamma(t_2 - t)}. \tag{2.19}$$

With this expression for $x(t)$, the solution to the integral is

$$\frac{1}{4}\int_{-\infty}^{t_2} \gamma\Big[\dot{x}(t) + \frac{k}{\gamma}x(t)\Big]^2 dt_{min} = \frac{1}{2}\frac{k}{k_B T}x^{(2)}, \tag{2.20}$$

provided that $x(-\infty) = 0$, $x(t_2) = x^{(2)}$. Comparing Eq.(2.20) with Eqs.(2.13) and (2.16) we see that the expected result has been verified. If we look at Eq.(2.6), and now consider the ratio of the forward and backward conditional probabilities, we find

$$\frac{P(x^{(2)}, t+\tau|x^{(1)}, t)}{P(x^{(1)}, t+\tau|x^{(2)}, t)} = \exp\Big\{ - \frac{k}{2k_B T}\frac{[(x^{(2)} - e^{-\beta}x^{(1)})^2 - (x^{(1)} - e^{-\beta}x^{(2)})^2]}{(1 - e^{-2\beta})}\Big\}, \tag{2.21}$$

where $\beta = k\tau/\gamma$. With some algebraic manipulation the term in square brackets becomes

$$[\cdots] = (1 - e^{-2\beta})[(x^{(2)})^2 - (x^{(1)})^2], \tag{2.22}$$

and finally,

$$\frac{P(x^{(2)}, t+\tau|x^{(1)}, t)}{P(x^{(1)}, t+\tau|x^{(2)}, t)} = e^{-\Delta U/k_B T}. \tag{2.23}$$

Here we have reverted to writing the conditional probabilities as $P(\cdots|\cdots)$ as opposed to $f_1(\cdots)$[16].

As will be seen in later sections, we can simplify experimental and computational results by considering ratios of quantities, this practice is sensible when considering trajectories and paths for stochastic systems. At first glance, Eq.(2.6) seems rather benign, almost trivially obvious, it is simply the ratio of two Boltzmann factors. However, it has been obtained by considering a ratio of conditional probabilities and *not* the ratio $P(x^{(2)})/P(x^{(1)})$. It is also important to reiterate that the conditional probability distribution function given by Eq.(2.12) requires only that $\dot{x}$ depend linearly on the force causing the velocity during the time interval of interest, τ, which allows for a time-dependent forcing, in general. This will be discussed in greater detail in Section 4. For now let us move on to considering the time to make these transitions, or follow paths, between initial and final positions. We turn to the work of Astumian, Bier, Derényi, et al.

2.2.3 First passage times: Astumian, Bier, and Derényi

We now consider the time to traverse a potential and show that it, like the probability to traverse, is important. A first passage time is a physical quantity that is measurable and intuitively accessible. The most common first passage time that is discussed is the mean first passage time (MFPT), it represents the average time it takes to pass from an initial position to a final position. The MFPT is a directly useful quantity, from a practical standpoint it represents the average amount of *time* for an event to occur. Another first passage time that has recently been discussed in the literature is the last-touch-first-touch-time (LTFTT). this is the time that elapses between the last crossing of the initial position and the first crossing of the final position [16, 29, 15, 13, 14]. This is very different from the MFPT, if $x_i < x_f$, the LTFTT does not include any time the particle may spend outside the space between the endpoints and we have $x_i < x(t) < x_f$. A cartoon showing what a LTFTT is can be found in Figure 2.1. As we have already seen with the conditional

probabilities, interesting relationships can be obtained by taking ratios of quantities. Almost unexpectedly, the ratio of the LTFTT between an initial and final position and it's spatial reverse is unity, $\mathrm{LTFTT_F} = \mathrm{LTFTT_R}$, the subscripts representing forward and backward interpretations of the direction of travel. This asserts that for stochastic Markov processes, direct "uphill" and "downhill" transitions will take equally long. We can think of this as being analogous to saying that it will take just as long to ascend as it will to descend a flight of stairs. Although that analogy isn't quite accurate due to the obvious lack of random noise, it elucidates a fundamental difference between deterministic and probabilistic systems, that our deterministic intuitions cannot always be relied upon!

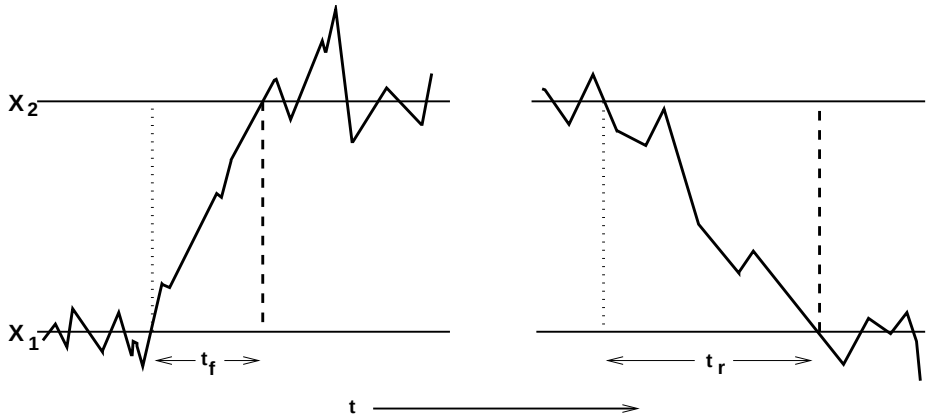

Figure 2.1. Hypothetical data of Brownian motion, or the state of a system represented in configuration space, representing a transition event. Here the LTFTT is represented by the time that is elapsed between the dotted and dashed lines. The image on the left shows a transition from x_1 to x_2 while the image on the right is an example of a transition in the opposite direction. Time progresses towards the right, the dotted line marks the time of the *last touching* of the initial position and the dashed line marks the *first touching* of the final position. On average, $t_f = t_r$.

The first mathematical derivation of the LTFTT was published by Derényi and Astumian [10]. The analysis begins with an expression for the MFPT which is then split into two events, a *wiggling* event that represents the fluctuations of the particle

around the initial position before last touching the initial position (first gate) and moving to the final position (second gate). The initial analysis follows from Section 5.2.7 of [36] where the associated Fokker-Planck equation is written down and from it the MFPT is derived for one reflecting boundary and one absorbing boundary. The justification for the reflecting boundary is that we insist that the particle not leave the interval in question. The absorbing boundary ensures that once the particle reaches the final position, it is not allowed back into the interval. The Fokker-Planck equation (FPE) describes the time evolution of the probability density $P(x,t)$. For an over-damped, Brownian particle in a potential $U(x)$, the FPE is

$$\dot{P}(x,t) = -J'(x,t). \tag{2.24}$$

The probability current is given by

$$J(x,t) = -\frac{1}{\gamma}U'(x)P(x,t) - DP'(x,t), \tag{2.25}$$

so that,

$$\dot{P}(x,t) = -\frac{\partial}{\partial x}\Big[-\frac{1}{\gamma}U'(x)P(x,t) - D\partial_x P(x,t)\Big]. \tag{2.26}$$

By comparing Eq.(2.26) with the general form given by Gardiner,

$$\dot{P} = -\partial_x[A(x,t)P(x,t)] + \frac{1}{2}\partial_x^2[B(x,t)P(x,t)], \tag{2.27}$$

we see that $A(x,t) = -U'(x)/\gamma$ and $B(x,t) = 2D$. We will now discuss the method outlined by Gardiner to obtain the MFPT in terms of A and B. Afterwards, the results will be applied to our system.

We consider a particle having position x at $t = 0$ and want to know how long it remains in the interval (a,b) where $a \leqslant x \leqslant b$. If the particle is in the interval (a,b) it has never left. The probability that at time t the particle is in the interval must be $\int_a^b dx' p(x',t|x,0) \equiv G(x,t) = P(T \geqslant t)$, where $t = T$ is the time when the particle is absorbed by one of the barriers. We can write down the initial conditions

$G(x,0) = \{1$ for a $\leqslant$ x $\leqslant$ b; 0 elsewhere. On arriving at one of the barriers, $x = a$ or b, we have $P(T \geqslant t) = 0$. That is, the MFPT T must be less than or equal to t since at that time the particle has been absorbed and $T \ngeqslant t$, equivalent to $G(a,t) = G(b,t) = 0$. Because we have specified the final conditions on x and are interested in the time development in the sense of the "initial time", the backward FPE is used, it's general form is slightly different than the forward-time FPE of Eq.(2.27),

$$\partial_t p(x',t|x,0) = A(x)\partial_x p(x',t|x,0) + \frac{1}{2}B(x)\partial_x^2 p(x',t|x,0), \tag{2.28}$$

when written in terms of $G(x,t)$ this becomes

$$\partial_t G(x,t) = A(x)G(x,t) + \frac{1}{2}B(x)G(x,t). \tag{2.29}$$

In order to cast Eq.(2.29) in terms of T we first define the mean of any function of T in terms of $G(x,t)$, the probability that $T > t$, as $\langle f(T)\rangle = -\int_0^\infty f(t)dG(x,t)$. Thus, the MFPT $T(x) = \langle T\rangle$ and is

$$\begin{aligned} T(x) &= -\int_0^\infty t\partial_t G(x,t)dt, \\ &= \int_0^\infty G(x,t)dt, \end{aligned} \tag{2.30}$$

after integrating by parts. Integrating the l.h.s. of Eq.(2.29) with respect to time we see that

$$\int_0^\infty \partial_t G(x,t)dt = -1, \tag{2.31}$$

since $G(x,\infty) = 0$ and $G(x,0) = 1$. A differential equation in $T(x)$,

$$A(x)\partial_x T(x) + \frac{1}{2}B(x)\partial_x^2 T(x) = -1, \tag{2.32}$$

is obtained by performing the same integration on the r.h.s. of Eq.(2.29). This is done by moving the integrations inside the partial derivatives and then replacing

the integrals over $G(x,t)$ with $T(x)$. The solution to Eq.(2.32) has the form $T(x) = v_1(x)y_1(x)+v_2(x)y_2(x)$ where the y_i's are the solutions to the homogeneous equation,

$$\partial_x^2 y(x) + \frac{2A(x)}{B(x)}\partial_x y(x) = 0, \tag{2.33}$$

which has two solutions, the first being an arbitrary constant, c_1, the second is $g(x)y_1(x)$. In general, $g(x) = C\int y_1^{-2}\exp(-\int P(\eta)d\eta)d\xi$, with $y_1 = c_1$ and $P(x) = 2A(x)/B(x)$. Now, $g(x)$ is

$$g(x) = c_2\int^x \frac{d\xi}{\psi(\xi)}, \tag{2.34}$$

where $\psi(x) = \exp(\int 2A(x)/B(x)dx)$. We must now find the v_i's,

$$\begin{aligned} v_1(x) &= -\int_{x_1}^x \frac{y_2(\xi)R(\xi)}{W(\xi)}d\xi \\ &\propto 2\int^x \Big[\int^\xi \frac{dz}{\psi(z)}\Big]\frac{\psi(\xi)}{B(x)}, \end{aligned} \tag{2.35}$$

$$\begin{aligned} v_2(x) &= \int_{x_2}^x \frac{y_1(\xi)R(\xi)}{W(\xi)}d\xi \\ &\propto 2\int^x \frac{\psi(y)dy}{B(y)}, \end{aligned} \tag{2.36}$$

where $R(x) = -2/B(x)$ and $W(x) \propto \psi(x)^{-1}$ is the Wronskian. The general solution with arbitrary lower limits on the integrals is

$$T(x) = c_1 2\int^x \Big[\int^y \frac{dz}{\psi(z)}\Big]\frac{\psi(y)dy}{B(y)} + c_2 2\int^x \frac{\psi(y)dy}{B(y)}\int^y \frac{dz}{\psi(z)}. \tag{2.37}$$

After manipulation, $T(x)$ becomes

$$T(x) = \frac{2\left[\left(\int_a^x \frac{dy}{\psi(y)}\right)\int_x^b \frac{dy'}{\psi(y')}\int_a^{y'}\frac{dz\psi(z)}{B(z)} - \left(\int_x^b \frac{dy}{\psi(y)}\right)\int_a^x \frac{dy'}{\psi(y'}\int_a^{y'}\frac{dz\psi(z)}{B(z)}\right]}{\int_a^b \frac{dy}{\psi(y)}}. \tag{2.38}$$

If we now let the boundary at a be reflecting, the boundary conditions are $\partial_x G(a,t) = 0$ and $G(b,t) = 0$ and the solution to Eq.(2.32) becomes,

$$T_a(x) = 2\int_x^b \frac{dy}{\psi(y)}\int_a^y \frac{\psi(z)}{B(z)}dz, \tag{2.39}$$

requiring that b be reflecting gives

$$T_b(x) = 2\int_a^x \frac{dy}{\psi(y)}\int_y^b \frac{\psi(z)}{B(z)}dz, \tag{2.40}$$

where the subscript on $T_i(x)$ refers to the relevant reflecting boundary. These expressions for the MFPT become much simpler in the context of Brownian motion in a potential. In this context we have $B(x) = 2D$ and $A(x) = U'(x)/\gamma$, and $\psi(x) = \exp[(U(x) - U(a))/k_BT]$. For Eqs.(2.39) and (2.40) the MFPT's become

$$T_a[x_i \to x] = \frac{1}{D}\int_{x_i}^x \int_a^y e^{[U(y)-U(z)]/k_BT}dzdy \quad \text{if } x_i \le x, \tag{2.41}$$

$$T_b[x_i \to x] = \frac{1}{D}\int_x^{x_i} \int_y^b e^{[U(y)-U(z)]/k_BT}dzdy \quad \text{if } x \le x_i. \tag{2.42}$$

We now assume that the MFPT is the sum of two separately defined times. Let us define the transition interval with the initial position a and final position b. The first of these two times represents the time during which the particle *wiggles* around the initial position a, defined as the time that elapses between the first and last touching of a. The second time represents the actual transition from a to b, precisely defined by the time that elapses between the last touching of a and the first touching of b. In [29] these are the "wiggling" and "instanton" times, they are derived from a quantity called the intra-well relaxation time (IRT). The IRT is the MFPT from an initial position given by the Dirac δ-function $\delta(x(0)-a)$, to a Boltzmann distributed final position, $B(x) = 1/Z\exp(-U(x)/k_BT)$, where $Z = \int_a^b \exp(-U(x)/k_BT)dx$ is the normalization constant. Obtaining the IRT is straightforward because of the additive nature of the MFPTs, that is, $T[a \to x_2] = T[a \to x_1] + T[x_1 \to x_2]$. This additivity allows us to write the IRT as $T[a \to B(x)] = T[a \to b] - T[B(x) \to b]$,

the first of these terms $T(a \to b)$ we have in Eq.(2.41), the second we now find. The MFPT for a Boltzmann distributed particle to reach the barrier at b is obtained by taking the average of Eq.(2.39) over $B(x)$ so that

$$T[B(x) \to b] = \frac{1}{DZ} \int_a^b \int_x^b \int_a^y e^{[-U(x)+U(y)-U(z)]/k_B T} dz dy dx. \tag{2.43}$$

To put $T[a \to b]$ into a similar form we reverse the order of integration and insert a well chosen 1 into the expression,

$$\begin{aligned} T[a \to b] &= \frac{1}{D} \int_a^b \int_a^y e^{[U(y)-U(x)]/k_B T} dx dy, \\ &= \frac{1}{D} \int_a^b \int_x^b e^{[U(y)-U(x)]/k_B T} dy dx, \\ &= \frac{1}{D} \int_a^b \int_x^b \left[\frac{\int_a^b e^{-U(z)/k_B T} dz}{Z} \right] e^{[U(y)-U(z)]/k_B T} dy dx, \\ &= \frac{1}{DZ} \int_a^b \int_x^b \int_a^b e^{[-U(z)+U(y)-U(x)]/k_B T} dz dy dx. \end{aligned} \tag{2.44}$$

Now we write the IRT as,

$$T[a \to B(x)] = \frac{1}{DZ} \int_a^b \int_x^b \int_y^b e^{[-U(z)+U(y)-U(x)]/k_B T} dz dy dx \tag{2.45}$$

It can also be shown that the IRT is exactly equal to $T[b \to a] - T[b \to B(x)]$,

$$T[a \to b] - T[B(x) \to b] = T[b \to a] - T[b \to B(x)]. \tag{2.46}$$

This equality suggests that certain passage times are independent of the direction of travel.

The LTFTT requires that the boundary condition at the final position be absorbing, this condition alters the probability distribution by allowing an outflow of probability at b. This is addressed in Ref. [29], it is argued that for a sufficiently deep well, this outflow will be small and can be neglected, thereby allowing the IRT to be approximated by that for reflecting boundaries. Instead of a Boltzmann distribution we now must use a stationary distribution within the well on account

of the absorbing boundary. We obtain an expression for the stationary distribution by setting the l.h.s. of Eq.(2.24) to zero. Thus, the probability current is constant, $J(x,t) = J(x) = C$, i.e. $dJ(x)/dx = 0$. The stationary distribution is the solution to the differential equation,

$$0 = -\frac{1}{\gamma}U'(x)P(x) - DP'(x). \tag{2.47}$$

Multiplying by the integrating factor, $\alpha(x)$, which is $\int e^{(U(x)/k_BT)}dx$, the stationary distribution is

$$S(x) = \frac{1}{Y}\int_x^b e^{[-U(x)+U(y)]/k_BT}dy, \tag{2.48}$$

where

$$Y = \int_a^b \int_x^b e^{[-U(x)+U(y)]/k_BT}dydx. \tag{2.49}$$

Taking this result and rewriting the IRT in terms of $S(x)$ we find that

$$T[a \leftrightarrow S(x)] = \frac{1}{DY}\int_a^b \int_x^b \int_y^b \int_z^b e^{[-U(z)+U(y)-U(x)+U(v)]/k_BT}dvdzdydx, \tag{2.50}$$

a result that at first glance seems only a more complicated version of Eq.(2.45), however, this can be shown to be identical to what is called the mean instanton time (MIT) between b and a stationary distributed position. In Ref. [29] the MIT is stated to be the average time to travel from b to $S(x)$ (or vice versa) after the last touching of the initial position. The expression for the MIT between two points a and b is,

$$\mathcal{I}[a \leftrightarrow b] = \frac{1}{DZ'}\int_a^b \int_y^b \int_z^b e^{[U(y)-U(z)+U(v)]/k_BT}dvdzdy, \tag{2.51}$$

where $Z' = \int_a^b e^{U(y)/k_BT}dy$ is a normalization factor which uses the inverted potential, $-U(y)$. The MIT between two points represents the average time that elapses between the last touching of the initial position and the first touching of the final position, we have already defined this as the last-touch-first-touch time (LTFTT). The time that is spent around one endpoint before touching it for the last time and

moving to the other endpoint is called the mean wiggling time (MWT), $\mathcal{W}[x_i]$. Also we have the following expression,

$$\mathcal{I}[b \leftrightarrow x] = T_b[b \rightarrow x] - \mathcal{W}[b]. \tag{2.52}$$

Eq.(2.52) tells us that the MFPT can indeed be thought of as the sum of two independent processes. One being the time that the particle spends around the initial position "wiggling" and the other being the average time for the particle to travel directly between the endpoints, the LTFTT. Since $\mathcal{I}[b \leftrightarrow x]$ is independent of the direction of travel we have the following equality, as predicted by microscopic reversibility at equilibrium,

$$\mathcal{I}[b \rightarrow x] = \mathcal{I}[x \rightarrow b], \tag{2.53}$$

or,

$$\text{LTFTT[b} \rightarrow \text{x]} = \text{LTFTT[x} \rightarrow \text{b]}. \tag{2.54}$$

Chapter 3
METHODS

To study the random fluctuations of a Brownian particle we have used an optical tweezer. An optical tweezer confines the particle to a localized region of space. Depending on the intended use, an optical tweezer apparatus can be rather simple or impressively complicated. Single-trap experiments are relatively simple to perform and require minimal equipment, the more complicated experiments involve multiple lasers, complicated optical systems, mechano-optical devices, and piezoelectric controls. To confine a single Brownian particle and study it's motion is perhaps the simplest and most basic use of the optical tweezer, but the extent of its relevance cannot be overstated, it is an indispensable tool [6]. We discuss the history of optical trapping in relation to colloidal particles as well as the physics of optical trapping. Following this, we describe our experimental apparatus and methods used to investigate the conditional probabilities and passage times previously discussed.

3.1 The history of optical trapping

The first evidence that mesoscopic particle dynamics could be controlled by the radiation pressure of light was presented in 1970 by Arthur Ashkin [1]. The intent of those first experiments was to show that radiation pressure from a mildly focused Gaussian laser, incident on a transparent sphere (diameter $\sim \lambda_{light}$), would produce observable, directed motion. Yes, Ashkin observed the motion he had predicted, motion along the laser's axis of propagation (the optical axis). But he observed another motion, an unexpected motion. This secondary particle motion tended towards the axis. There was a radial force pulling the particle into the

beam-axis if it drifted away. With the beam blocked, the particle would freely diffuse, but when the laser was unblocked the particle would return to the axis! Because of this restoring force, perpendicular to the optical axis, Ashkin was able to create the first stable, three-dimensional optical trap using two opposing laser beams. The two opposing beams balanced the forces along the optical axis while they each imparted restoring forces perpendicular to the optical axis. The two-beam optical trap, which creates stronger trapping forces than are typical for a single trap with a given laser power, has been used for interesting experiments involving the stretching/pulling of DNA/motor proteins [34, 65]. This increased trap strength is accomplished by effectively canceling the axial forces that would tend to push the particle *downstream*. Following those first examples of optical trapping Ashkin showed that a particle could be levitated by balancing the axial radiation pressure and the downward force due to gravity [2]. The trap of greatest significance to the biological sciences, and most relevant to this work, was not developed until 1986.

During a period of difficulty, while trying to create an atomic trap, Ashkin and his colleagues tried their experiment on colloidal particles to test their "failing" method. They were successful. Though at that time they weren't able to trap atoms, they were able to trap colloids. Their new trap was both unique and more accessible to the biological sciences. What they showed was that stable, three-dimensional trapping was possible with a single, highly focused laser beam [22]. When a high numerical aperture (NA) microscope objective is used to focus the laser, the direction of the axial force (that would typically push the particle "downstream") becomes dependent upon where along the optical-axis the particle is, relative to the focal point. If the particle is beyond the focal point, it is pulled towards the focal point. If the particle is in front of the focal point, it is again pulled towards the focal point. With this restoring force and the already known radial restoring force, the first stable, three-dimensional optical trap using a single laser was created. Once

Ashkin and his colleagues had shown that it was possible to trap particles $\sim 1\mu m$ in size, they began trapping biological materials such as viruses and cells. In these first experiments they found that their laser was literally killing the objects they were trapping due to the excessive heating from the laser, a process termed "opticution" [3]. The opticution dilemma was resolved by switching from a visible laser to an infrared laser. Furthermore, the use of a high-resolution, optical microscope allowed them to trap and observe biological processes indefinitely [3, 5, 4]. There has been a tremendous amount of work using optical tweezers and two excellent reviews discussing the many varied experiments can be found in Refs. [6] and [46]. The primary aim of this work is to emphasize that from experiments with single optical traps, our results are fundamental and universally applicable to systems described by over-damped Gaussian-Markov processes.

3.2 The physics of optical trapping

Arthur Ashkin's interest in optical trapping was sparked by this back-of-the-envelope calculation: consider a continuous wave laser, with power $P = 1$ W, incident upon a particle with radius approximately the wavelength of the laser $r \sim \lambda_{light}$. If we assume that a fraction, $q = 0.1$, of the light is reflected by the particle, the force that arises from the radiation pressure (due to the momentum change of the photons) will be $F_{rad} \sim 2qP/c \approx 10^{-10}N$ or 10^{-5} dyne, where c is the speed of light in vacuum. This is a small force, in fact it is a tiny force. But tiny forces can push tiny objects and a particle of diameter $d = 1\mu\text{m} \sim 2\lambda_{\text{light}}$, with density $\rho_p = 1g/\text{cm}^3$, will experience an acceleration of $a = 1.2 \times 10^8 \text{cm/s}^2$, that's roughly 10^5 times the acceleration due to gravity. This calculation is useful insomuch as it points out that laser light is sufficient to exert significant forces on microscopic objects. We present two different explanations of how a stable three-dimensional

optical trap/tweezer can be created from a highly focused laser: the first explanation is more quantitative in nature, it deals with the force experienced by a dielectric sphere in an electric field; the second is more qualitative and intuitive, by tracing light rays as they are refracted by the particle, and noting the momentum change of the light, the restoring forces are uncovered.

A general method for calculating the force on a trapped particle of arbitrary size R (R may be $>$, $<$, or $= \lambda_{light}$) is given by Tlusty et al. in [62] and presented here. We work under the assumption that the particle to be trapped is both spherical and made of a linear dielectric material. The force that arises from the interaction of an electromagnetic field with a dielectric particle is[40, 43],

$$\mathbf{F} = \oint_{\mathcal{S}} \overleftrightarrow{\mathbf{T}} \cdot d\mathbf{a} - \epsilon_0 \mu_0 \frac{d}{dt} \int_{\mathcal{V}} \mathbf{S} d\tau, \tag{3.1}$$

where ϵ_0 is the electric permittivity of free space, and μ_0 is the magnetic permeability of free space. The integrals are over the surface and volume of the particle, $\mathbf{S}$ is the Poynting vector, and $\overleftrightarrow{\mathbf{T}}$ is the *Maxwell stress tensor* with elements

$$T_{ij} = \epsilon_0 \left(E_i E_j - \frac{1}{2} \delta_{ij} E^2 \right) + \frac{1}{\mu_0} \left(B_i B_j - \frac{1}{2} \delta_{ij} B^2 \right). \tag{3.2}$$

The indices i and j represent the cartesian coordinates x, y, and z so that there are nine elements in the tensor. For a Gaussian laser (TEM_{00} mode) the most significant terms in the tensor come from the electrostatic part. So $T_{ij} \approx \epsilon_p / 4\pi [E_i E_j - \frac{1}{2} \delta_{ij} E^2]$, where the permittivity of the dielectric particle is $\epsilon_p = 4\pi\epsilon_0$. Since the particle is a linear dielectric, we write the force that it experiences in an electric field as the derivative of the dipole interaction with respect to its position, $-\frac{1}{2} \int E_0 \cdot P d\tau$, where the polarization $P = \chi E$ and the dielectric susceptibility is $\chi = (\epsilon - 1)/4\pi$. We use the unperturbed electric field, E_0, because we assume that the particle is weakly polarized, $\chi \ll 1$.

With these simplifications, the energy associated with the dipole interaction now becomes

$$\begin{aligned} W &= -\frac{1}{2}\int E_0 \cdot P d\tau \\ &= -\alpha \int I d\tau, \end{aligned} \tag{3.3}$$

where $\alpha = (\epsilon_p - \epsilon_0)/\epsilon_0$, and the unperturbed intensity (energy density) is $I = \epsilon_0 E_o^2/8\pi$. The force exerted on the particle can be found by taking the gradient of W with respect to a change in the particle's position, $\mathbf{F} = \nabla W$. To evaluate this gradient we need a reasonable expression for I. Tlusty et al. use an azimuthally symmetric intensity,

$$I(\rho, z) = I_0 \exp\left(-\frac{\rho}{2\omega^2} - \frac{z^2}{2\omega^2\varepsilon^2}\right), \tag{3.4}$$

here I is in cylindrical coordinates, ρ is the radial coordinate, z is the coordinate parallel to the beam axis, the dimensions of the beam waist are ω and $\omega\varepsilon$ in the transverse and axial directions, respectively. The anisotropy of the beam is expressed by the coefficient ε, for $\varepsilon = 1$ the trap will be symmetric in ρ and z. The origin is chosen to be coincident with the focal point ($\rho = z = 0$).

We expect that the force will be restoring for the region of space near the origin, however, not necessarily for all space because of the finite nature of the trap. To illustrate this for the isotropic trap the radial component of the restoring force we have

$$\begin{aligned} F(r) &= \partial_r W(r) \\ &= 4\pi\alpha I_0 \frac{\omega^4}{r^2} e^{-\left(R^2+r^2\right)/2\omega^2} \left[\frac{Rr}{\omega^2}\cosh\left(Rr/\omega^2\right) - \sinh\left(Rr/\omega^2\right)\right]. \end{aligned} \tag{3.5}$$

This expression is obtained by taking the radial component of the gradient of Eq.(3.3) in spherical-polar coordinates. Also, the particle radius and position have

been normalized to the beam waist so that $a = R/\omega$ and $u = r/\omega$. The restoring force is[1]

$$F(r) = 4\pi\alpha I_0 \frac{\omega^2}{u^2} e^{-\left(a^2+u^2\right)/2} \left[au\cosh(au) - \sinh(au)\right]. \tag{3.6}$$

We now consider this force in the limits of big and small R relative to the size of the focal point ω. If $R \ll \omega$ this corresponds to $a \to 0$ in Eq.(3.6) and we have

$$\begin{aligned} F_{R\ll\omega} &\simeq V\alpha\nabla I \\ &= \frac{4}{3}a^3\alpha I_0\omega^2 u e^{-u^2/2}, \end{aligned} \tag{3.7}$$

the particle is assumed to be small enough that the dipole interaction energy is simply the product of the intensity at the particle's center of mass and its volume, hence the force being proportional to $V\nabla I$. Figure 3.1 contains a plot of the restoring force $F_{R\ll\omega}$, it is implicit that when positive, this force acts to push the particle *towards* $r = 0$. The restoring force is approximately linear near the origin and increases to it's maximum at $r = \omega$ where $F_{max} = 4/3\pi\alpha I_0 a^3\sqrt{e}$. As r increases from ω the force weakens until it is negligible at $r/\omega \approx 5$.

When R is large compared to ω the restoring force is quite different for the isotropic trap. In the limit that $R \gg \omega$ we have

$$F_{R\gg\omega} \simeq 2\pi\alpha I_0\omega^2 \frac{a}{u} e^{-(u-a)^2/2}. \tag{3.8}$$

The symmetry in the system causes the force to be negligible for small r, that is, only at the fringes of the trap will there be a significant restoring force. The effect of this is clearly visible in Figure 3.2, the force is negligible until r approaches R where it peaks at $r = R$ with maximum $F_{max} \simeq 2\pi\alpha I_0\omega^2$.

In practice, optical tweezers aren't isotropic and it is unreasonable to assume that $\varepsilon = 1$. However, it is instructive to consider the simpler case of $\epsilon = 1$. In

[1]It should be noted that the exact form of Eq.(3.6) differs from the published result in [62] by a factor of $1/u^2$. Through private communication with the authors we have verified a typographical error in the published article.

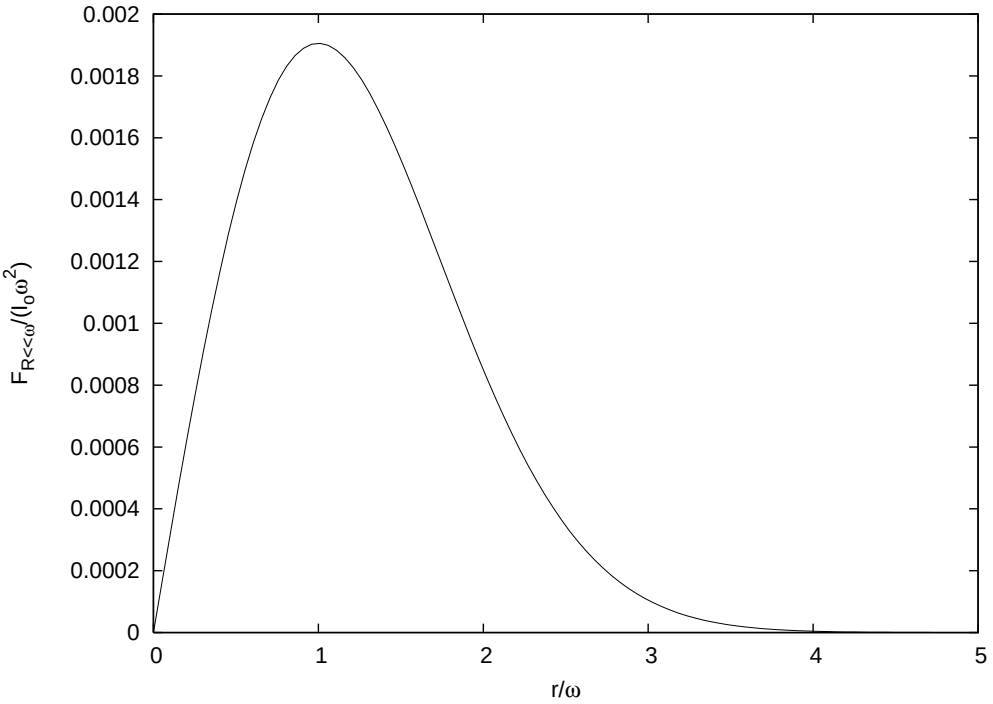

Figure 3.1. The restoring force in the limit $R \ll \omega$ with constants $\alpha = I_0 = \omega = 1$ in Eq.(3.7). To accommodate the limit of small R the particle radius is set such that $R/\omega = 0.1$ in this plot.

our experiments, the particle motion in the direction of laser propagation has been ignored. That said, Tlusty et al. further analyze the anisotropic tweezer. From their analysis of an anisotropic tweezer it may be concluded that the azimuthal symmetry in the force around the axis of propagation is preserved.

A more qualitative analysis using ray diagrams can also be used to explain how an optical tweezer works. We show this for a dielectric particle with index of refraction n_p, surrounded by a medium with index of refraction n_m, where $n_p > n_m$. We will conclude that the particle will be in stable equilibrium near the focal point. Each of the following arguments is rooted in momentum conservation. Photons incident on the particle will be refracted, because the particle does not, in general, have parallel sides. The transmitted photons will exit the particle in a direction differing from their incident direction and the net change in momentum of the photons will point

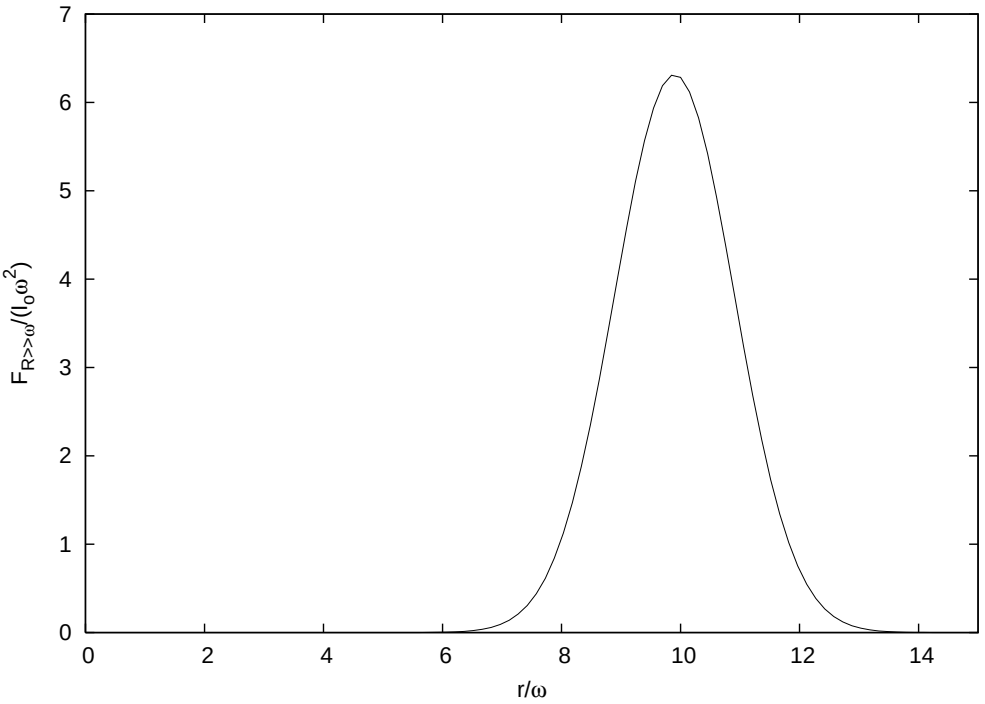

Figure 3.2. The restoring force in the limit $R \gg \omega$ with the constants $\alpha = I_0 = \omega = 1$ in Eq.(3.8). For this plot $R/\omega = 10$, which is representative of the limit for large R.

in the direction of the net force exerted on the particle. This is nothing more than an assertion of Newton's Second Law, $\mathbf{F} = \dot{\mathbf{p}}$. When the particle is located radially away from the beam-axis we need to be certain that there is a force that is directed towards the beam-axis. Because of the Gaussian profile of the laser, the rays that will have the most significant contribution are those located near the center of the beam. Figure 3.3 shows a central ray incident on a particle with it's center located left of the beam-axis. The incident ray is refracted towards a line normal to the point on the surface where the ray and particle intersect, the ray follows a straight line through the particle (n_p is constant throughout the particle), finally, at the exiting surface the ray is refracted away from the line normal to the surface/ray intersection. We see that the exiting ray points away from the beam-axis and still downward. This indicates that the momentum change $\Delta\mathbf{p}$ must have a component

pointing *towards* the beam-axis, as we expect. At present we are not concerned with the scattering force that is parallel or anti-parallel to the beam-axis, only the transverse force, labeled $F_{gradient}$. The reflected rays have been ignored and are assumed to not significantly contribute to the net force.

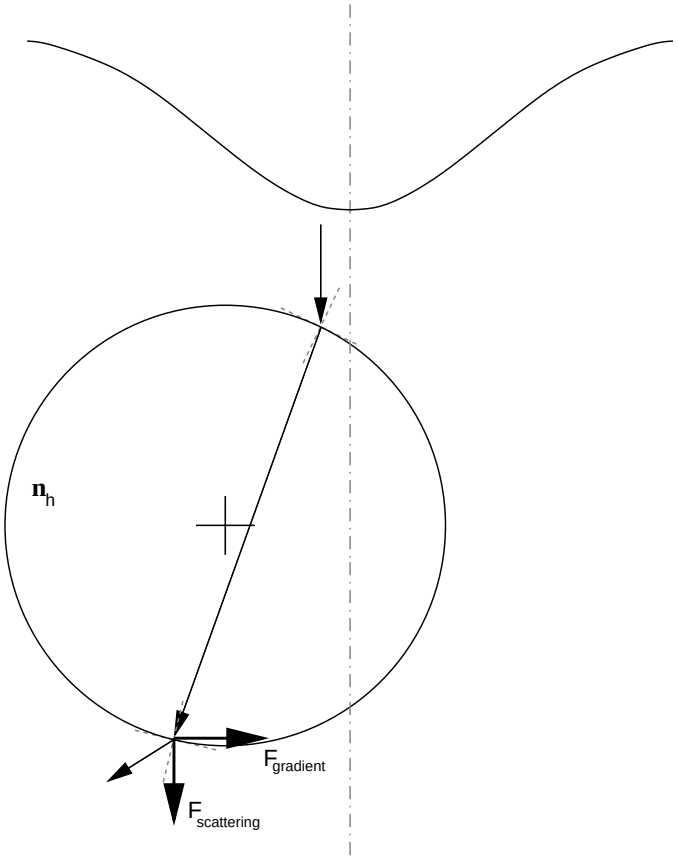

Figure 3.3. A single ray can sufficiently describe the force acting on a particle displaced from the beam-axis. Here n_h indicates that the particle has a larger index of refraction than the incident medium.

We now consider the situation in Figure 3.4, here the particle center is located on the beam-axis but below the focal point. The force on the particle can be described by two incident rays. For a symmetric particle and a Gaussian laser, the transverse forces cancel and we have $\Delta p_{transverse} = 0$. All arguments about how the light is refracted through the particle are identical to those used for the situation where the

particle is off-axis. For this arrangement, the highly divergent nature of the beam will cause the incident rays to be refracted in such a way that there will be a net force that pulls the particle towards the focal point. Here we see that the change in direction for each ray has a significant component parallel to the beam-axis, thus there is a restoring force that will pull the particle back toward the focal point. An analogous treatment can be carried out to yield a similar restoring force when the particle is centered above the focal point.

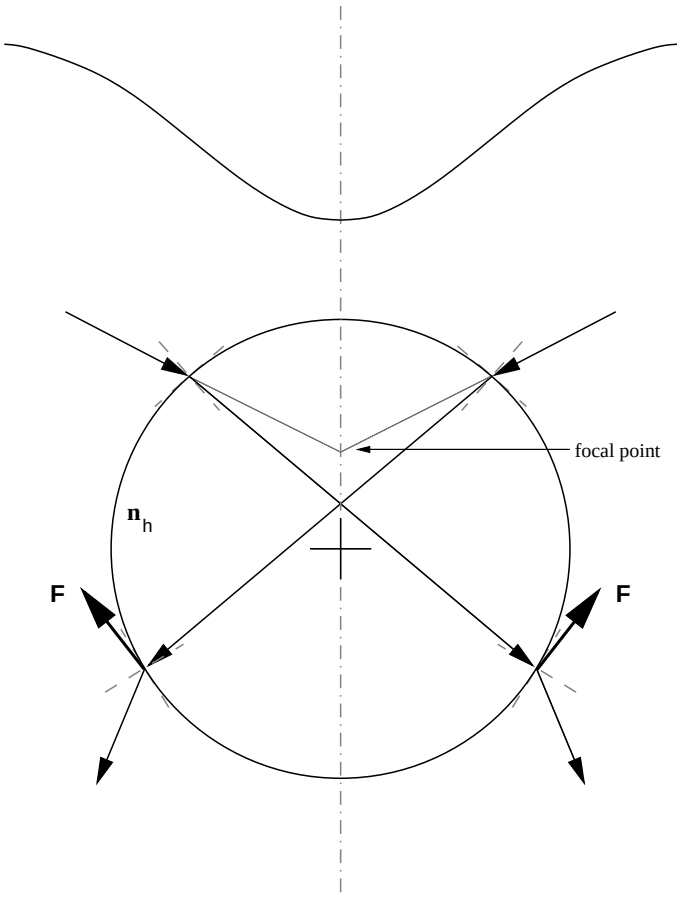

Figure 3.4. A ray trace for a particle with its center below the focal plane of the laser. If the particle center is above or below the focal point two rays are needed to show that the forces perpendicular $F_{gradient}$ sum to zero while the forces parallel to the beam-axis $F_{scattering}$ are additive and push the particle towards the focal point. Here the particle center is below the focal point.

Our treatment here has served to explain the primary behavior of a dielectric particle in an optical tweezer. There are a number of different analytical treatments describing this phenomenon, however, recognizing that the particle will be trapped within a potential well is the most important aspect to our work. For every theoretical treatment of an optical trap there is at least one different design for an actual optical trap. Our apparatus is perhaps the most simple, relying upon only a few components.

3.3 Experimental setup

Experiments involving optical tweezers are widely varied in complexity, ranging from the simplest, described here, to the most elaborate experiments that make measurements on single molecules [45, 46, 59, 65] or the multiple trap experiments of [30, 58, 55, 48]. To make measurements of the position of a trapped Brownian particle we use only a few components in our apparatus. The necessary items and their functions are listed in Table 3.1, excluding an optical table/bench and the necessary mounts and hardware.

Figure 3.5 shows a schematic of an optical tweezer. The laser is set to the approximate height of the rear microscope port and we use the two mirrors to fine-tune the alignment. The two lenses are chosen so that the laser is expanded to fill the back aperture of the microscope. The dichroic mirror is mounted inside the microscope. Depending on whether or not we wish to look at the laser, we use either a neutral density filter or a laser filter in front of the camera. The camera is attached to a computer for image collection and later analysis. The specific items that we have used in the experiments described in this work were chosen for a holographic optical tweezer (HOT) apparatus that was never made to function properly. HOTs can be used to make arbitrary arrays of optical tweezers using a single laser [30].

Item	Description
Laser	The light source for the trap.
Microscope objective	This serves two purposes, the first is to focus the laser and create the trap, the second is to allow the trapped particle to be observed.
Observation device	The means of monitoring the trapped particle. Typically a camera, alternatively a quadrant photo-diode.
Dichroic mirror	This mirror reflects/directs the laser into the objective while allowing light of different wavelengths to pass through to the observation device.
Filter	A filter to prevent laser light from reaching the observation device and interfering with particle detection/location.
Illumination source	A source of illumination for the trapped particle.
Sample	A sample containing the particle/s to be trapped.

Table 3.1. The necessary items for an optical tweezer apparatus.

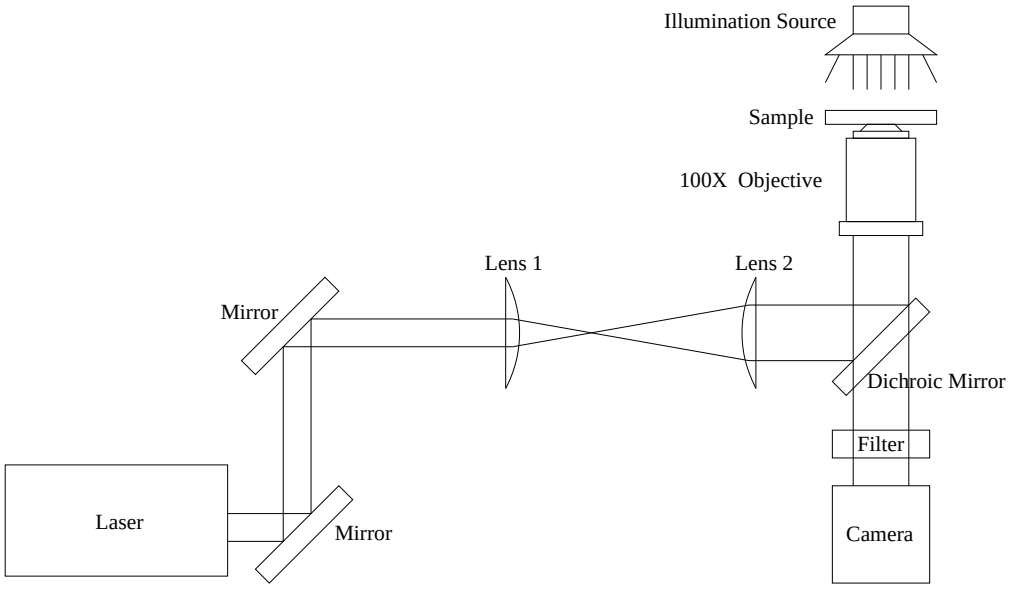

Figure 3.5. A simple optical tweezer apparatus.

Following is a detailed description of each of the pieces of equipment used.

3.3.1 Laser

The ATC-Semiconductor Devices ATC 53-250 is a Russian manufactured, green, diode laser with a maximum rated power of 250 mW. The specifications supplied with the laser are listed in Table 3.2.

Parameter	Specification
Wavelength	0.53 ± 0.03 μm
Maximum laser power	0.25 W
Transverse mode	TEM_{00}
Diameter of beam	< 0.5 mm
Divergence of beam	< 0.002 rad
Position stability of light spot	< 0.0002 rad
Stability of laser power (for 1 hour)	$< \pm 5\%$
Operating temperature	29°C
Maximum pump current	< 4000 mA

Table 3.2. The specifications supplied with the laser from the manufacturer ATC-Semiconductor Devices.

The maximum rated power of $0.25W$ is not terribly important for trapping a single particle, in fact, trapping a single particle can be accomplished with a laser power P_{trap} in the 1 to 2 mW range. Figure 3.6 is a comparison of the current drawn by the laser and the measured laser power. All experiments were performed with a fixed current, thus the stability at a given power is of much greater importance than the output power as a function of current drawn. Since $P_{trap} \sim 1$ mW for a single trap, laser stability at higher powers was never a concern, measurements were made to check laser output power at a fixed current, over long times. The claimed laser stability of $\pm 5\%$ was not found to be always true after an hour of operation, as can be seen in Figure 3.7. Another property of the laser that is crucial in trapping experiments is the beam profile.

This laser is designed to produce a Gaussian beam profile, the TEM_{00} mode. For the majority of optical trapping experiments a symmetric trap is preferred. The

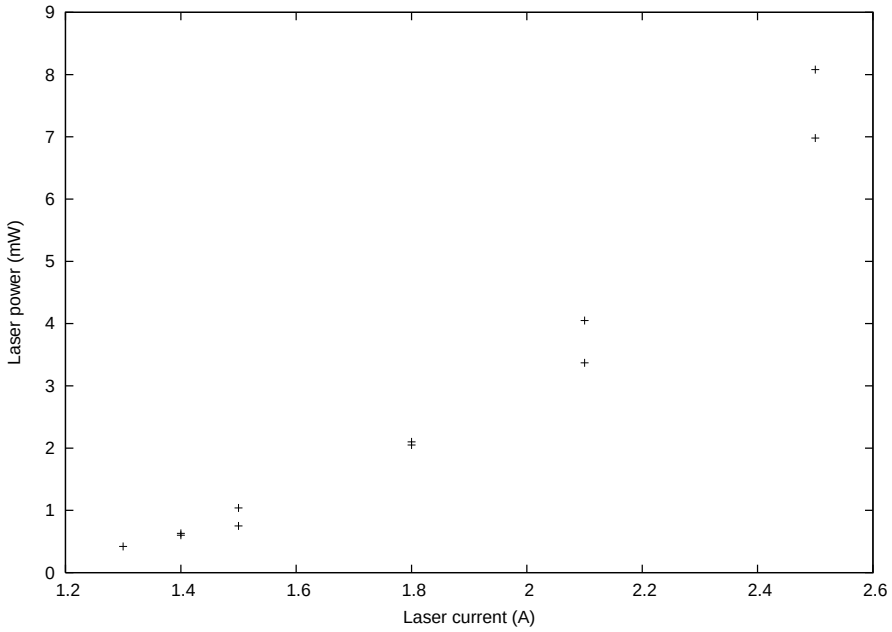

Figure 3.6. Measured laser power vs. current for the ATC 53-250.

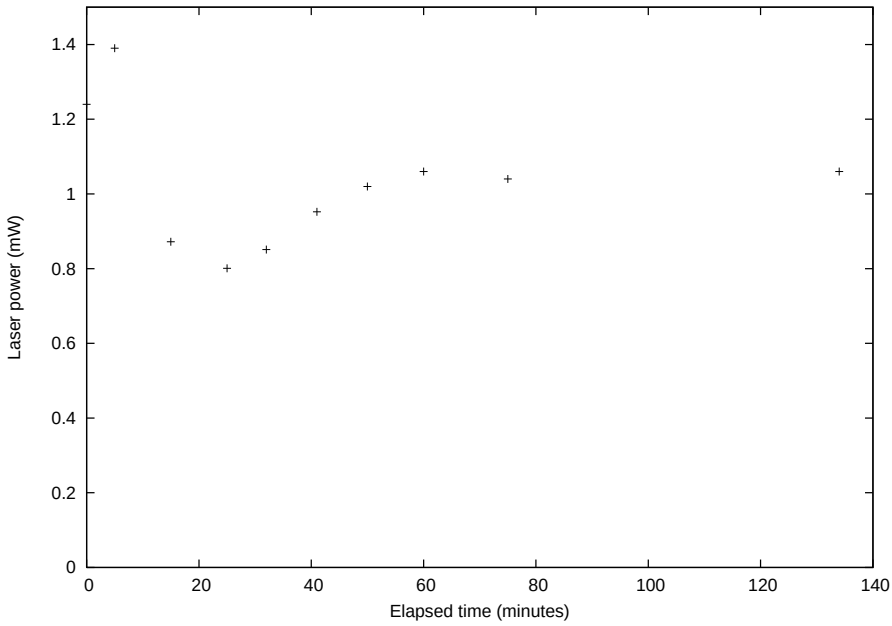

Figure 3.7. Measured laser power as a function of time with the laser current set to 1.5 A for the ATC 53-250.

TEM_{00} mode is azimuthally symmetric about the beam-axis so that with properly aligned optics a symmetric trap is produced. The symmetry of the laser was checked by collecting a two-dimensional image of the laser and analyzing its brightness. By placing a mirror over the microscope objective, the focused spot was reflected back through the objective enabling us to study the shape of the focused spot. A neutral density filter was placed in front of the camera to prevent pixel saturation in the camera. The center of the bright spot was found and the pixel values for the column and row intersecting this spot-center pixel have been plotted in Figure 3.8. We see that the spot is symmetric with an approximately Gaussian profile that trails off to a constant minimum. Figure 3.9 shows the central peak of the laser profile plotted against a Gaussian curve having zero mean and $\sigma = 12.9$ pixels. For some trapping experiments, where motion in only one dimension is considered, ensuring the trap be equally strong in the x and y directions is unnecessary, however, symmetry about x *or* y is. Properly aligned optics are crucial to effective trapping and will be discussed in greater detail in the following sections.

3.3.2 The microscope

The second most important component in an optical tweezer apparatus is the microscope objective (the laser being the most important). High numerical aperture (NA) objectives are used to maximize trap stiffness. The NA of an objective is a measure of its resolving power and is related to the index of refraction of the medium the objective is operating in, n_{med}, as well as the maximum half-angle α of the cone of light that can enter or exit the objective. These two quantities are related by $\text{NA} = n_{med} sin(\alpha)$. To increase the NA of an objective they are designed to operate with a liquid of higher index of refraction between the objective and the cover slip, such as water or immersion oil. We can imagine the NA relating to ϵ in Eq.(3.4), a higher NA will produce a more localized trap, similarly we can

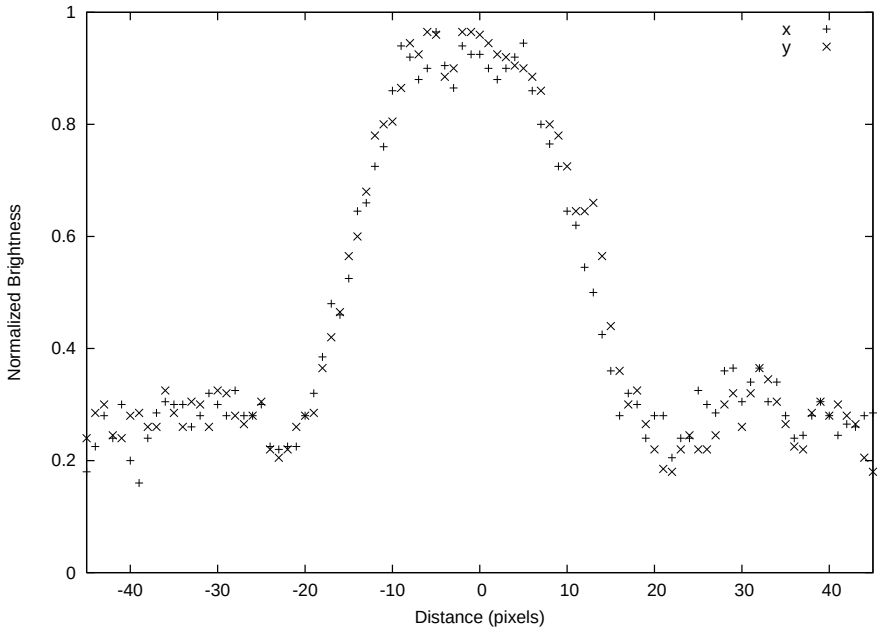

Figure 3.8. Cross-sections of the focal point of the laser in the x- and y-direction, the normalized brightness is obtained by dividing each pixel value by a threshold value of 200 (a white pixel has value 255).

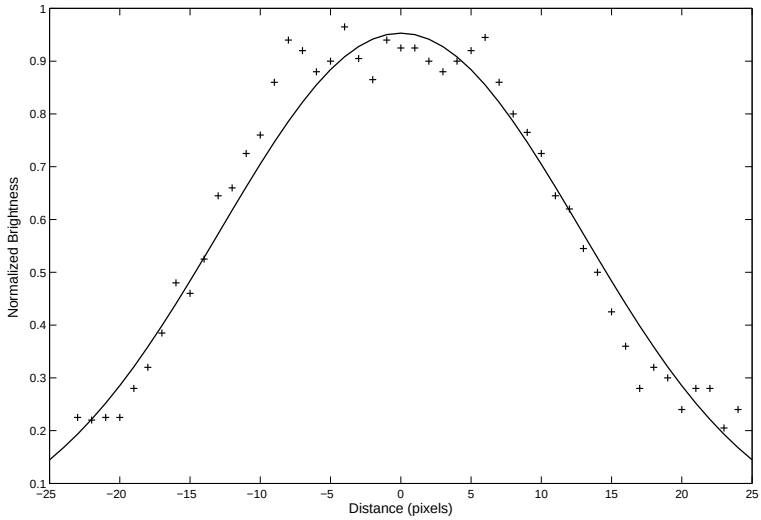

Figure 3.9. The central peak of the laser profile plotted against a Gaussian curve, $\sigma = 12.9$ pixels and the normalized brightness is obtained by dividing each pixel value by a threshold value of 200 (a white pixel has value 255).

imagine the rays from the edges of the field incident on the particle will have an even more pronounced refraction through the particle, thereby increasing the restoring force. Smith et al. obtained an expression for the trap width as a function of numerical aperture that justifies the use for a high NA objective [60]. Their expression, $w_{trap} \geq 1.22\lambda/n_{med}\sqrt{[(n_{med}/\mathrm{NA})^2 - 1]}$, where λ is the laser wavelength, gives an ideal minimum value for $w_{trap}(NA)$. Clearly, as the NA is increased is w_trap decreased creating a stronger, more localized trap.

In our apparatus we use a Nikon TE-2000U inverted microscope. Our objective is an infinity corrected, 100X immersion-oil objective with a NA of 1.4. The use of inverted microscopes in optical tweezer experiments is rather ubiquitous because of the ease with which they can be accessed with a laser. The laser is directed into the microscope through either a viewing or camera port beneath the objective and directed to the back-aperture of the objective with a dichroic mirror. However, there are many examples of working optical tweezers that do not rely upon a dedicated microscope body. Such systems incorporate the objective and dichroic mirror by mounting them in a rack system within the optical-train [12, 51, 60, 65]. There are two benefits to an arrangement that does not use a microscope body, the first being monetary, the second being safety. One less piece of equipment makes the apparatus less expensive. As regards safety, a microscope with eyepieces in place poses an inherent risk because with the laser directed into the microscope, highly focused laser light can be directed into the eyepieces causing irreparable damage if looked into with an unprotected eye. In our arrangement, the laser enters the microscope through the rear port, level with the optical table. The sample is viewed with a camera mounted to an auxiliary port on the microscope as a safety measure, this allows us to direct all light from the sample to the camera, and none to the eyepieces.

3.3.3 Latex beads

We have used Latex microspheres from Interfacial Dynamics Corporation for our trapping experiments. The spheres have diameter 970 nm and are classified as *surfactant-free NIST calibrated standard/sulfate white polystyrene latex* (batch number C-106.1). At $20°C$ the particles have the following properties, index of refraction for light of wavelength $\lambda = 590$ nm is $n_{590} = 1.591$, density $\rho = 1.055$ g/cm^3, and mass 4.033×10^{-12} g. Samples were prepared by making a solution of the spheres in distilled water and placing a droplet ~ 100 μl on a glass microscope slide. The samples were covered with a 0.17 mm cover slip and sealed with conventional nail-polish. Specific concentrations were never measured. Since trapping of only a single particle was desired, a minimal concentration was preferred. A low particle concentration minimized the likelihood of stray particles drifting into the trap.

3.3.4 The optics

We have used a minimum of optics in our apparatus to maximize efficiency. We used only enough components to ensure the laser was collimated and expanded to the fill the back aperture of the objective. We used two plano-convex lenses to form a telescope and anti-reflection mirrors to direct the laser into the rear microscope port. These components were purchased from a variety of suppliers, ThorLabs, Newport, and Oriel. Inside the microscope we used a dichroic mirror from ChromaTechnology that designed to be reflecting for the wavelength of our laser $\sim$ 530 nm. When collecting images of the laser we placed a neutral density filter before the camera to block 90% of the light reaching the camera. During trapping experiments we used a filter that prevented the laser light from reaching the camera, this facilitated observation of only the trapped particle and not the laser light diffracted by the particle.

3.4 Data collection and image processing

Images were captured with a camera connected directly to a computer. We have used an iBOT FireWire WebCam. The decision to use a web-cam was one of economy and versatility. The camera itself is inexpensive and does not require the use of after-market software or a frame-grabber for pre-processing. This camera is capable of collecting grayscale images at a rate of 30 frames per second (fps) at a resolution of 640x480 pixels, comparable to many of the available CCD cameras. Although the camera is capable of collecting color images, this was an unnecessary feature; in our experiments, only grayscale images were collected. For image collection we have used the freely available Coriander software, developed by Damien Douxchamps, which runs on most Linux operating systems [35]. This software has become much more stable in recent years and version 2.0.0-pre6 proved quite robust. We were able to collect images at 30 fps at full resolution for as long as necessary. We typically collected a minimum of 15,000 images. To verify that we were collecting images at 30 fps we collected images and stored them with a date-stamp accurate to 1 ms. We found that, on average, the time between collecting two successive images was constant and $\Delta i \simeq 33$ ms corresponding to a frame rate of 30 fps. To save computation time during image analysis only the relevant portion of the images were kept. Image cropping was performed on the entire batch of images using a `bash` shell script [61] and the program Convert. Convert is contained within the freely available ImageMagick software suite[42]. After the images were cropped to a square region containing only the trapped particle, tracking software was used to locate the particle centers. The coordinates of the particle centers were stored for further analysis.

Particle tracking has been carried out by two separate methods. The first used software that we have written in the C programming language, the second uses freely

distributed software for MATLAB. This software was ported from IDL by Blair and Dufresne [17], the IDL software was originally developed by Grier, Crocker, and Weeks [24]. This third-party software is more advanced than our software allowing for the tracking of multiple particles through a sequence of images. Our software only allows for a single particle to be tracked. We have used the MATLAB software for the image analysis of the experiments presented here. We chose to use the MATLAB software because, with our camera, we could only collect two image formats at $30fps$, .raw or .ppm. MATLAB processes .ppm's directly whereas our software requires .png formatted images, this meant that to use our software we converted all images into .png format before cropping. Yet another extra step in the image analysis procedure. When planning our experiments we anticipated the need to track multiple particles so, rather than implement that feature into our software, we chose to use the available MATLAB software. The MATLAB tracking protocol is outlined in this section and a comparison with our software can be found in Appendix A.

Particle tracking is accomplished by a suite of programs that filter and then find features of a user specified size, l_f. First the image containing the features to track is loaded into MATLAB as a two-dimensional array having the same dimensions as the image size in pixels, Figure 3.10. After the image has been loaded it is then smoothed and filtered with the program `bpass.m`, this convolves the image with a Gaussian spot of size l_f and subtracts off the background light. The net effect of this is that the image is now comprised of white Gaussian spots where the features were, with a completely black background. Figure 3.11 is `bpass.m`'s output for the image in Figure 3.10 with a particle size of 15 pixels. There is a black border around this image to ensure that only features that can be fully accounted for are accepted, that is, features cannot extend beyond the edge of the image. After the image has been filtered, a coarse feature locater is applied using the program `pkfnd.m`. This

locates features to within single pixel accuracy by locating the brightest pixels above a minimum value b_{min}, care is taken to avoid selecting two pixels within l_p of each other by only selecting the brightest within that region, the results of this are seen in Figures 3.11 and 3.12, in each of these the approximate center is highlighted at $(61, 59)$. Finally, sub-pixel accuracy is obtained with `cntrd.m`. This last program calculates the *center of mass* around each of the previously located features, this is taken as the true location of the feature. For the particle shown in Figure 3.10 `cntrd.m` locates the center at $(60.68, 58.77)$. As was stated earlier, this software is capable of tracking multiple particles through a sequence of images, but that was never needed in our work because we have only looked at the motion of single particles and have taken care to only have one particle in each image. To implement this software for a sequence of images we wrote a simple program for MATLAB using the aforementioned functions. We processed all images in a directory and output the x and y coordinates to a text file for future analysis.

3.5 General trapping procedure

Experiments performed using a static, non-fluctuating potential were carried out in the following manner.

1. The laser was turned on and set to the desired power at least one hour prior to data collection. This was done to ensure laser stability, which should have been within $\pm 5\%$ after one hour.

2. The symmetry of the laser profile was checked. This was accomplished by placing a cover slip over the objective with a mirror on top of that to observe the focused laser through the web-cam. Any necessary adjustments to the optics were made at this point.

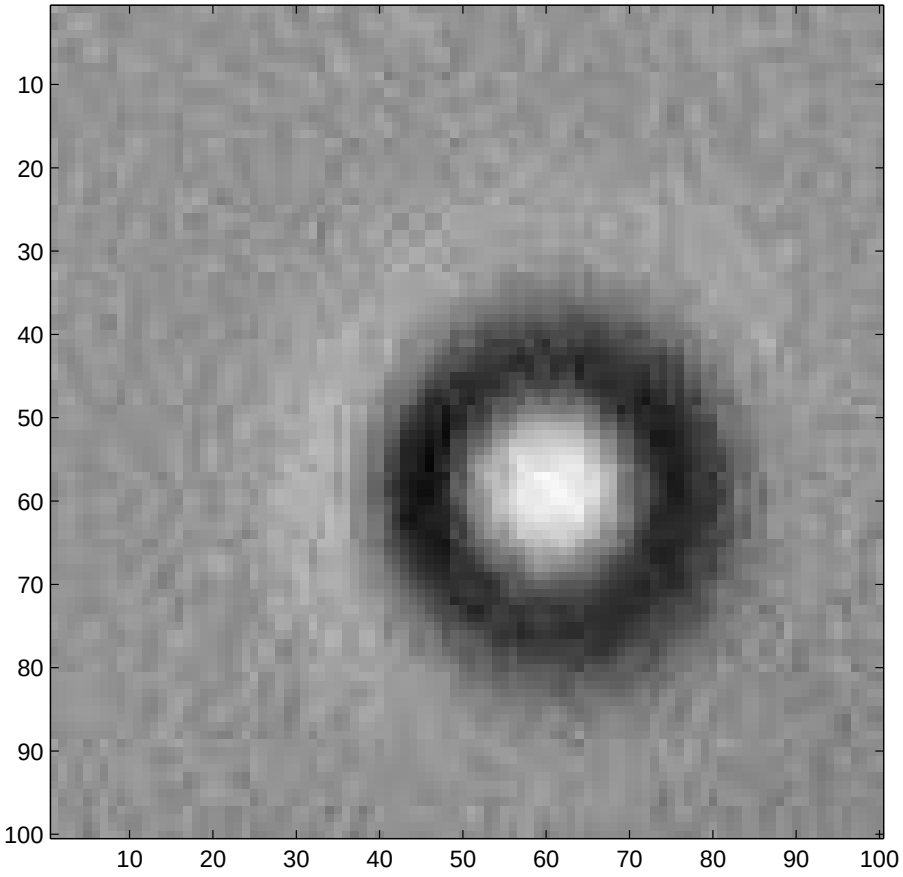

Figure 3.10. A typical cropped (100x100 pixel) image containing a single particle.

3. A sample was prepared by placing a small droplet of solution containing the latex beads on a glass slide and sealed with a cover slip and nail-polish. After a few minutes the nail-polish was dry.

4. The mirror and cover slip were removed from the microscope and replaced with the prepared sample. With the laser blocked from entering the microscope we used the eyepieces to focus the sample and locate a latex bead for trapping.

5. Once the sample was focused we opened the rear port to the microscope to allow the laser to focus on the sample. We also directed the light away from

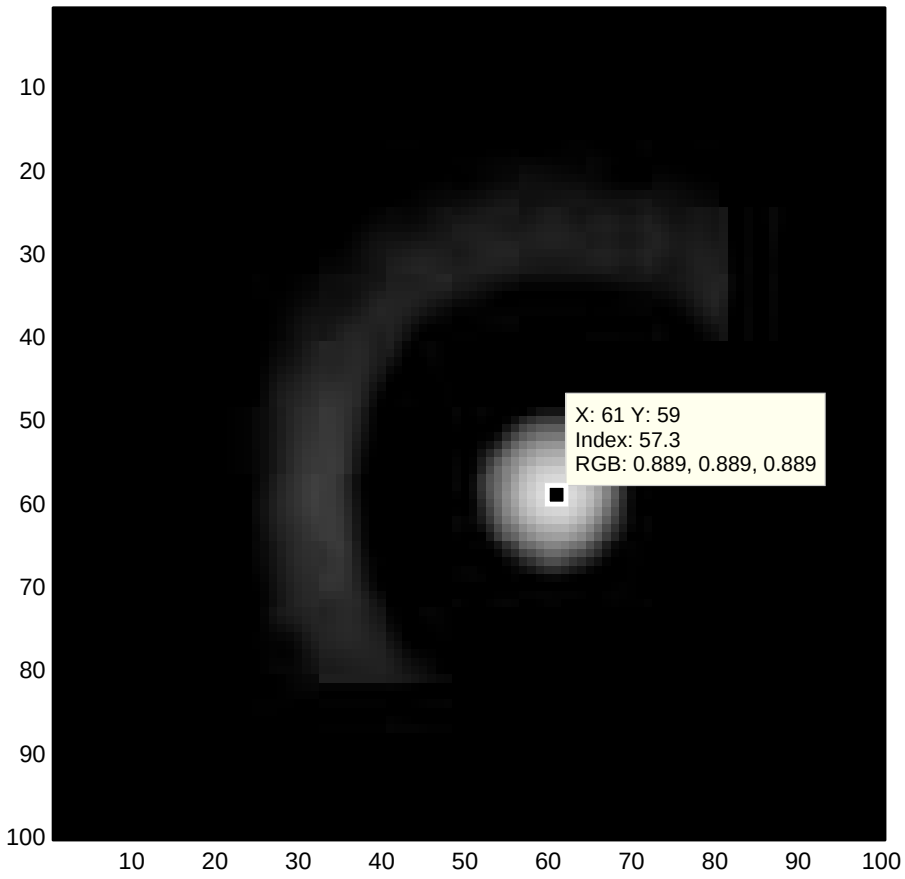

Figure 3.11. After `bpass.m` is run with a particle size of 15 pixels we get this (100x100 pixel) image.

the eyepieces and into the camera. To collect good images we made certain that the laser filter was in place in front of the camera.

6. With `Coriander` running on our computer we were able to observe the focused sample on the computer in real time. We kept the laser focus centered on the camera center so that we knew where to direct particles whenever it was necessary to "go fishing" for a particle to trap. At this point, if a particle could be observed in the field of view we moved the sample stage on the microscope and tried to direct the particle into the trap. If a particle could not be seen we either continued moving the stage (slowly)

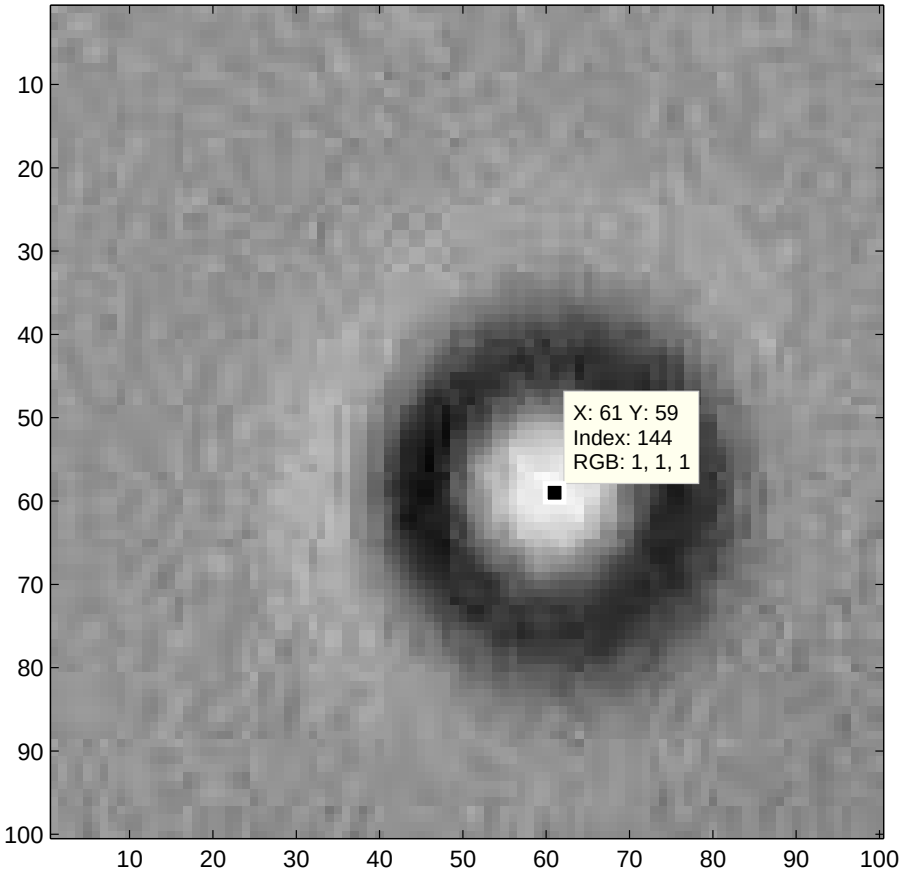

Figure 3.12. The original cropped (100x100 pixel) image again with the approximate particle center labeled at $(61, 59)$.

to find a particle or we blocked the laser and looked through the eyepieces while translating the stage until a particle to trap was located.

7. With a particle trapped, we specified where to store the images and began saving images. We typically collected a minimum of 15,000 images (at 30 fps this takes approximately 8.5 minutes). During this time it was important that another particle did not drift into the trap, something that was problematic with higher concentrations of latex beads. If a second particle did drift into the trap the experiment was stopped, the images discarded, the laser blocked to release the trapped particles, before trapping another.

3.6 Results and analysis

After the images had been processed and we had a text file containing the coordinates of the particle center for each image collected, we analyzed the motion of the trapped Brownian particle. We began by determining the trap stiffness k_{trap} and the probability distributions $P(x)$ and $P(y)$, these were separately used to calculate the potential energy. Next, we looked at conditional probabilities, i.e. $P(x_f, t + \Delta t|x_0, t)$, and LTFTTs. Finally, for verification, we performed computer simulations of a Brownian particle in an analogous potential well.

3.6.1 Trap characterization

In Section 2.2.2 we wrote the force due to the potential as $-kx$ in the Langevin equation (Eq.(2.5)), this choice was not arbitrary and is a valid approximation of the potential created by a Gaussian laser in an optical tweezer [12, 51, 60, 65]. By assuming the force arises from the gradient of a scalar potential we have $\mathbf{F} = -\nabla U$, and for $F_x = -kx$ the potential may be written as $U(x) = 1/2kx^2 + C$, k is the trap stiffness (spring constant) and C is an arbitrary constant. From the equipartition theorem we may relate the average potential that the particle experiences to the energy associated with a single microscopic degree of freedom $k_BT/2$, in terms of the trap stiffness and the mean square displacement of the particle this is

$$\frac{1}{2}k\langle x^2\rangle = \frac{1}{2}k_BT. \tag{3.9}$$

This equality means that we can obtain the trap stiffness from the absolute temperature T and $\langle x^2\rangle$, note that x is measured relative to the equilibrium point of the trap. A segment of raw data is shown in Figure 3.13 with the deviation of the particle from the average position, $\bar{x}$, measured in nanometers. To convert from pixels to physical units we obtain images of a particle once it has settled on the cover-slip and is immobile. We then use an image analysis program, i.e. TheGimp, to measure the

particle size in pixels before calculating the length of a pixel in nanometers, typical values are $1 pixel \sim 30 - 40$ nm.

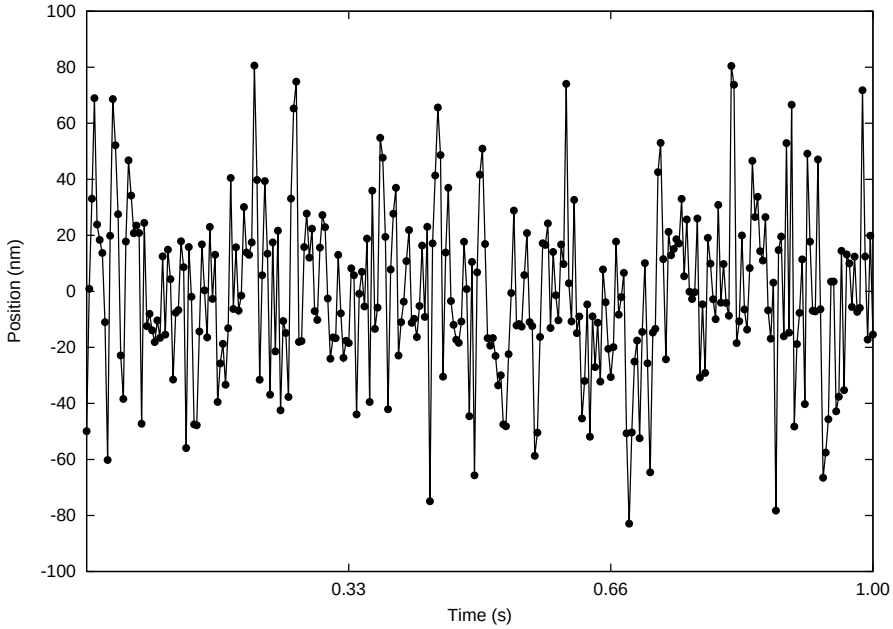

Figure 3.13. A sample of raw data. We have plotted the deviation of the particle from the average position (nm) versus the frame number, the time-span between each frame is 1/30 s.

For a system that is not dissipating energy we may write the probability of finding the particle at x, independent of time, as being proportional to the associated equilibrium Boltzmann factor,

$$P(x) = \frac{1}{Z} e^{-U(x)/k_B T}, \tag{3.10}$$

where the normalization condition $\int P(x)dx = 1$ allows us to write $Z = \sqrt{2\pi k_B T/k}$. By taking the logarithm of $P(x)$ we have an expression for the potential $U(x) = -k_B T \log[ZP(x)]$.

We have written our analysis software using the C programming language, the first program we use is `stats.c` and it calculates the trap stiffness, the probability distribution, and the potential. Obtaining the trap stiffness is straightforward in

practice. From our raw data, which is a list of N pixel coordinates with values x_p ($1 \leq p \leq N$), we begin by locating the mean position of the particle, $\bar{x} = N^{-1} \sum x_p$, and next calculate the mean square displacement using $\langle x^2 \rangle = N^{-1} \sum (x_p - \bar{x})^2$. Running `stats.c` with the complete data-set, partially shown in Figure 3.13, we find $\langle x^2 \rangle = 0.854$ pixel^2 $= 784$ nm^2 and use $\langle x^2 \rangle$ to calculate the trap stiffness using Eq.(3.9). We find $k \approx 5 \times 10^{-3}$ pN/nm $= 5 \times 10^{-6}$ N/m. With this value for k we can obtain a theoretical probability distribution by using Eq.(3.10). An experimental probability distribution is obtained by parsing the raw data into bins of a predetermined size, bin_w, and then counting the number of times we find the particle in each bin and dividing each of these by the number of points in the set. With N data points we find the particle in the jth bin n_j times ($\sum n_j = N$) and the probability that the particle is in the jth bin to be $P_j = n_j/N$. Using the same data from Figure 3.13 the P_j's were obtained and in Figure 3.14 have been plotted against the theoretical $P(x)$ using the value of k obtained above. This figure shows good agreement between the theoretical curve of Eq.(3.14) and the experimentally obtained distribution. How well the fit *looks* is dependent upon bin_w. As bin_w is made smaller the resolution becomes greater but the n_j's decrease as the N points are distributed over more bins, consequently the resulting distribution will appear less Gaussian. This is a purely statistical effect, there is a limit to how finely a given set of data can be resolved. In the following sections bin_w will vary and this will always be noted. In Section 4 we will discuss in greater detail the significance of the choice of bin_w. We next consider the conditional probabilities associated with spatial transitions.

3.6.2 Conditional probabilities

In Eq.(2.4) of Section 2.2.2 we equated the ratio of two spatially inverted conditional probabilities to $e^{-\Delta U/k_B T}$ for a non-dissipative system. This equality asserts

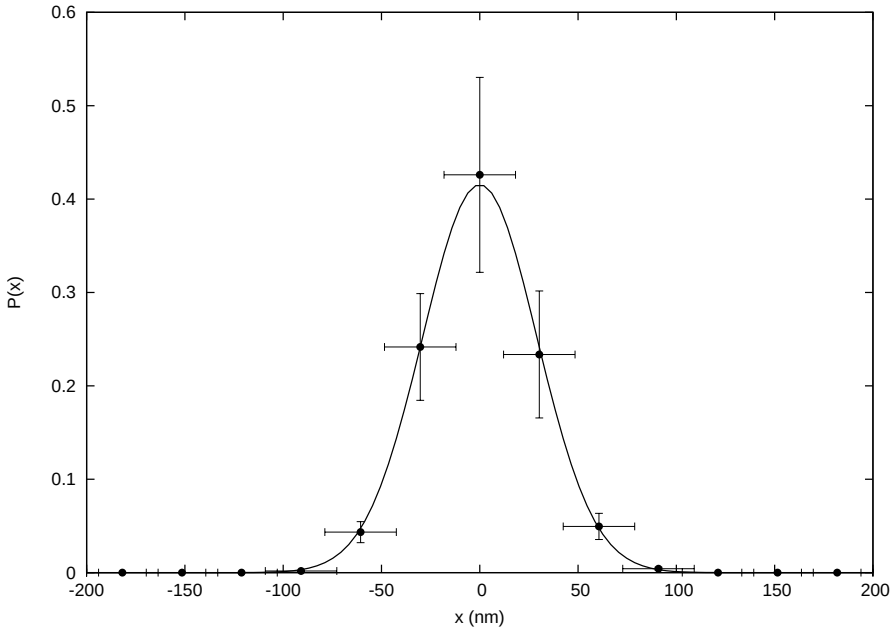

Figure 3.14. Probability distribution for the raw data of Figure 3.13. The data has been parsed into bins 1 pixel ($\sim$ 30 nm) in width and the curve is obtained from Eq.(3.10).

that we can know the shape of a potential from measurements of conditional probabilities. That is, by observation of fluctuations we can know something of the underlying potential (energy landscape) that is influencing the behavior of the system. From the position data of a Brownian particle we can investigate specific transitions. To do this we have created the program `conditionals.c` and an accompanying set of functions in `objects.h`, the code fore these can be found in Appendix B. To calculate the conditional probability given by $P(x_f, i + \Delta i|x_0, i)$, where i is the "frame number" and Δi is the time interval during which the transition takes place, we look at the $(i + \Delta i)$th point to check that the condition is satisfied. We use data that has been parsed into bins and, starting at the first data point, loop through the entire set. If $x_i = x_0$, we add 1 to a counter: $c_0 \rightarrow c_0 + 1$, and if $x_{i+\Delta i} = x_f$, we add 1 to a second counter, $c_{f,0}$. After looping through the entire data set we have the conditional probability, $P(x_f, i + \Delta i|x_0, i) = c_{f,0}/c_0$. Reversing x_f and x_0

and repeating the process gives us $P(x_0, i + \Delta i | x_f, i) = c'_{f,0}/c'_f$ where the primed counters refer to the reverse transition. We calculate the ratio of these conditional probabilities for a variety of x_f's, x_0's, and a range of Δi's ($2 \leq \Delta i \leq 10$) (only bins containing at least 5% of the total number of points, i.e. $n_j \geq 0.05N$, are considered acceptable transition endpoints). We then compare these ratios with the decaying exponential mentioned earlier, when doing this we can use either $U(x) = 1/2kx^2$ or $U(x) = -k_B T \log(ZP(x))$, the former being obtained from the mean square displacement of the particle and the latter coming from the number of occurrences in a particular bin (the n_j/N's from Section 3.6.1).

The program `conditionals.c` outputs a data file named `cond_probs_#_xdat.c` (for a raw data file named `xdat`, and the `#` represents the maximum value Δi_{max} considered) containing a list of quantities relevant to the conditional probabilities we have calculated. Figure 3.15 is a plot comparing the theoretical expression for the ratio of two conditional probabilities with experimental data using two different methods for obtaining $U(x)$. To make this plot we have used the same data shown in Figures 3.13 and 3.14, we have kept $bin_w = 1$ pixel and have averaged the ratios of conditional probabilities for the values of Δi in the range $2 \leq \Delta i \leq 10$. For this choice of bin_w each method of obtaining the potential gives good agreement with the theoretical curve $e^{-\Delta U/k_B T}$, however, not all choices of bin_w lead to such good agreement. This subtlety as well as the significance of our choice of Δi will be discussed in greater detail in Section 4.

3.6.3 Transition times

By looking at spatial transitions we can obtain mean first passage times (MFPTs), the mean *wiggling* times (MWTs), and the direct transition times, either LTFTTs or LOFOTs (defined in Section 1.3. We describe here the general method used to extract transition times. The term LOFOT is used to describe all direct transitions.

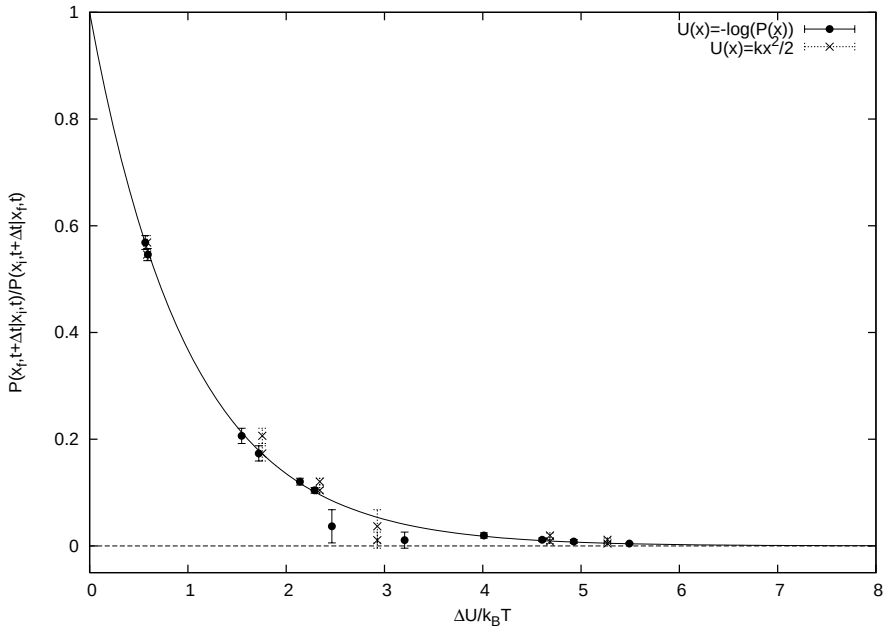

Figure 3.15. A comparison of experimental results for the ratio of conditional probabilities where we have used two methods to obtain the potential. In this plot the data has been parsed into bins having width 1 pixel and we have averaged the ratios of conditional probabilities for $2 \leq \Delta i \leq 10$, that is, the transition takes place over $2 - 10$ frames.

Note that the LTFTT represents the limiting case where the observation interval is sufficiently small, i.e. LOFOT $\rightarrow$ LTFTT as $\Delta t \rightarrow m/\gamma$, the velocity relaxation time. The program we have written to extract each of these transition times from a data set is `passage_times.c`, like `conditionals.c`, this also uses functions contained in `objects.h`. Again, the data is parsed into bins before the MFPTs, LOFOTs, and MWTs are obtained. We have used a method similar to that used to obtain conditional probabilities in that we loop through the entire data set looking for the number of times a particular set of conditions are met. In the case of transition times we tabulate the duration of each transition rather than specifying a time interval for the transition *a priori*. Because MFPT = MWT + LTFTT, we need only determine two of these with the third being derived from this expression for the MFPT. We have explicitly obtained the MFPTs and LOFOTs and take their

difference to obtain MWTs. To extract the MFPT for a transition from x_i to x_f, we loop through the data and when $x_p = x_i$ we start a "timer" (a counter that we add to), and we let this run, for l steps, as soon as $x_{p+l} \geq x_f$ (for $x_i < x_f$) we record l, the time the particle takes to make the transition, for each occurrence in the data set. The average of the l's gives us the MFPT. Although obtaining the LOFOT is more cumbersome, we are able to quantify both it and the MFPT simultaneously. To measure the LOFOTs we use a second "timer" that restarts whenever $x_{p+j} \leq x_i$ (again, $x_i < x_f$), where x_{p+j} is an intermediate step between the first touching of x_i and the first touching of x_f, $1 \leq j \leq l$. If $p + j_{max}$ is the last time the particle touches x_i, $x_{p+j_{max}} = x_i$, and the LOFOT is expressed as $l - j_{max}$. Finally, we average the $(l - j_{max})$'s to obtain the average LOFOT. We calculate the forward and reverse MFPTs and LOFOTs for a given pair of endpoints in a single pass through the data to be more computationally economical. A different set of rules is used depending upon whether or not x_p is equal to x_i or x_f. After we have obtained the LOFOTs, we compare the forward and reverse times by plotting them against the line $y = x$ for reference.

We were unable to measure actual LTFTTs because the timescales that these occurred on in our trapping experiments were much shorter than we could resolve with our camera. A few quick calculations elucidate this. First, the mean displacement of a freely diffusing, micron sized Brownian particle in aqueous solution is described by the relation, $\langle x \rangle = \sqrt{2Dt}$. The diffusion constant, D, comes from the Einstein relation, Eq.(2.2), and is $D = k_B T/\gamma \approx 5 \times 10^5 \text{nm}^2/\text{s}$. For the time interval relevant to our experimental setup, the time between two successive frames collected with our camera, 1/30 s, we find $\langle x \rangle \approx 180$ nm. This distance is similar to the particle's maximum displacement from its equilibrium position in the trap. From Eq.(2.51) we find LTFTT = 1.3 ms for a particle in a trap of stiffness $k = 5 \times 10^{-6}$ N/m starting at the origin, $x_i = 0$, and moving to the position $x_f = 60$ nm, this is nearly 25 times

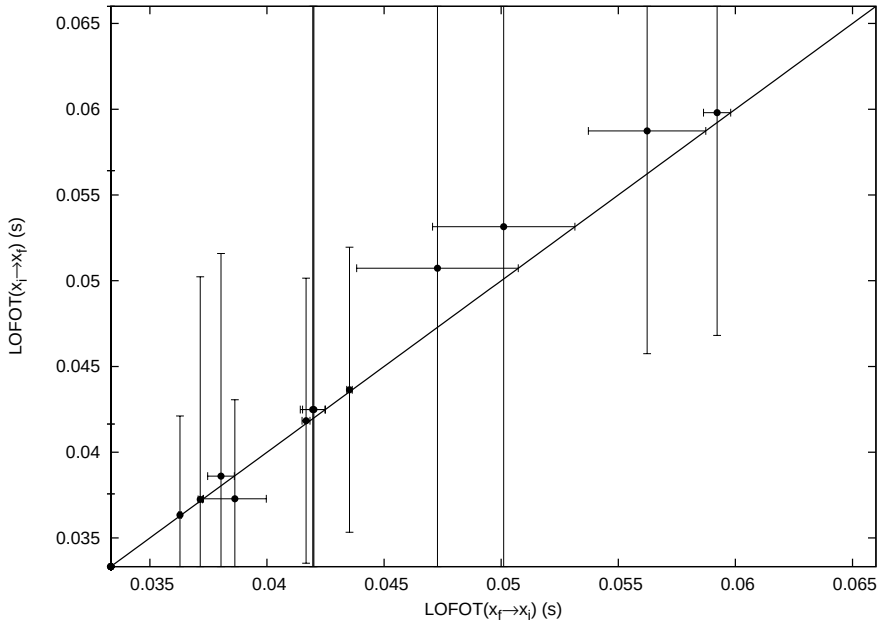

Figure 3.16. A comparison of LOFOTs from the same experiment as Figure 3.15 plotted against the line $y(x) = x$. For this experiment $k \approx 5 \times 10^{-6}$ N/m, a relatively strong trap, which means the particle is confined to a smaller region and has a correspondingly narrow probability distribution. We had to use $bin_w = 0.4$ pixel to get even the few transitions shown here.

shorter than the 33 ms we can experimentally resolve. Interestingly, regardless of the time interval between measurements, we find that irrespective of direction, the observed transitions take equal times. While we cannot claim experimental measurements of LTFTTs, this feature is equality between LOFOTs in general is worth noting.

Figure 3.16 shows the measured LOFOTs for the data shown in Figures 3.13 and 3.15. We have set $bin_w = 0.4$ pixels in these figures because this is a relatively strong trap ($\langle x^2 \rangle$ is small). Narrower bins were used so that there would be more bins. Smaller bins were also chosen to minimize the likelihood that transition times were not, on average, spanning only two successive data points. That said, bins were kept large enough that sufficiently large bin populations (large n_j's) were maintained—to serve as statistically significant transition endpoints. For the

data shown in Figure 3.16, we found the average deviation of the ratio of LOFOTs from unity to be 0.02. Similarly, the average of the ratio of transition times was $\langle \mathrm{LOFOT}_{\downarrow}/\mathrm{LOFOT}^{\uparrow} \rangle = 1.0005$, the up and down arrows represent up- and down-well transitions, respectively. Only ratios were taken of LOFOTs that were spatial reverses of one another. Looking at Figure 3.16, there is an undeniable correlation between the measured LOFOTs and the data collection rate. This will be discussed in detail in the following section where we describe simulations of our optical trapping experiments. In these simulations we specify all parameters, including the sampling rate, to further investigate LTFTTs.

3.6.4 Monte-Carlo simulations

The system we have studied can be accurately represented by a Monte Carlo simulation. Monte Carlo simulations follow a set of rules, the outcome of at least one rule being determined at random. By simulating a Brownian particle in a parabolic potential well, under the same conditions as our experiments, we were able to test our analysis methods. The original motivation for doing this was to check whether or not we were accurately obtaining values for trap spring constants; if we specify in a simulation that the potential have the form $U(x) = 1/2k_x x^2$, we should be able to obtain that same potential by correctly analyzing the simulated position data. We were, in fact, able to conclude that $k_{data} = k_x$, where k_{data} is the value we obtained from our analysis of the simulated data. What was originally intended to be a simple check to ensure our analysis was working proved to be invaluable because in a simulation we could arbitrarily specify any parameters of interest. In order to simulate Brownian motion in a potential well we have used a method described by Gillespie [39], outlined below.

Gillespie begins by writing the Langevin equation, in the over-damped limit, with the random noise term $f(t)$ we have

$$\gamma v = -\frac{dU(x)}{dt} + \sqrt{2k_B T\gamma} f(t). \tag{3.11}$$

Eq.(3.11) is, of course, Eq.(2.5), excluding the inertial term. To put this into a self-consistent form that can be simulated, Gillespie begins by writing the finite difference,

$$x(t+dt) - x(t) = \Xi(dt; x, t). \tag{3.12}$$

We require that Ξ be independent of all times prior to t, a smooth function of its variables, and that $\Xi(dt \to 0; x, t) \to 0$. A direct consequence of the last requirement is that if we separate dt into n sub-intervals,

$$\Xi(dt; x, t) = \sum_{i=1}^{n} \Xi(dt/n; x, t). \tag{3.13}$$

We use the central limit theorem to argue that the sum of these n statistically independent random variables is *normal*. Because of this normality we write the mean and variance of $\Xi(dt; x, t)$ as

$$\begin{aligned} \langle \Xi(dt; x, t) \rangle &= \sum_{i=1}^{n} \langle \Xi(dt/n; x, t) \rangle \\ &= n \langle \Xi(dt/n; x, t) \rangle, \end{aligned} \tag{3.14}$$

and,

$$\begin{aligned} \Delta\Xi(dt; x, t) &= \sum_{i=1}^{n} \Xi(dt/n; x, t) \\ &= n \Delta\Xi(dt/n; x, t). \end{aligned} \tag{3.15}$$

Furthermore, we can show that a function $y(x)$ that is smooth and obeys $y(x) = ny(x/n)$, for all integers n, must have the form $y(x) = Ax$ ($A \neq A(x)$). Noting that when a derivative with respect to x is taken, $y'(x) = y'(x/n)$, and that for large n we have $y'(x) = y'(x/\infty) = y'(0)$, it follows that $y(x)$ must have the form Ax. This

is further elaborated by Gillespie in his book[38]. In light of this, Eqs.(3.14) and (3.15) must have the forms

$$\langle \Xi(dt; x, t) \rangle = A(x, t)dt, \tag{3.16}$$

$$\Delta\Xi(dt; x, t) = D(x, t)dt, \tag{3.17}$$

respectively. Since $\Xi(dt; x, t)$ is a normal random variable with the properties that $\langle \mathbf{N}(\alpha, \beta^2) \rangle = \alpha$ and $\Delta\mathbf{N}(\alpha, \beta^2) = \beta$, we see that $\Xi(dt; x, t) = \mathbf{N}(A(x, t)dt, D(x, t)dt)$ may also be written as

$$\Xi(dt; x, t) = A(x, t)dt + D^{1/2}(x, t)dt^{1/2}\mathbf{N}(0, 1). \tag{3.18}$$

Comparing Eqs.(3.11) and (3.18) we see that $A(x, t) = U'(x, t)$, $D = 2k_BT\gamma$, and that $f(t)$ must be $(dt)^{-1/2}\mathbf{N}(t)$. Next, Eq.(3.11) is written with $f(t)$ replaced by $1/(dt)^{1/2}N(t)$, the normal distribution with zero mean and a standard deviation of unity. We now have the update equation,

$$x(t + dt) = x(t) - \frac{1}{\gamma}U'(x)dt + \sqrt{2k_BT\gamma}(dt)^{\frac{1}{2}}N(t). \tag{3.19}$$

By specifying an initial value for $x(0)$ we are left the task of iterating through this equation any number of times to obtain $x(t)$.

To simulate our experiments, the potential is chosen to have the form $U(x) = 1/2kx^2$ to represent the quadratically approximated potential well of an optical tweezer. The derivative of the potential enters Eq.(3.19) as $-kx$, a Hooke's law type force. To compare with the experimental results discussed in Section 3, we set the trap stiffness to $k = 0.005 \times 10^{-6} N/m$. All other constants have been fixed and are listed in Table 3.3.

The choice of the sampling time interval, Δt_s, is the most critical to being able to resolve the subtle motions of a Brownian particle. There are two relevant timescales, the characteristic (velocity) relaxation time of the particle due to viscous

Parameter	Value		Description
T	298	K	Absolute temperature
k_B	1.38×10^{-23}	Nm/K	Boltzmann Constant
	1.38×10^{-2}	pNnm/K	
D	4.90×10^{5}	nm^2/s	Diffusion constant (k_BT/γ)
η	8.9×10^{-3}	poise	Water viscosity
	0.89×10^{-9}	pNs/nm^2	
r	0.50	μm	Particle radius
γ	8.4×10^{-6}	pNs/nm	Coefficient of viscous drag ($6\pi\eta r$)

Table 3.3. The parameters and their values used in our simulations. These were all held constant, only the trap stiffness k and the sampling time Δt_s were varied.

forces $\tau_c h = m/\gamma \approx 6 \times 10^{-11}$ s, and the relaxation time for a trapped particle $\tau_{trap} = \gamma/k \approx 10^{-3} - 10^{-4}$ s. Another less relevant timescale is the period of undamped motion in the trap $T_{trap} = \sqrt{m/k} \approx 10^{-4} - 10^{-5}$ s. It is sensible to make Δt_s very small, and we have used both 10^{-5} s and 10^{-6} s with equal success. Using Einstein's result that $\langle x^2 \rangle = 2Dt$, Eq.(2.1), we can estimate the particle's displacement as a function of elapsed time. Results for various times are listed in Table 3.4. We can also estimate the mean displacement from using the equipartition

$\langle x \rangle$ (nm)	t (s)
180	30^{-1}
100	10^{-2}
30	10^{-3}
10	10^{-4}
3	10^{-5}
1	10^{-6}

Table 3.4. Values for the mean displacement as a function of time for a freely diffusing Brownian particle.

theorem, Eq.(3.9). If $k = 0.005$ pN/nm, we expect $\langle x^2 \rangle = 823$ nm^2, which roughly translates to a mean displacement of approximately 29 nm. Comparing this distance with the distances in Table 3.4, it is clear that in order to accurately resolve the particle's motion we must keep $\Delta t_s \leq 10^{-3}$ s.

Our code for the simulation, `brownies.c`, can be found in Appendix B, it outputs two `.txt` files, one is the position data for each time step, the other is the position data that would be obtained from a data acquisition rate of 30 Hz, the speed of our camera. From this data we can produce identical plots to those shown in the previous sections as evidenced in Figures 3.17, 3.18, and 3.19. Clearly, these plots are similar to those obtained from experiments, Figures 3.13, 3.15, and 3.16.

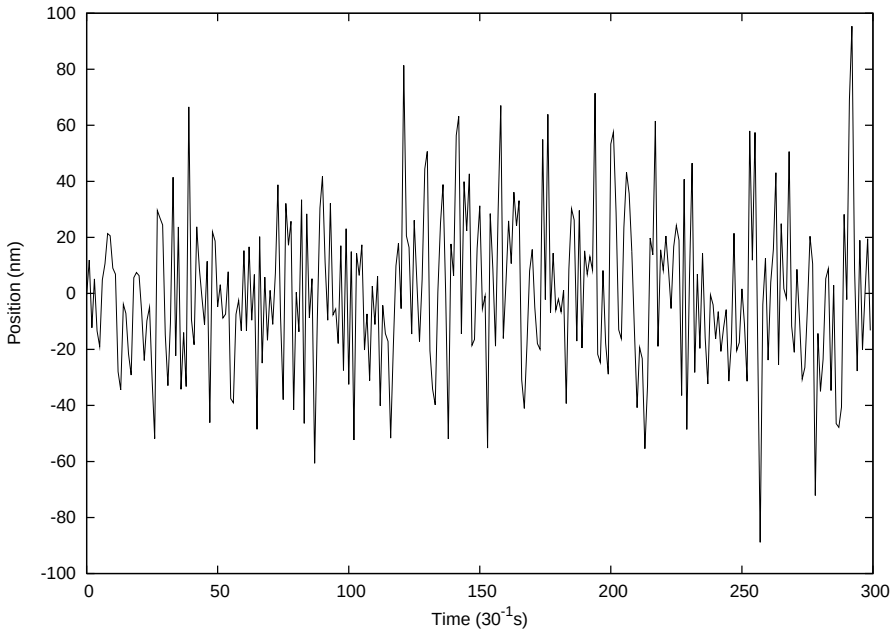

Figure 3.17. Raw data from a simulation with $k = 0.005 \times 10^{-6}$ N/m, sampled at 30 fps.

A quantitative comparison between our experimental results and those obtained from the simulations is neither insightful nor useful because of the coarseness of the experimental sampling in both time and space. However, a deeper understanding of the fundamental processes can be gained by comparing the quantities associated with spatial transitions as the timescales are varied. With $\Delta t_s = 10^{-5}$ s, we can resolve the LTFTTs and recover the theoretical predictions of Eq.(2.50) [29]. Looking at Table 3.4 it is certainly impossible to study the motions of a Brownian particle

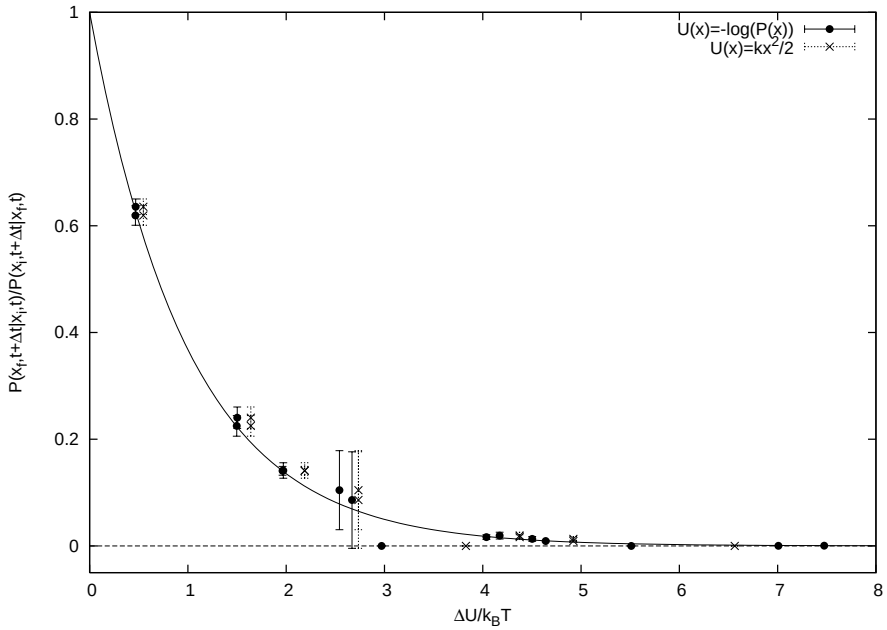

Figure 3.18. A plot of the ratio of conditional probabilities obtained from a simulation with parameters the same as Figure 3.15. The simulated data has been sampled at a rate of 30 fps, $k = 0.005 \times 10^{-6}$ N/m, $\Delta t_{max} = 10 frames = 0.33\bar{3}$ s, and the data has been parsed into bins with $bin_w = 18$ nm.

when observations are made at 30 Hz, $\langle x \rangle = 180$ nm at that time interval! Only with the transition endpoints separated by a distance many times greater than the predicted $\langle x \rangle$ for a given observation time interval can LTFTTs accurately be resolved. The vast amount of information that is lost with such a slow sampling rate, 30 fps, is quite apparent when we look at the movements of the particle between two successive samplings. If we simulate the system with $\Delta t_s = 10^{-5}$ s there will be 3333 simulation steps during any 0.033 s interval (the time interval for a 30 fps collection rate). Figure 3.20 shows an example 3333 "missed" steps, it is quite apparent that sampling rates on the order of 10 to 100 KHz are necessary here.

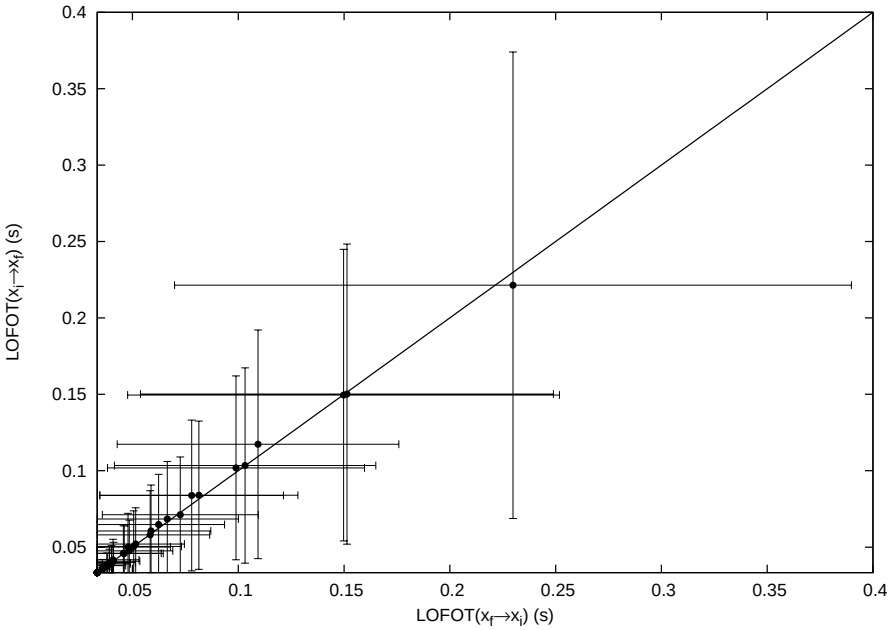

Figure 3.19. Simulated LOFOTs where data is sampled at 30 fps with each Brownian step taking $\Delta t_s = 10^{-5}$ s and $k = 0.005 \times 10^{-6}$ N/m.

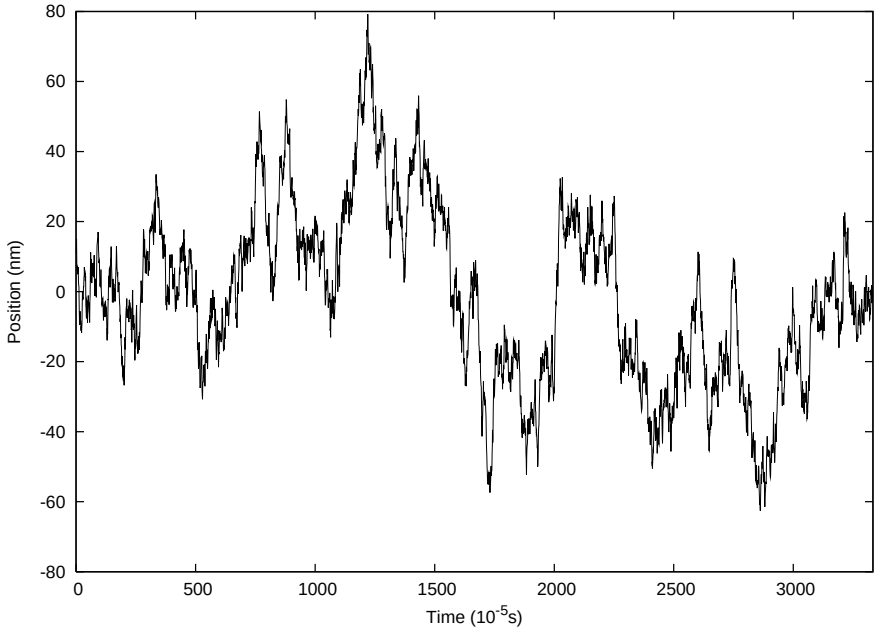

Figure 3.20. The 3333 data points obtained between the first and second 30 Hz samplings when $\Delta t_s = 10^{-5}$ s.

Chapter 4
DISCUSSION

In most optical trapping experiments the desired quantities are the particle's displacement from the equilibrium position as a function of time, the mean square displacement from the equilibrium position, and the force exerted on the particle by the trap. The optical tweezer is treated as a mass attached to an ideal spring (because the force is assumed to be linear) undergoing random fluctuations about the equilibrium position. If another object is attached to the trapped particle, i.e. DNA, kinesin, polymerase, and a force is exerted on the attached object by another means, the position of the particle in the trap will tell something of how hard the particle is being pulled. This is remarkable, it has allowed scientists to measure the forces exerted by tiny particles which are orders of magnitude smaller than the trapped particles, this is evidenced by the extensive work already done. This review article [6] by Arthur Ashkin does a wonderful job of discussing a vast array of work done with optical traps.

The experiments mentioned above represent amazing applications of the optical tweezer. In the preceding chapters we have shown that by studying the fluctuations of a particle trapped in a simple optical tweezer, fundamental quantities like the potential energy can be extracted. This chapter contains a discussion of our methods and the results we have obtained.

4.1 Bin sizing

The size of the bins the data is sorted into can affect the apparent results of a given experiment. As the bin size, bin_w, decreases the particle will be found less

frequently in any given bin. This affects the probability of finding the particle in a particular location. In order to minimize artifacts of bin sizing we have collected sufficiently large data sets so that when $bin_w \sim 1$ pixel (1 pixel $\sim 30 - 40$ nm), the bins will be sufficiently populated. Also, the analysis of the conditional probabilities and LOFOTs has been carried out on bins (transition endpoints) populated with at least 3% of the data points. This 3% threshold might seem small, but for a data set of 30×10^3 points the analysis is carried out on bins containing no fewer than 900 data points. As the data sets become larger the issue of low bin populations becomes less significant. For example, in a simulation with $N = 30 \times 10^6$ data points, 3% of N is 900×10^3, certainly a statistically significant number. As the bin size is increased, there is a decrease in the number of occupied bins. Consequently, there are fewer relevant data points in any analysis of the binned data. When we compare two plots of the probability distribution where different bin sizes are used, we see a marked difference in the *appearance* of the probability distribution—a variation in the height of the probability distribution. Although obtaining the probability distribution requires binning the data, obtaining the potential does not. It is not necessary to bin the data to obtain $\langle x^2 \rangle$ and the trap stiffness. Furthermore, we can compare the probability distribution from the binned data with the curve predicted from the mean square displacement (trap stiffness). Figure 4.1 shows the scaled probability distributions for four bin widths. We see that by scaling the binned data, the probability distribution is the same, as it must be. Each plot of binned data is compared to the same Gaussian fit of the non-binned data having two fit parameters, the mean and standard deviation, $\mu = -3 \times 10^{-13} \pm 0.032$ pixels and $\sigma = 3.78 \pm 0.02$ pixels^2, respectively. It is necessary to scale the binned probability distributions because as the bin width increases, the probability of finding the particle in a given bin also increases. This is because there is a greater interval that is accounted for in each bin.

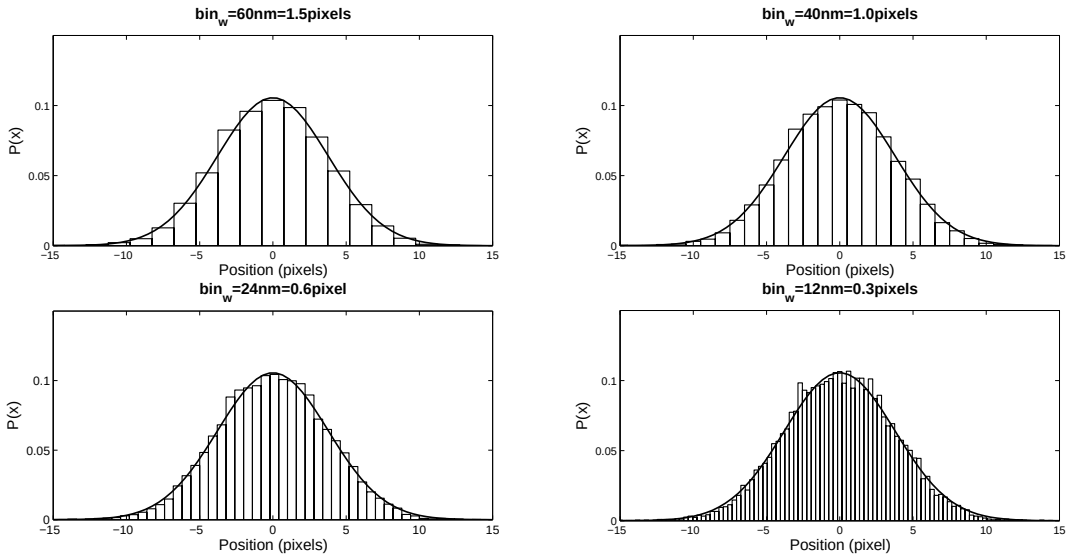

Figure 4.1. The binned experimental data after scaling. The same Gaussian curve is plotted against each binned distribution.

Surprisingly, the binning becomes less significant when we consider conditional probabilities, i.e. $P(x_f, t+\Delta_t|x_i, t)$. When we compare the ratio of two conditional probabilities with the expected $e^{-\Delta U/k_BT}$, there is a weaker agreement when we use $U(x) = 1/2k_x x^2$ as opposed to $U_j = -k_BT \log P_j$. This occurs for the same reason as the discrepancy between a given P_j and the corresponding $P(x_j)$ (x_j represents the location of the center of the j'th bin) that arises from the binning of the data. It is more pronounced with larger bins since $P_j \neq P(x_j)$, clearly we must write $P_j = \int_l^r P(x)dx$ where l and r are the boundaries of the j'th bin, as the distance $r - l$ increases for wider bins, so too will $P_j - P(x_j)$. Even though varying the bin width can have a pronounced, albeit superficial, effect upon the *apparent* agreement with the theoretical expression for the potential, it is of little concern. In fact, it is more sensible to use smaller bins to minimize the discrepancy between the P_j's and the $P(x_j)$'s, thus giving a more accurate value for ΔU when comparing the ratio of conditional probabilities with the corresponding ratio of Boltzmann factors.

There is little to say about data binning in regard to investigations into LOFOTs because we only compare one LOFOT with it's spatial inverse. If we are not comparing the LOFOTs with any other quantities, the effects of bin sizing are essentially negligible. The notable effects that increasing the bin width has on the LOFOTs are that the number of observable transitions decreases while the number of times a transition is observed increases. In spite of this, the ratio of two LOFOTs remains unity. However, when we compare simulated LOFOTs with theoretical predictions we find that any binning of the data leads to discrepancies. We can imagine two extreme scenarios. First, if the particle is last observed at the lower limit of the initial bin and first observed in the final bin at its upper edge or beyond, this would tend to give a longer than expected LOFOT. The second case corresponds to a transition from the upper edge of the initial bin to the lower edge of the final bin, on average contributing a smaller than expected transition time. Of course, these effects are mitigated by the large number of occupancies in a particular bin. That said, because in general we expect the lower potential side of the bins to be more frequently populated than the higher potential side, we expect some potential deleterious effects. In order to draw more accurate comparisons with the results of Eq.(2.51) we simulated specific transitions, this will be discussed in Section 4.3.

4.2 Conditional probabilities

The general condition for detailed balance is

$$P_{eq}(x_f)P(x_i, t+\Delta t|x_f, t) = P(x_i)_{eq}P(x_f, t+\Delta t|x_i, t). \tag{4.1}$$

This of course may be rearranged to yield Eq.(2.4) with its r.h.s. independent of Δt. For systems undergoing random fluctuations without time varying forces, the potential is clearly obtainable from the conditional probabilities. From our experiments with optical traps we provide experimental evidence that for times

much longer than the average time to transition from x_i to x_f, this relation is true. Conversely, from our simulations we have shown that for very short time intervals, $\Delta t <$ LTFTT, the equality also holds. This latter result is owed to long simulations and the fact that a sufficiently large data set is required to observe a given transition taking place over short times intervals. Nevertheless, the equality is verified for all relevant values of Δt. It is easy to see that for all time intervals the equality must hold, this is because the system is unconditionally described by Boltzmann statistics. It might seem that for very short time scales the equality could break down, that is, if $\Delta t <$ LTFTT it might be impossible for the particle to traverse the interval in question during such a short time interval. However, while improbable, it is not impossible. Furthermore, the LTFTT represents the *average* time for the direct transition; the difference between improbable and impossible cannot be overstated. Similarly, the conditional probability that a transition take place over a *large* potential difference, ΔU, over a *short* Δt, will certainly be small for the uphill transition, but it will be larger for the downhill transition, and these are related by the multiplicative factor $e^{-\Delta U/k_B T}$. It is also important to note that for those conditional probabilities that are smallest, there is the chance that even in a large data set, a specific transition will not take place at all. Unobserved transitions have always been disregarded by us. In Figure 4.2 we have varied Δt to test the robustness of Eq.(2.4) (or Eq.(4.1)). These plots differ from previous plots of conditional probability ratios in that we have used only a single Δt instead of averaging the ratios for a full range of Δt's, $1 \leq \Delta t \leq \Delta t_{max}$. In each of these figures the condition derived from detailed balance is satisfied, as well as directly from Onsager and Machlup by Bier et al. We expect to see a more broad range of ΔU represented as Δt is increased and this should come as no surprise, the probability of observing the particle undergo an energetic fluctuation of $5k_B T$ during a time interval of 10^{-5} s is incredibly small compared to the case where the time interval is

33 ms. This can be verified by a quick calculation using Eq.(2.6). However, the ratio of an uphill conditional probability and its downhill counterpart remains invariant with respect to the observation time interval. It is quite apparent from Figure 4.2 that our assumption that averaging the ratio of conditional probabilities for multiple values of Δt is both reasonable and sensible; if the result is valid for any single value of Δt it makes no difference whatsoever if we average over many. This may prove useful with data sets of limited size where there may not be a statistically significant number occurrences of a transition for a single Δt. By averaging over many Δt's a more accurate representation of the system's behavior can be obtained.

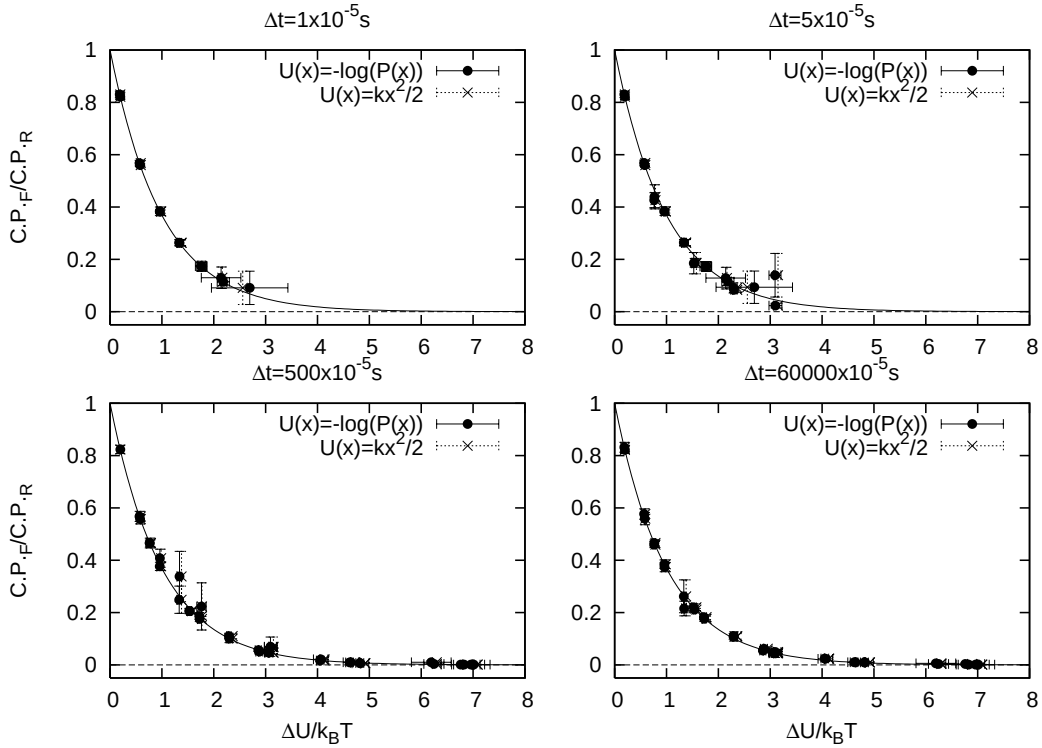

Figure 4.2. Plots of ratios of conditional probabilities for multiple values of Δt. These results are from a simulation with trap stiffness of 0.005 pN/nm and a bin width of 18 nm. Note that C.P.$_F$ and C.P.$_R$ refer to $P(x_f, t + \Delta t|x_i, t)$ and $P(x_i, t + \Delta t|x_f, t)$, respectively.

We also note that there is no correlation between the two different methods used to obtain the potential and the method used for obtaining the conditional probabilities. When we calculate the potential using the equipartition theorem we obtain the mean square displacement from the mean-position of the particle in the trap, $\langle x^2 \rangle$. This *only* requires that we calculate the displacement from the mean-position for each point in the raw, non-binned data set. The only similarity shared between the conditional probabilities and the potential obtained from taking the logarithm of the Boltzmann distribution is that the data must first be binned, in one case to have definite endpoints for the transitions, in the other it is to obtain the probability distribution.

Although the temporal behavior of a stochastic system described by Boltzmann statistics cannot be obtained through consideration of the conditional probabilities, it does elucidate the fact that independent of observation time intervals, the relative likelihood of a transition occurring in the forward or reverse sense can be obtained. There has been previous work investigating the Kramers rate theory for escape from a well and/or barrier hopping by McCann et al. [50]. In an experiment using two narrowly separated optical traps, the motion of a single Brownian particle was used to accurately obtain the three-dimensional potential, as perceived by the particle. The experiment was carried out for a variety of barrier heights and the transition rates (frequencies) were obtained from the average dwell times of the particle within each well. Using the transition rates they verified that the ratio of the transition rates was directly related to $e^{-\Delta U/k_B T}$. This work differs from our work with conditional probabilities in that it only considers a single timescale, the transition rate for a given double-well potential, to obtain one feature of the potential, the barrier height. In our experiments we have shown that measurements made at arbitrary intervals affords a general means of obtaining the potential difference between any two positions/states the particle was observed at/in.

Our relating the ratio of two spatially inverted conditional probabilities to the potential difference between the endpoints may be seen as a special case of a generalized fluctuation-dissipation (F-D) theorem. Generally speaking, F-D theorems have the form, $P(\alpha)/P(-\alpha) = e^{-\Delta\alpha}$, where α is generally the entropy produced during a path [32, 27]. These F-D theorems relate a quantity produced during a path of specified duration (temporal, spatial, number of fluctuations), α, to the ratio of the probability of observing the path and its microscopic reverse. By relating conditional probabilities as we have done places no requirements upon the specific path on the $e^{-\alpha}$ part of the F-D relation, we have implicitly shown that, as long as the endpoints remain unchanged, the probabilities for any forward and reverse paths, i.e. any choice of Δt, may be used to obtain the potential difference between the endpoints, ΔU. This path independence is an indication that the probability distribution is stationary. It should be stated that the F-D theorems are typically generalized to account for time varying potentials/forces, or irreversible processes, thus the probability distributions are generally *not* stationary. This is discussed in greater detail in Section 5.2.2.

The inherent generality of the conclusion we have drawn from detailed balance means that for any system undergoing Gaussian-Markov random fluctuations, not necessarily colloidal Brownian particles, information about the underlying potential is obtainable. Furthermore, even in a limited sampling of data, there may be more information contained in an analysis of conditional probabilities than of the probability distribution. For a set containing N points there will be only N *bits* of information available when we consider the probability distribution (each point is independent). When we consider the conditional probabilities there are exactly $(N-1)!$ *bits* corresponding to the number of pairs of points in the set. In spite of this, it is unclear whether or not the potential will be better described with a smaller set of points from conditional probabilities alone.

4.3 Transition times

In order to measure the last-touch-first-touch time multiple measurements of the particle's position during a single transition are necessary. There is no other way. We were unable to directly measure this in our experiments due to equipment limitations. In spite of its being a very well defined quantity, the LTFTT is a difficult quantity to measure which has no doubt contributed to its not being discussed in the literature, except infrequently and in a theoretical context only [29, 16, 15]. The curious result we wanted to uncover in our trapping experiments was that the LTFTT is identical for uphill and downhill transitions. To our surprise we did uncover an equality between last *observed* first *observed* spatially inverted transition times. What can we conclude from this? In Figure 3.16 we see that the observed transitions occur between one and two measurements (the time interval between measurements was $\Delta t = 33$ ms). Because this $\Delta t \gg$ LTFTT, it is very likely that the particle will not only complete the transition from x_i to x_f in one direction but also complete the reverse transition all during a single Δt. If this is the case, we should not expect to find that on average, after observing the particle at x_i, that our first observation of the particle at x_f should take the same amount of time as the transition from x_f to x_i. We should not expect this to be the case because with a large Δt, successive observations of the particle's position, described by Boltzmann statistics, are uncorrelated. Thus, given $U(x_f) > U(x_i)$, if we make an observation of the particle at x_f, it will take fewer observations (samplings) to find the particle at x_i than if it were traveling in the opposite direction, because $P(x_i) > P(x_f)$. From this we might conclude that the LOFOT relates the frequency that the particle visits x_i to x_f, but that quantity is $P(x_f|x_i)/P(x_i|x_f)$. In point of fact, whether the LOFOT has general significance is undetermined. We reiterate our qualitative explanation for the equality of the LOFOTs, in a given observed transition, the particle will

likely leave the interval between x_i and x_f many times at both transition endpoints. Furthermore, we assume that on average the particle makes a different number of unobserved fluctuations outside the interval at x_i than at x_f. Because we do not observe some number of fluctuations at x_i and x_f, for both forward and backward transitions, the LOFOT is, on average, found to be independent of the transition direction.

When we simulated our experiments but sampled from the finely discretized data at a rate of 30Hz and performed identical analyses to those with our experimental data, we found the results to be indistinguishable. However, when we looked at the entire data-set in which the time interval between the successive points was orders of magnitude smaller, $\Delta t_s \sim 10^{-5} - 10^{-7}$s, we were able to compare simulated LTFTTs with theoretical predictions made in [29]. We alluded to problems with the binning of data to obtain LTFTTs in a previous section. To compare with the well defined expression of Eq.(2.51) we had to analyze our data by starting and stopping our "stopwatch" at the crossings of the endpoints rather than allowing crossing of or entering/exiting a bin to be sufficient. Artifacts from the binning can be minimized by making the bins as small as possible, however, it was better to not bin at all.

We found that simulations carried out with a reflecting barrier at the initial point of the transition allowed us to accurately consider very specific transitions a statistically significant number of times, a minimum of $N = 5 \times 10^4$ in each direction. Simulations were carried out as described in Section 3.6.4 with the additional condition that the particle was confined to the interval of interest. Two aspects of the simulation that affected the accuracy of our results were the protocol used to make the reflecting boundary and the sampling time Δt_s. As we decrease Δt_s the effects of the rules governing the reflecting boundary condition become less significant as the particle tends to move shorter distances between simulation steps. Ultimately we used $\Delta t_s = 10^{-7}\text{s} = 0.1\mu\text{s}$ for our simulations of transitions. Using

such a short time step is especially necessary for transitions spanning distances less than ~ 30 nm. For longer intervals, the transitions take long enough that accurate results could be obtained with $\Delta t_s = 1$ μs, however, we opted to use the same Δt_s for all transition intervals. To make a reflecting boundary condition we found that accurate results could be obtained from a variety of different sets of rules. The most obvious, and simplest, is to assume an elastic collision in which the particle returns to the point it was at before colliding with the boundary, i.e. $x_j \to x_{j-1}$. Another approach, perhaps more physically sensible, is to calculate the distance *beyond* the boundary the particle *would* have traveled, and put the particle that same distance away from the boundary on the allowed side, i.e. $x_j = x_b + (x_b - x'_j)$ where x'_j is the location the particle would have moved to in the absence of the reflecting boundary and x_b is the location of the boundary, which in this example is the *bottom* of the transition interval. A third method we tried was to note when the particle would collide with the boundary but then, without allowing time to step forward, iterate until we find the particle at an allowed position, that is, if x_j is not allowed, we stop time, put the particle at x_{j-1} and calculate new values of x_j until we find one that is allowed, we select this first allowed value and restart time. This last method seems physically unreasonable but interestingly gives accurate results. In these simulations we considered two times, one being the total transition time, from $t = 0$ when $x(0) = x_i$ to the the time when $x(t_f) = x_f$, so that $\langle t_f \rangle$ = MFPT and the second corresponding to the LTFTT which had its "stopwatch" restarted each time the particle collided with the boundary at the initial position so that $\langle t_f - t_{lt} \rangle$ = LTFTT where t_{lt} is the time of the last touching/crossing of x_i. From the difference of these two times we obtained the mean wiggling time (MWT). Results comparing theoretical values for LTFTTs and those obtained from simulations have been tabulated in Table 4.1. Our simulations were all within 5% of the theoretical LTFTTs, for each reflecting boundary condition protocol described. This

indicates that, for very small Δt_s, the specifics of the boundary condition only give subtly different results. The effects are further minimized when we consider longer transition intervals, if we look at the transition from $x_i = 0$ nm to $x_f = 70$ nm there is a discrepancy of only 4μs between the LTFTTs using methods 1 and 2 whereas the difference with the theoretical result is perhaps more relevant with a magnitude of $\sim 25\mu$s. Whether this is of much significance is uncertain, as it equates to less than a 1.5% discrepancy between simulation and theory. Rather than arising from the region surrounding the boundary, this discrepancy is likely due more to the fact that either longer transitions must be observed a greater number of times (N should be increased) to maintain good statistics, or Δt_s should be varied. To further investigate this we simulated the specific transition just discussed with $N = 1 \times 10^5$ while keeping $\Delta t_s = 10^{-7}$ s and found $\overline{\mathrm{LTFTT}_1} = 1.699$ ms and $\overline{\mathrm{LTFTT}_2} = 1.698$ ms. Because the former is identical to and the latter is near to the values in Table 4.1 we consider this sufficient evidence that, at least for this potential, setting $N > 5 \times 10^5$ is not necessary. To check whether changing Δt_s would have any effect, we increased it to 10^{-6} s, a factor of 10 greater than was used in the simulations that generated the data in Table 4.1. Changing Δt_s in this manner, for this transition/interval, is certainly reasonable since the LTFTT is approximately 1.7 ms, roughly $1700 \times \Delta t_s$. Interestingly, keeping $N = 5 \times 10^4$ while using $\Delta t_s = 10^{-5}$s gives a more accurate result, we found $\overline{\mathrm{LTFTT}_1} = 1.736$ ms and $\overline{\mathrm{LTFTT}_2} = 1.738$ ms, with discrepancies of 0.4% and 0.5%, respectively, certainly an improvement over the previously obtained 1.6% and 1.8%. It would seem that shorter time-steps should give better results for all transitions, however, this does not seem to be true (it certainly is for the smaller intervals, but not longer). It is also worth noting that even when $N \to 2.5 \times 10^5$ there is no improvement in accuracy, which may bring us full-circle. Obtaining better results at this point likely requires better, more precise simulations. For transitions occurring on intervals of 10 nm we see

the greatest discrepancies as well as no overlap between theoretical LTFTTs and simulated LTFTTs (after uncertainties are accounted for). This could arise from inaccuracies in the numerical methods used to evaluate the theoretical expressions. Another potential source of error in the simulations is that inertia has not been accounted for, typically this is a safe practice, but when $\Delta t_s \to \Delta t_v = m/\gamma$ there is potential for concern. Considering inertial effects could lead to more accurate results for short transitions where a small number of Brownian *kicks* are required. Conversely, for long transitions, the inertial effects are less significant because there will be many *kicks* in the direction opposite the transition.

There has been recent theoretical work looking at transition times by Bezrukov et al. in which they consider a quantity called the direct translocation time, which is equivalent to our LTFTT. Their direct translocation also takes the same time to transition in each direction [15, 14]. In the first of these papers, the subject of particles traversing membrane channels is investigated, especially channels with potential difference between the ends. This is of course no different from the transitions we have already considered for an arbitrarily shaped potential (with force proportional to the gradient of the potential). Their analysis, however, is different, relying upon a propagator representing the probability density of finding the particle within the channel at position x with velocity v given that it was at one of the channel ends at $t = 0$, with velocity v_0. They define the flux of particles exiting the channel given that they had entered the opposite end. Next, the average over the distribution of initial velocities on entering the channel is obtained. This averaged flux is integrated over all time ($0 \leq t \leq \infty$) to obtain the "translocation" probability density. Finally, the probability density of the "direct translocation time" is the averaged flux divided by the "translocation" probability density. This procedure is carried out for a flux in each direction and, by invoking detailed balance, the two probability densities are shown to be identical. While their initial analysis limited

		Theoretical	Simulation	
x_i (nm)	x_f (nm)	LTFTT (ms)	$\overline{\text{LTFTT}_1}$ (ms)	$\overline{\text{LTFTT}_2}$ (ms)
0	10	0.034	0.0358 ± 0.0003	0.0355 ± 0.0002
0	20	0.138	0.1390 ± 0.0005	0.1398 ± 0.0008
0	30	0.316	0.3162 ± 0.0018	0.3149 ± 0.0018
0	40	0.572	0.5674 ± 0.0028	0.5653 ± 0.0014
0	50	0.905	0.8933 ± 0.0026	0.8939 ± 0.0031
0	60	1.301	1.2828 ± 0.0065	1.2848 ± 0.0026
0	70	1.729	1.6986 ± 0.0054	1.6946 ± 0.0136
10	20	0.034	0.0357 ± 0.0003	0.0357 ± 0.0002
10	30	0.138	0.1394 ± 0.0008	0.1390 ± 0.0006
10	40	0.313	0.3117 ± 0.0008	0.3131 ± 0.0014
10	50	0.56	0.5550 ± 0.0023	0.5548 ± 0.0030
10	60	0.869	0.8615 ± 0.0034	0.8576 ± 0.0046
10	70	1.22	1.1985 ± 0.0053	1.2008 ± 0.0068
20	30	0.034	0.0356 ± 0.0003	0.0356 ± 0.0003
20	40	0.137	0.1385 ± 0.0006	0.1379 ± 0.0006
20	50	0.309	0.3076 ± 0.0011	0.3075 ± 0.0017
20	60	0.544	0.5391 ± 0.0025	0.5385 ± 0.0033
20	70	0.828	0.8190 ± 0.0058	0.8183 ± 0.0028
30	40	0.034	0.0357 ± 0.0003	0.0354 ± 0.0002
30	50	0.136	0.1370 ± 0.0005	0.1372 ± 0.0005
30	60	0.304	0.3024 ± 0.0021	0.3020 ± 0.0011
30	70	0.526	0.5209 ± 0.0016	0.5206 ± 0.0019
40	50	0.034	0.0356 ± 0.0002	0.0356 ± 0.0003
40	60	0.135	0.1357 ± 0.0008	0.1358 ± 0.0006
40	70	0.297	0.2959 ± 0.0010	0.2945 ± 0.0011
50	60	0.034	0.0352 ± 0.0003	0.0352 ± 0.0003
50	70	0.134	0.1346 ± 0.0004	0.1349 ± 0.0008
60	70	0.034	0.0352 ± 0.0002	0.0352 ± 0.0002

Table 4.1. Tabulated values for the theoretical LTFTTs calculated with Eq.(2.51) and simulated LTFTTs for a particle trapped in a potential well with stiffness $k = 0.005$ pN/nm. In the simulations each transition was observed 5×10^3 times, this was performed ten separate times. The over-bar indicates that we have averaged the LTFTTs for each direction of travel and then averaged over the ten different simulations. The subscripts 1 and 2 denote that a different condition was used to determine the particle's location after colliding with the reflecting boundary. Method 1 puts the particle back where it was before colliding with the barrier whereas method 2 notes that a collision occurred and iterates with time stopped, until an allowed location for the particle to move to is obtained.

itself to the space within a membrane channel, they go on to assert that the result holds for any potential which tends to infinity as $x \to \pm\infty$ and in the region of the membrane, is equal to the potential contained therein. While we know that a bounding potential is not necessary in practice, and the LTFTT, or direct translocation time, is defined by the region in which the transition occurs, it facilitates observing single particle dynamics over long times, as we have already seen. Bezrukov et al. conclude that a method identical to ours, studying a colloidal particle, may be used to measure/observe such transitions, they also mention two specific experimental "tests" of the theory [11, 57]. The first referenced paper [11] only discusses digital video microscopy insomuch as it relates to observing and resolving non-interacting and interacting particles in contact with each other, there is no discussion of time resolution, something paramount to the observation of transitions. The second paper deals with a bistable transition in a single molecule experiment. The molecules are dyed so that if laser light is incident, and the molecules are in one of the two states (folded or unfolded), photons are emitted with energy corresponding to the conformational state. This experiment is limited by what seems to be observation of the transition endpoints only. However, an estimate is obtained for an upper bound on the transition time, though this too, is unhelpful in the analysis of the transition process.

Again, we are reminded of the difficulties faced in studying transition times. For single molecules, colloidal particles, and biological processes these transitions typically occur on timescales that are too small to accurately resolve experimentally. A conundrum lies in the apparent ease with which probabilistic information about a transition is obtained while temporal information about the same transition is virtually unobtainable. Furthermore, we might expect that by separating the MFPT into two different times, the LTFTT and the MWT, retrieving the relative probabilities of uphill and downhill transitions could be possible.

4.4 Connections between conditional probabilities and LTFTTs

Detailed balance allows us to relate the ratio of conditional probabilities in the elegant and concise form of Eq.(2.4). This ratio tells us of the relative frequency of transitions occurring in a particular direction. Oftentimes, frequencies are inversely related to timescales. With this in mind, we have sought a relationship between the relative frequency of the direction of a transition and the average transition time, whether the MFPT, $\mathcal{W}$ (MWT), or some combination that may or may not include the LTFTT. The r.h.s. of Eq.(2.4) depends only upon the difference in the potential at the endpoints of the transition interval so we hypothesized that the mean wiggling times might somehow be related to ΔU, also in a simple manner. From our data it was straightforward to obtain the mean wiggling times since $\mathcal{W} = \text{MFPT} - \text{LTFTT}$ for a given transition. Comparing the $\mathcal{W}$s with ΔU was, however, inconclusive. The expression for $\mathcal{W}$ is similar to that for the LTFTT (Eq.(2.51)) [29] and equally difficult to deal with analytically. We write the mean wiggling time at a point x_f for a transition starting at x_f and ending at x_i,

$$\mathcal{W}_{x_i}[x_f] = \frac{1}{DZ'_{[x_i,x_f]}} \int_{x_i}^{x_f} \int_{y}^{x_f} \int_{x_i}^{z} e^{[U(y)-U(z)+U(v)]/k_BT} dv dz dy, \tag{4.2}$$

the subscript x_i denotes the location of the transition endpoint and $Z'_{[x_i.x_f]}$ is the same normalization factor that appears in the LTFTT expression. For an arbitrary potential, the expression is quite complicated to solve analytically. For a quadratic potential the integrals yield error functions that do not readily lend themselves to solution, but an analytic expression is easily obtained when we consider a linear potential.

Consider the linear potential $U(x) = \rho x$, with ρ being an arbitrary constant. We define our transition endpoints as x_i and x_f ($x_i, x_f \neq \pm\infty$). Furthermore, we require that $U(x_i) \leq U(x) \leq U(x_f)$. Writing the expressions for the $\mathcal{W}$s, where $\alpha = \rho/k_B T$, we have

$$\mathcal{W}_{x_i}[x_f] = \frac{1}{DZ'_{[x_i,x_f]}} \frac{1}{\alpha^2} e^{\alpha x_i} \left\{ -2(x_f - x_i) + \frac{1}{\alpha} \left[e^{\alpha(x_f - x_i)} - e^{-\alpha(x_f - x_i)} \right] \right\}, \quad (4.3)$$

$$\mathcal{W}_{x_f}[x_i] = -\frac{1}{DZ'_{[x_f,x_i]}} \frac{1}{\alpha^2} e^{\alpha x_f} \left\{ -2(x_f - x_i) + \frac{1}{\alpha} \left[e^{\alpha(x_f - x_i)} - e^{-\alpha(x_f - x_i)} \right] \right\}. \quad (4.4)$$

Noting that $Z'_{[x_i,x_f]} = -Z'_{[x_f,x_i]}$ we see that

$$\begin{aligned} \frac{\mathcal{W}_{x_i}[x_f]}{\mathcal{W}_{x_f}[x_i]} &= e^{\alpha(x_i - x_f)}, \\ &= e^{-\Delta U/k_B T}. \end{aligned} \quad (4.5)$$

This is the simple and useful relationship we sought, it relates the times that the particle spends around the initial point of a transition before making the actual direct transition to the potential difference between the transition endpoints. To verify this result we used simulations identical to those described in the previous section that consider specific transitions, many times. We used a potential analogous to gravity to make our results experimentally accessible. The effective mass for a buoyant particle is $m_b = m_p - m_w$, where m_p and m_w are the mass of the particle in air and the mass of a volume of water equal to the particle volume, respectively. By calculating the MFPTs and the LTFTTs we were able to obtain the $\mathcal{W}$s. These times were then compared with predicted theoretical values. Table 4.2 contains the mean wiggling times and their ratios for simulations of a Brownian particle in water and subject to gravitational forces only. The buoyancy of the particle effectively reduces the mass from $m_p \simeq 0.5$ pg to $m_b \simeq 0.03$ pg. For these simulations we found that using the more physically sensible reflecting boundary condition that places the particle at its last allowed location when collisions with the boundary occur, and not the condition that allowed further iterations until an allowed position was

		Theory			Simulation		
Δx	$e^{-\Delta U/k_BT}$	$\mathcal{W}_a[b]$	$\mathcal{W}_b[a]$	$\mathcal{W}_a[b]/\mathcal{W}_b[a]$	$\mathcal{W}_a[b]$	$\mathcal{W}_b[a]$	$\mathcal{W}_a[b]/\mathcal{W}_b[a]$
50	0.9968	1.614	1.619	0.9968	1.674	1.694	0.9882
100	0.9936	6.560	6.602	0.9936	6.637	6.772	0.9801
150	0.9904	14.875	15.020	0.9904	14.982	15.021	0.9974
200	0.9872	26.270	26.612	0.9872	26.343	26.733	0.9854
250	0.9840	41.002	41.670	0.9840	41.046	41.911	0.9794
300	0.9808	59.055	60.211	0.9808	60.024	61.034	0.9836
350	0.9776	80.142	81.974	0.9776	79.511	80.702	0.9852
400	0.9745	104.518	107.254	0.9745	104.331	107.212	0.9731
450	0.9713	132.137	136.035	0.9713	132.460	136.588	0.9698
500	0.9682	162.804	168.149	0.9682	164.072	168.496	0.9737

Table 4.2. A tabulated comparison of theoretical and simulated mean wiggling times. The units for Δx are *nano*-meters, and the $\mathcal{W}$s are all in *milli*-seconds. In the simulations each transition was observed 2.5×10^4 times and we have assumed a perfectly reflecting boundary, i.e. the particle is placed back at its position prior to colliding with the boundary.

found, only gave physically meaningful results. Again, symbolically, the boundary condition is $x_j \to x_{j-1}$ if $x_i \nleq x_j \nleq x_f$. It is difficult to compare the individual $\mathcal{W}$s because, for $\Delta x < 300$nm, $\mathcal{W}_{x_i}[x_f] - \mathcal{W}_{x_f}[x_i] \lesssim 1$ ms, and certain $\delta\mathcal{W}$s are of similar size. Because of this we have opted to make a more quantitative comparison with the ratios of these times. In Figure 4.3 the ratios have been divided by $e^{-\Delta U/k_BT}$ and plotted against the line $y = 1$. All ratios considered here fall within 0.8% of the expected values, verifying the accuracy of the simulation as well as the validity of the theoretical predictions.

Our showing that for linear potentials the ratio of mean wiggling times is equal to the ratio of Boltzmann factors for the transition endpoints is evidence that, in general, the ratio of MFPTs will not give the same result. However, because we have only showed this to be true for linear potentials, we are unable to make any claims as to the general relevance of the ratio. It may be that the ratio is related to $e^{-\Delta U/k_BT}$, but in a yet unknown manner. There is still merit in studying the problem

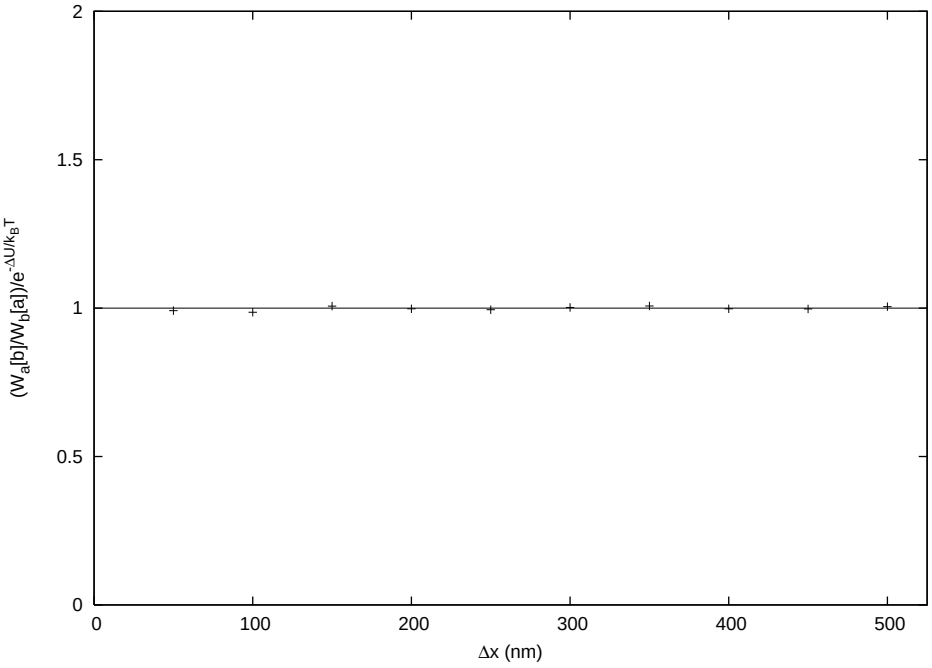

Figure 4.3. The ratios of simulated mean wiggling times for a Brownian particle in a linear potential. For this special case the ratio should equal $e^{-\Delta U/k_B T}$. Here the ratio of $\mathcal{W}$s is divided by the decaying exponential and plotted against Δx which is proportional to ΔU. For a simulation time step $\Delta t_s = 10^{-6}$ s and $N = 2.5 \times 10^4$ transitions for each Δx the fit is very good being within 0.86% for each Δx.

of the linear potential. It is quite accessible from theoretical, computational, and experimental perspectives, which makes it a practical tool for the classroom.

4.5 Summary

In a stochastic system, all events are probabilistic, moving from point x_i to point x_f in any context will always be somewhat uncertain. As we have already seen with the conditional probabilities, interesting subtleties exist within the simple framework of a transition between two points. To the untrained eye it may seem bizarre that, from observations made over any time interval, the same information can be obtained about a transition, a transition that on average will take a certain time. But this is not bizarre at all, it is simply a consequence of microscopic reversibility.

Once we recognize the implications of microscopic reversibility, we see that just as the probabilities for forward and backward transitions are related, so too are the timescales on which these transitions occur. From microscopic reversibility we have asserted that the time to traverse the interval between x_i and x_f will, on average, be identical during a *direct* crossing (LTFTT) and shown this to be true. That the LTFTT is identical for uphill and downhill travel is rather counter-intuitive. In the context of a staircase, what we are saying is that it will take, on average, the same amount of time to step to a higher step as it will to a lower step, given this, we quickly conclude that stepping up n steps will take the same time as stepping down n steps on average, something we don't typically expect to be true. This can also be cast in terms of paths and actions. Here we find that the most probable uphill and downhill paths are simply reverses of each other, and thus occur on identical timescales.

Chapter 5
FINAL REMARKS

5.1 Overview

Investigations into the random, confusing sea of fluctuations that define stochastic systems and processes can be quite complicated. The inherent *noisiness* of these systems challenges our physical intuitions about the world we live in. That said, from our experiments we have been able to show, for the first time, that for a time independent potential, the ratio of two conditional probabilities that are spatial inverses is equal to $e^{-\Delta U/k_B T}$, for all time intervals separating the initial and final positions. This is an interesting result, not just because it holds for long timescales but because it also holds for very short timescales. For this relationship to be true even when we consider a time interval shorter than the average time to move directly across the interval (LTFTT) is curious, indeed. We must remember that what matters is the *relative* likelihoods and not the *absolute* probabilities of observing either transition during the time interval. In regard to transition times, we have primarily been concerned with the last-touch-first-touch time (LTFTT). From microscopic reversibility we can argue that the LTFTT will, on average, be the same for transitions in the forward and reverse directions. We simulated colloidal particles in quadratic potential wells analogous to those used in our experiments to verify this. By temporally discretizing the Langevin equation we simulated the Brownian motion of the particle using a time step shorter than the LTFTT. To ensure that we were recovering the LTFTTs from our simulations and not any other transition time we compared these with times obtained from the theoretical predictions of Astumian and Derényi [29]. We found that in our experiments, where images were collected

at 30 fps, we observed a similar phenomenon to the equality of LTFTTs in that the LOFOTs also obeyed this, strangely. Our explanation for this equality is straightforward. When the time between observations is greater than the LTFTT for the transition of interest, we assume that the particle leaves the transition interval an undetermined number of times between our last observation of it at the initial position and our first observation of it at or beyond the final position. Our justification for the equality of the LOFOTs is that observation of the transition requires that we observe the particle at or beyond each endpoint so that, on average, we miss the same number of excursions outside the interval at each endpoint regardless of which transition direction we are interested in. This is simply a curious result and should not overshadow the more important result that, for the first time, there has been verification of the directional independence of the LTFTTs.

There are two primary shortcomings of this work. The first is the lack of an experimental measure of a LTFTT. We were limited in our experimental apparatus and could not measure this. In hindsight, specifically tailoring an experiment to measure LTFTTs may not be possible within our lab. The second shortcoming of this work is its focus on time-independent potentials. Although we can think of many naturally occurring processes as taking place within a constant potential, there is significant interest in *nonequilibrium* fluctuations with time varying forces performing irreversible work, as we will further discuss in the next section. Before moving on, let us emphasize the generality of the work of Onsager and Machlup. We have exploited what was termed the "one-gate conditional probability" [52] to show that for a non-fluctuating potential, this expression holds for arbitrary timescales. If, however, a dissipative force and/or fluctuating potential is present, the expression for the conditional probabilities must be dealt with in a more general manner, the path that is followed between the transition endpoints must be accounted for, as in Eq.(2.13), discussed by Bier et al. [16], and more recently by Astumian [8, 9].

5.2 Extending this work

The nature of scientific research is that it is an ongoing, self-sustaining process. As new measurements are made, new questions are asked, revised measurements are made, and so on. Here we offer a brief discussion of potential, justifiable, and reasonable extensions of our work.

5.2.1 Furthering our experiments

The first, most obvious, experiment that should follow this work is one that shows the equality between uphill and downhill LTFTTs. There are a number of experiments that could be carried out to measure these transition times. Using a single colloidal particle in an optical trap is one possibility, provided the sampling rate is sufficiently fast ~ 10 kHz. Sampling at this rate is unachievable with digital video microscopy (using cameras to collect images) so the use of photo-diodes is necessary. Alternatively, the motion of the particle is further damped using a more viscous medium (this decreases the distance the particle is likely to travel during any time interval). One recent experiment that may allow transitions to be observed over longer lengths and timescales can be found in Refs. [64, 25, 58, 55]. In these experiments a single optical tweezer is reconfigured into a linear trap. If one end of the trap could be given a higher potential than the other, motion along the trap may be tuned so that transitions could be observed over long time periods. Another experiment would place a colloidal particle in a volume such that its motion in the direction of the gravitational force can be observed. As we have already seen, certain transitions in a gravitational potential occur on timescales greater than 1/30 s, a typical camera frame-rate. Experimental verification of the equality between LTFTTs would further emphasize the validity of considering the MFPT as two distinct processes, one occurring at the initial position of the transition, the

mean-wiggling time $\mathcal{W}_{x_f}[x_i]$ (MWT), and the other representing the actual direct transition, the last-touch-first-touch time LTFTT$[x_i \to x_f]$.

5.2.2 Generalized fluctuation-dissipation theorems

The ratio of conditional probabilities in relation to $e^{-\Delta U/k_B T}$ is often referred to as a fluctuation-dissipation theorem. We have relied upon this to provide us with a way of relating the random fluctuations during a reversible process to the energy fluctuation during that process. In our system this was the potential energy difference, ΔU, between the transition endpoints. Two concepts fundamental to the fluctuation-dissipation theorem were described by Onsager in Refs. [53, 54, 52]. These concepts are microscopic reversibility and the regression hypothesis. From the principle of microscopic reversibility, which relies upon mechanical equilibrium and not necessarily thermodynamic equilibrium, we can relate the average likelihood of fluctuations occurring in a "forward" direction to the "reverse" direction. In fact, on average, a process and its microscopic reverse are equally probable. From Ref. [52], the regression hypothesis states that if we find a system in a state possibly arising from an irreversible process, it will regress in a manner exactly as it would if it had arrived at that state by fluctuations alone. This is analogous to the assumption that a system/process is Markovian, i.e. the future of the system depends only upon its present state. This is important if we wish to generalize the fluctuation-dissipation theorem to irreversible processes in which there is a perturbation, not the special case in which we construct an initial state, by any means, and observe the response of the system.

In recent years numerous generalized fluctuation-dissipation relations have been derived and extensively discussed [44, 21, 26, 27, 28, 7, 41, 8, 9]. However, in the late 1970's Bochkov and Kuzovlev derived a generalized fluctuation-dissipation relation which preceded those of Jarzynski and Crooks [19, 20]. The results of Bochkov and

Kuzovlev were derived under conditions such that the initial state of the system is in equilibrium, the irreversible process that is present as the system follows a specific path ψ_F to its final state is started at $t = 0$, that is, an external [time varying] force, X_t, is applied as ψ_F begins. The value of the applied force during ψ_F is defined as X_t^F, the initial and final values are X_0^F and X_τ^F, respectively. In the context of the reverse path, ψ_R, the system begins at the final state of the forward path in the absence of any external force. At this instant the external force is switched on with value $X_0^R = X_\tau^F$ and ψ_R is followed. At each instant during the reverse path we have $X_t^R = X_{\tau-t}^F$. To summarize, we begin with a canonical distribution and let the system follow a forward path while applying a varying force, we then repeat, beginning at the final position while applying the varying force in a time reversed sense. Bochkov and Kuzovlev related the probabilities of observing the ψ_i's to the energy difference between the initial and final states of the system. Their result may be written in terms of the external work W_{ext} that is performed on the system by the applied force

$$\frac{P(\psi_F)}{P(\psi_R)} = e^{W_{ext}/k_BT}. \tag{5.1}$$

It is important to note that W_{ext} is actually the combined work of the applied force X and the potential U along the path. A similar result was later obtained by Crooks [26], however, in this formulation the system is taken to be described by a distribution in which there *is* an applied force present before ψ_i begins; for the forward realization we have the system equilibrated in the presence of X_0^F, and for ψ_R we would have X_0^R. The relation that Crooks obtained, while similar to that of Bochkov and Kuzovlev (Eq.(5.1)) is not exactly identical,

$$\frac{P(\psi_F)}{P(\psi_R)} = e^{(W-\Delta F)/k_BT}. \tag{5.2}$$

Rather than having W_{ext} in the exponential, we have $W-\Delta F$, W represents only the work performed by the applied force and ΔF is the free-energy difference between

the equilibrium endpoint states. Another relation, obtained by Jarzynski in Ref. [44] is

$$\langle e^{-W/k_BT} \rangle = e^{-\Delta F/k_BT}, \tag{5.3}$$

where W is the external work and the entire l.h.s. represents the average of the "work probability" for many experiments carried out over a specific and finite time, and ΔF is the free energy difference between the initial and final states. This result is experimentally accessible because of its averaging over many experiments. The significance of Jarzynski's result is that, despite each experimental realization giving a different result, the average result will tell us something of the underlying equilibrium free energy difference! This is tremendously powerful and there have been numerous experimental tests of this [49, 23, 18].

Let us examine one experiment designed to test Jarzynski's result, Eq.(5.3). Bustamante and colleagues stretched a single RNA molecule to induce a configurational change of the molecule such that the final state was "far from thermodynamic equilibrium" [49]. By attaching micron sized latex beads to each end of the RNA and pulling on them, Bustamante et al. could exert measurable forces and do measurable work on the molecule. The pulling was carried out by holding one bead with a micro-pipette and moving it at known rates while the second bead was trapped in an optical tweezer. Forces were determined by the position of the bead within the trap. Typically, RNA is folded over onto itself and virtually never found in a straightened configuration. Application of external work causes the molecule to lengthen and stretch, that is, transition to the final state of higher potential energy. Bustamante et al. showed that when the pulling was carried out at rates greater than that for reversible transitions (reversibility requires that the work performed in stretching the RNA is equal to the work done by the RNA in contracting), experiments carried out far from equilibrium allowed them to obtain the equilibrium free energy

difference between the initial and final configurations, as predicted. Experiments of this nature do elucidate those rarely visited places on the energetic landscapes of nature by driving processes, making them more favorable in order to study them. However, the terminology that this work has given rise to, "nonequilibrium paths" and "nonequilibrium trajectories" for single particles not in equilibrium is misleading, at minimum. It is unclear what a "nonequilibrium path" might actually be. There is always a finite probability that any particular path may be realized, even at equilibrium. Exactly what it means for a single particle to not be in equilibrium is both ambiguous and unclear. What is overlooked is that these experiments are actually carried out in mechanical equilibrium, although they do dissipate energy and are not in thermodynamic equilibrium[7]. The necessary condition for mechanical equilibrium is that of force balance. At every instant in time, the net force is zero and inertial effects are not present, or negligible. This is the case for small objects in aqueous solution, the world of low Reynolds number, as we have extensively discussed. We can further emphasize this with a few calculations of the power dissipated during these single molecule experiments. The characteristic thermal relaxation time for a Brownian particle can be expressed as $\tau_{ch} = m/\gamma$, for a nm sized object $\tau_{th} \sim 10^{-12}$s. We estimate the power associated with thermal noise using $P_{th} = k_B T/\tau_{th}$ and find $P_{th} \sim 10^{-}9$ W, or 1 nW. To estimate the power dissipated by the external force during a typical AFM pulling experiment we use values from Ref. [7], with a force applied at a rate as large as 100 nN/s we might expect an enormous amount of power will be dissipated, but with the pulling taking place over 100 nm we calculate a power $P_{ext} = \Delta x dF/dt \sim vF \sim 10^{-14}$ W. Similarly, we can obtain a different estimate by relating the power dissipated to the measured energy difference, in Ref. [49] $\Delta F \approx 60 k_B T$, when this is measured on an interval of 30 nm we find an average force of ~ 8 pN which in turn can be used to calculate a power from $P_{ext} = F^2/\gamma \sim 10^{-12}$W. Noting that the pulling is

performed at a rate of $\sim$ 50 pN/s over a distance $\Delta x \approx 30$ nm we conclude that the power dissipated by the pulling force is $P_{ext} = \Delta x dF/dt \sim 1.5^{-19}$ W. What is obvious from these calculations is that these experiments, proclaimed to take place far from equilibrium, remain in mechanical equilibrium. Consequently, microscopic reversibility and Onsager's regression hypothesis are most certainly obeyed.

Using the conditional probabilities presented herein we can imagine alternative methods to those described above for obtaining information about the energetic landscapes of small, fluctuating systems. Astumian suggested that an external field could be used to skew the intrinsic potential of the system[7] by making unfavorable transitions more favorable and frequent. Imagine a Brownian particle in a double-well potential containing stable points at a and b, these are separated by a barrier with height β relative to U_a with $U_0 = U_b - U_a$. Let us take β to be large enough that transitions between a and b are effectively nonexistent. To measure U_0 by observing the particle's motion is rather difficult, it would not be sensible to wait for transitions between a and b to occur enough times to be statistically significant. We can circumvent this by applying an external field, thereby altering the probability distribution in a predicted manner. The net effect is to make transitions from a to b more favorable. Ideally, the combined potential (net force) is such that transitions are equally favorable in either direction. Through successive observations of the transition, over *any* timescale, we can extract the equilibrium potential difference between the two stable points, U_0, by using nothing more than the ratio of the spatially inverted conditional probabilities. We offer this as an alternative means of obtaining the energy differences between those places/states that are frequently and infrequently visited. In general, we are interested in the likelihood that an event will occur, this tells us something of the thermodynamics governing the system, however, what it won't necessarily tell us is the timescales that the event occurs on.

We conclude by returning to the work of Onsager and Machlup. In their treatment of irreversible processes their only assumptions were that the processes be linear, "i.e. that the fluxes depend linearly on the forces that 'cause' them," and that the fluctuations be Gaussian[52]. In the context of Brownian motion, the claim is that the "flux" (particle velocity) is linearly related to the *net* force causing it, the imperative word being *net*. This means that the Onsager-Machlup description of irreversible processes may be used to describe time varying systems. Astumian has emphasized this point extensively [7, 8, 9]. One prediction is that symmetry related paths can be elegantly related, in a form similar to all of the the generalized fluctuation-dissipation theorems,

$$\frac{P(S_i)}{P(S_j)} = e^{(S_j - S_i)/\gamma}, \tag{5.4}$$

where S_i and S_j are the actions for two symmetry related paths. Using data from our simulations we have verified this result for short paths. We say "short" paths because, using binned data, our initial investigations have only considered three and four step paths, not those of arbitrary length. Figures 5.1 and 5.2 are comparisons between the ratio of probabilities of forward and reverse realizations of the paths, ψ_F and ψ_R, with associated actions S_F and S_R, respectively. We have plotted the line $e^{(S_R - S_F)/\gamma}$ for reference. The agreement between these is excellent and warrants further investigation. Extending this result, Eq.(5.4), to systems with time-varying forces could, rather than further split the various generalized fluctuation-dissipation theorems, serve to unify them.

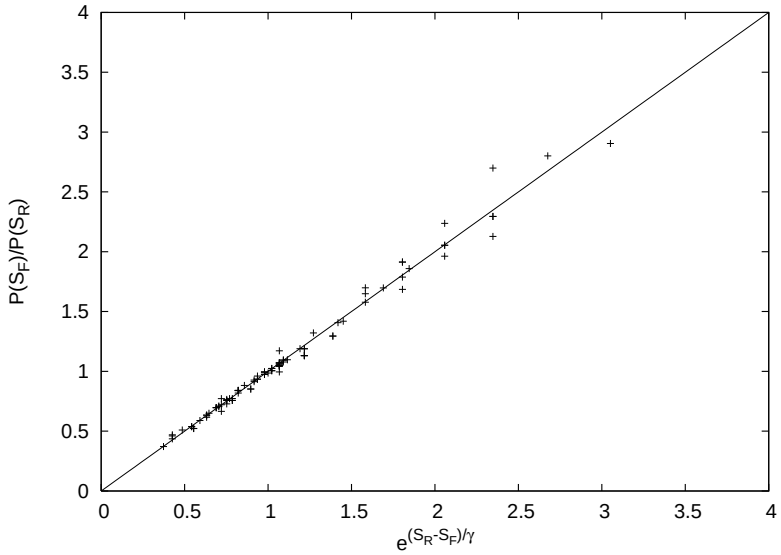

Figure 5.1. The ratio of the probabilities for forward and reverse, ψ_F and ψ_R, realizations of 100 randomly chosen three-step paths is equal to the exponential of the difference between the Onsager-Machlup thermodynamic actions for each path. The data for this plot was obtained from a simulation of a trapped Brownian particle with $k = 0.005$ pN/nm.

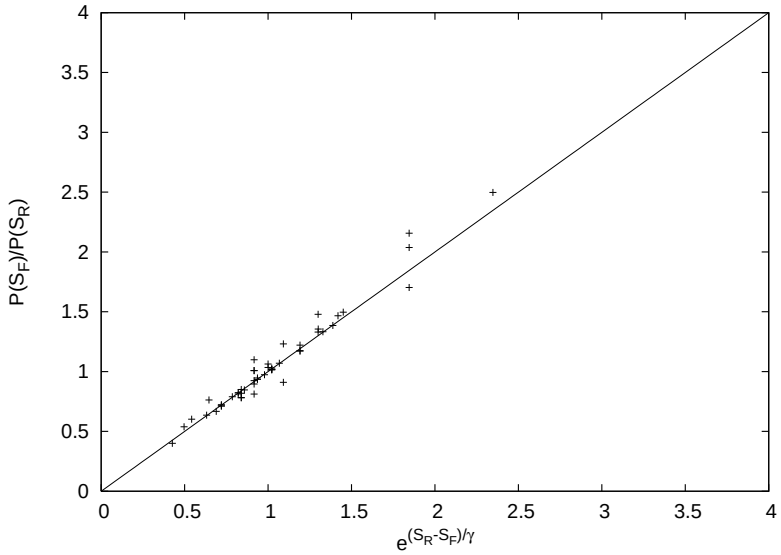

Figure 5.2. The ratio of the probabilities for forward and reverse, ψ_F and ψ_R, realizations of 50 randomly chosen four-step paths is equal to the exponential of the difference between the Onsager-Machlup thermodynamic actions for each path. The data for this plot was obtained from a simulation of a trapped Brownian particle with $k = 0.005$ pN/nm.

REFERENCES

[1] A. Ashkin. Acceleration and trapping of particles by radiation pressure. *Phys. Rev. Lett.*, 24(4):156–159, Jan 1970.

[2] A. Ashkin and J. M. Dziedzic. Optical levitation by radiation pressure. *Applied Physics Letters*, 19(8):283–285, 1971.

[3] A Ashkin and JM Dziedzic. Optical trapping and manipulation of viruses and bacteria. *Science*, 235(4795):1517–1520, 1987.

[4] A. Ashkin and J.M. Dziedzic. Internal cell manipulation using infrared laser traps. *Proc. Nat. Acad. Sci. USA*, 86(20):7914–7918, 1989.

[5] A. Ashkin, J.M. Dziedzic, and T. Yamane. Optical trapping and manipulation of single cells using infrared laser beams. *Nature*, 330(6150):769–771, 1987.

[6] Arthur Ashkin. Optical trapping and manipulation of neutral particles using—lasers. *Proceedings of the National Academy of Sciences*, 94(10):4853–4860, 1997.

[7] R. Dean Astumian. The unreasonable effectiveness of equilibrium theory for interpreting nonequilibrium experiments. *American Journal of Physics*, 74(8):683–688, 2006.

[8] R. Dean Astumian. Equilibrium theory for a particle pulled by a moving optical trap. *The Journal of Chemical Physics*, 126(11):111102, 2007.

[9] R. Dean Astumian. Symmetry relations for trajectories of a brownian motor. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 76(2):020102, 2007.

[10] R. Dean Astumian and I. Derényi. Fluctuation driven transport and models of molecular motors and pumps. *Eur. Biophys. J.*, pages 474–489, 1998.

[11] J. Baumgartl and C. Bechinger. On the limits of digital video microscopy. *Europhys. Lett.*, 71:487–493, 2005.

[12] John Bechhoefer and Scott Wilson. Faster, cheaper, safer optical tweezers for the undergraduate laboratory. *American Journal of Physics*, 70(4):393–400, 2002.

[13] Alexander M Berezhkovskii and Sergey M Bezrukov. Site model for channel-facilitated membrane transport: invariance of the translocation time distribution with respect to direction of passage. *Journal of Physics: Condensed Matter*, 19(6):065148 (10pp), 2007.

[14] Alexander M. Berezhkovskii and Sergey M. Bezrukov. Counting translocations of strongly repelling particles through single channels: Fluctuation theorem for membrane transport. *Physical Review Letters*, 100(3):038104, 2008.

[15] Alexander M. Berezhkovskii, Gerhard Hummer, and Sergey M. Bezrukov. Identity of distributions of direct uphill and downhill translocation times for particles traversing membrane channels. *Physical Review Letters*, 97(2):020601, 2006.

[16] Martin Bier, Imre Derényi, Marcin Kostur, and R. Dean Astumian. Intrawell relaxation of overdamped brownian particles. *Physical Review E (Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics)*, 59(6):6422–6432, 1999.

[17] Daniel Blair and Eric Dufresne. The matlab particle tracking code repository. http://physics.georgetown.edu/matlab/.

[18] V. Blickle, T. Speck, L. Helden, U. Seifert, and C. Bechinger. Thermodynamics of a colloidal particle in a time-dependent nonharmonic potential. *Physical Review Letters*, 96(7):070603, 2006.

[19] G.N. Bochkov and Yu.E. Kuzovlev. *Sov. Phys. JETP*, 45:125, 1977.

[20] G.N. Bochkov and Yu.E. Kuzovlev. Nonlinear fluctuation-dissipation relations and stochastic models in nonequilibrium thermodynamics. *Physica*, 106A:443–479, 1981.

[21] V. Chernyak, M. Chertkov, and C. Jarzynski. Dynamical generalization of nonequilibrium work relation. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 71(2):025102, 2005.

[22] Steven Chu, J. E. Bjorkholm, A. Ashkin, and A. Cable. Experimental observation of optically trapped atoms. *Phys. Rev. Lett.*, 57(3):314–317, Jul 1986.

[23] D. Collin, F. Ritort, C. Jarzynski, S.B. Smith, I. Tinoco Jr, and C. Bustamante. Verification of the crooks fluctuation theorem and recovery of rna folding free energies. *Nature -London-)*, 7056:231–234, 2005.

[24] J. C. Crocker and D. G. Grier. Methods of digital video microscopy for colloidal studies, 1996.

[25] J. C. Crocker, J. A. Matteo, A. D. Dinsmore, and A. G. Yodh. Entropic attraction and repulsion in binary colloids probed with a line optical tweezer. *Phys. Rev. Lett.*, 82(21):4352–4355, May 1999.

[26] Gavin E. Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. *J. Stat. Phys.*, 90(5/6):1481–1487, 1998.

[27] Gavin E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E (Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics)*, 60(3):2721–2726, 1999.

[28] Gavin E. Crooks. Path-ensemble averages in systems driven far from equilibrium. *Physical Review E (Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics)*, 61(3):2361–2366, 2000.

[29] Imre Derényi and R. Dean Astumian. Intrawell relaxation time: The limit of the adiabatic approximation. *Physical Review Letters*, 82(13):2623–2627, 1999.

[30] Eric R. Dufresne and David G. Grier. Optical tweezer arrays and optical substrates created with diffractive optics. *Review of Scientific Instruments*, 69(5):1974–1977, 1998.

[31] A. Einstein. On the motion-required by the molecular kinetic theory of heat-of small particles suspended in a stationary liquid. *Ann. D. Phys.*, 17:549–560, 1905.

[32] Denis J. Evans, E. G. D. Cohen, and G. P. Morriss. Probability of second law violations in shearing steady states. *Phys. Rev. Lett.*, 71(15):2401–2404, Oct 1993.

[33] R. Feynman, R. Leighton, and M. Sands. *The Feynman Lectures on Physics*, volume I, chapter 46. Addison-Wesley, second edition, 1964.

[34] J. T. Finer, R. M. Simmons, and J. A. Spudich. Single myosin molecule mechanics: piconewton forces and nanometre steps. *Nature*, 368(6467):113–119, Mar 1994. 10.1038/368113a0.

[35] Coriander: The Linux GUI for IEEE1394/Firewire cameras, 2007. http://damien.douxchamps.net/ieee1394/coriander/index.php.

[36] C.W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer, third edition, 2004.

[37] Stephen Gasiorowicz. *Quantum Phyisics*. John Wiley & Sons, Inc., 1996.

[38] Daniel T. Gillespie. *Markov processes-An introduction for physical scientists*. Academic Press, San Diego, CA, 1st edition, 1992.

[39] Daniel T. Gillespie. The mathematics of brownian motion and johnson noise. *American Journal of Physics*, 64(3):225–240, 1996.

[40] David J. Griffiths. *Introduction to Electrodynamics*. Prentice Hall, 1999.

[41] Jordan Horowitz and Christopher Jarzynski. Comparison of work fluctuation relations. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(11):P11002, 2007.

[42] ImageMagick. A software suite able to convert, edit, and compose images. http://www.imagemagick.org/script/index.php.

[43] John David Jackson. *Classical Electrodynamics*. John Wiley & Sons, 1999.

[44] C. Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14):2690–2693, 1997.

[45] D. Keller, C. Bustamante, G. Wuite, S. Smith, and M. Young. Single-molecule studies of the effect of template tension on t7 dna polymerase activity. *Nature*, 404(6773):103–106, 2000.

[46] Matthew J. Lang and Steven M. Block. Resource letter: Lbot-1: Laser-based optical tweezers. *American Journal of Physics*, 71(3):201–215, 2003.

[47] Richard L. Liboff. *Introductory Quantum Mechanics*. Addison Wesley, 1998.

[48] J. Liesener, M. Reicherter, T. Haist, and H.J. Tiziani. Multi-functional optical tweezers using computer-generated holograms. *Opt. Comm.*, 185:77–82, 2000.

[49] Jan Liphardt, Sophie Dumont, Steven B. Smith, Jr. Tinoco, Ignacio, and Carlos Bustamante. Equilibrium Information from Nonequilibrium Measurements in an Experimental Test of Jarzynski's Equality. *Science*, 296(5574):1832–1835, 2002.

[50] L. I. McCann, M. Dykman, and B. Golding. Thermally activated transitions in a bistable three-dimensional optical trap. *Nature*, 402(6763):785–787, Dec 1999. 10.1038/45492.

[51] D. N. Moothoo, J. Arlt, R. S. Conroy, F. Akerboom, A. Voit, and K. Dholakia. Beth's experiment using optical tweezers. *American Journal of Physics*, 69(3):271–276, 2001.

[52] L. Onsager and S. Machlup. Fluctuations and irreversible processes. *Phys. Rev.*, 91(6):1505–1512, Sep 1953.

[53] Lars Onsager. Reciprocal relations in irreversible processes. i. *Phys. Rev.*, 37(4):405–426, Feb 1931.

[54] Lars Onsager. Reciprocal relations in irreversible processes. ii. *Phys. Rev.*, 38(12):2265–2279, Dec 1931.

[55] M. Polin, Y. Roichman, and D. Grier. Auto-correlated colloidal interaction measurements with extended optical traps. *Phys. Rev. E—accepted 2008*, 2007.

[56] E. M. Purcell. Life at low reynolds number. *American Journal of Physics*, 45(1):3–11, 1977.

[57] E. Rhoades, M. Cohen, B. Schuler, and G. Haran. Two-state folding observed in individual protein molecules. *Journal of the American Chemical Society*, 126(45):14686–14687, 2004.

[58] Y. Roichman and D.G. Grier. Projecting extended optical traps with shape-phase holography. *Opt. Lett.*, 31(11):1675–1677, 2006.

[59] M. Schnitzer, K. Vischer, and S. Block. Force production by single kinesin motors. *Nature Cell Bio.*, 2:718–723, 2000.

[60] Stephen P. Smith, Sameer R. Bhalotra, Anne L. Brody, Benjamin L. Brown, Edward K. Boyda, and Mara Prentiss. Inexpensive optical tweezers for undergraduate laboratories. *American Journal of Physics*, 67(1):26–35, 1999.

[61] `bash`. In the Linux operating system `bash` is one of many *shells* that act as the interface through which the user passes commands to the kernel. Programs and routines, or *shell scripts* can be written within the shell to perform tasks and jobs.

[62] Tsvi Tlusty, Amit Meller, and Roy Bar-Ziv. Optical gradient forces of strongly localized fields. *Physical Review Letters*, 81(8):1738–1741, 1998.

[63] G. E. Uhlenbeck and L. S. Ornstein. On the theory of the brownian motion. *Phys. Rev.*, 36(5):823–841, Sep 1930.

[64] Ritu Verma, J. C. Crocker, T. C. Lubensky, and A. G. Yodh. Entropic colloidal interactions in concentrated dna solutions. *Phys. Rev. Lett.*, 81(18):4004–4007, Nov 1998.

[65] Mark C. Williams. Optical tweezers: Measuring piconewton forces, 2002. http://www.biophysics.org/education/williams.pdf.

Appendix A

OUR TRACKING SOFTWARE

We wrote a C program to locate the center of a bright, round object within a .png image. The specific program we describe here is capable of processing multiple sequentially numbered files so that images may be processed in batches. The program is named `multi_implay.c` and the code for it may be found in Appendix B.

Code description

Upon reading the pixel data from the grayscale image into a pointer of size $W \times H$, where W and H are the width and height of the image in pixels, we locate the brightest 2% of the pixels and store these locations in `nbright`. Once we have these pixels we select an approximate center by calculating the total brightness in a square of length `dim` around each of the `nbright` pixels, the brightest is selected and its location is stored as `app_center`. We then calculate the centroid of a square of radius l around `app_center` and use that as our particle center.

This code is not very sophisticated. It does not deal well with speckled images or images containing multiple bright regions. However, it was designed as such, we expected each image to contain one and only one bright spot that is brightest in the center and fades radially.

Comparison with the MATLab software

We can easily compare our software with the MATLAB software. First we calculate the center of the particle that we already analyzed in Sec. 3.4, Fig. 3.10.

When we run our software on this image we find the center to be at (60.50,57.92), the MATLAB software found the center to be at (60.68,58.77). The x-coordinate differs by a mere 0.18 pixels whereas in the y-coordinate the difference is a more significant 0.85 pixels. Whether or not this is meaningful is impossible to tell without analyzing many images and comparing the centers from each method with the images themselves and/or the relevant statistics that result from each set of coordinates, i.e. $\langle x \rangle$, $\langle y \rangle$, $\langle x^2 \rangle$, and $\langle y^2 \rangle$. Table A.1 contains the coordinates obtained from thirty images using each method, we see that on average there is a difference of approximately one pixel between our C program and the MATLAB program. For reference we have plotted the thirty points obtained in using each method in Figs. A.1 and A.2. We have analyzed 1500 images from a trapping experiment to compare the trap stiffnesses that would be obtained from each tracking method. We found $\langle x_C^2 \rangle = 3.370$ pixels^2, $\langle x_M^2 \rangle = 3.149$ pixels^2, $\langle y_C^2 \rangle = 3.559$ pixels^2, $\langle y_M^2 \rangle = 3.478$ pixels^2, $k_{x_C} = 0.996 \times 10^{-3}$ pN/nm, $k_{x_M} = 1.066 \times 10^{-3}$ pN/nm, $k_{y_C} = 0.996 \times 10^{-3}$ pN/nm, and $k_{y_M} = 0.965 \times 10^{-3}$ pN/nm. The difference between the trap stiffnesses are $\delta k_x = 0.70 \times 10^{-4}$ pN/nm and $\delta k_y = 0.22 \times 10^{-4}$ pN/nm, these differences are small, however in relation to the trap stiffnesses they are less than 10%. If these difference are deemed inconsequential we are free to choose which software to use. Also, if the data is to be binned and we on average we have $|p_i^C - p_i^M| < bin_w$, where p_i^C is the coordinate obtained from our C program, p_i^M is that obtained from the MATLAB software, i is either x or y, and bin_w is the width of the bin used in the analysis, we are certainly at liberty to choose either method.

Image	x_C	x_M	$\lvert\delta x_i\rvert$	y_C	y_M	$\lvert\delta y_i\rvert$
1	80.99	82.22	1.23	72.98	73.35	0.37
2	80.00	81.09	1.09	76.00	76.78	0.79
3	80.00	80.82	0.82	71.00	71.63	0.63
4	77.00	77.88	0.88	75.00	75.88	0.88
5	80.99	82.19	1.20	77.00	77.73	0.73
6	84.99	85.76	0.76	78.00	78.84	0.85
7	83.96	83.81	0.16	78.99	80.02	1.03
8	79.99	80.35	0.37	78.00	79.00	1.00
9	81.99	82.85	0.86	77.01	77.99	0.99
10	81.00	82.07	1.07	73.99	74.67	0.68
11	80.01	81.19	1.19	69.02	70.21	1.19
12	77.03	77.98	0.95	67.98	68.05	0.07
13	79.01	79.96	0.95	73.00	73.62	0.63
14	79.97	79.84	0.13	73.99	74.70	0.71
15	78.01	78.71	0.70	71.98	72.25	0.27
16	78.01	78.85	0.84	73.98	74.96	0.97
17	76.99	77.74	0.75	77.99	78.57	0.58
18	78.00	78.90	0.90	79.97	80.30	0.33
19	78.01	79.03	1.02	78.02	79.44	1.41
20	77.98	78.40	0.42	76.99	77.92	0.92
21	78.03	79.11	1.09	77.02	78.38	1.36
22	79.99	80.57	0.58	72.03	73.37	1.34
23	76.03	77.09	1.07	76.02	77.39	1.37
24	75.99	76.45	0.46	79.02	80.69	1.67
25	79.98	80.61	0.63	80.01	81.35	1.34
26	78.03	79.22	1.19	82.00	82.75	0.75
27	76.99	77.89	0.90	77.02	78.14	1.13
28	79.01	80.60	1.59	77.01	78.39	1.37
29	81.00	82.25	1.25	81.01	82.30	1.29
30	80.01	81.53	1.51	82.02	83.15	1.14

Table A.1. A comparison of the positions of thirty cropped images of a latex bead fixed to a cover-slip. We have first analyzed the images using our software written in C and second using the MATLAB software. Each value listed above is a coordinate within the analyzed image, measured in pixels, x_C and x_M correspond to the x coordinate obtained with our C program and the MATLAB software, respectively. The average deviations were $\delta x = 0.89$ pixels and $\delta y = 0.93$ pixels.

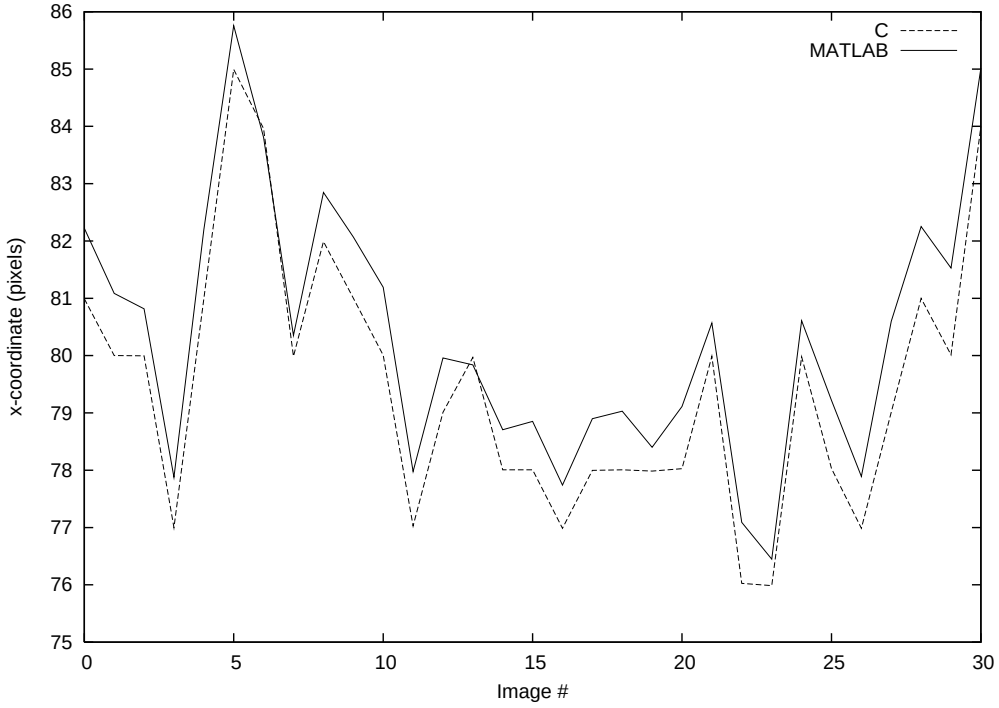

Figure A.1. A plot of the thirty x-coordinates listed in Table A.1.

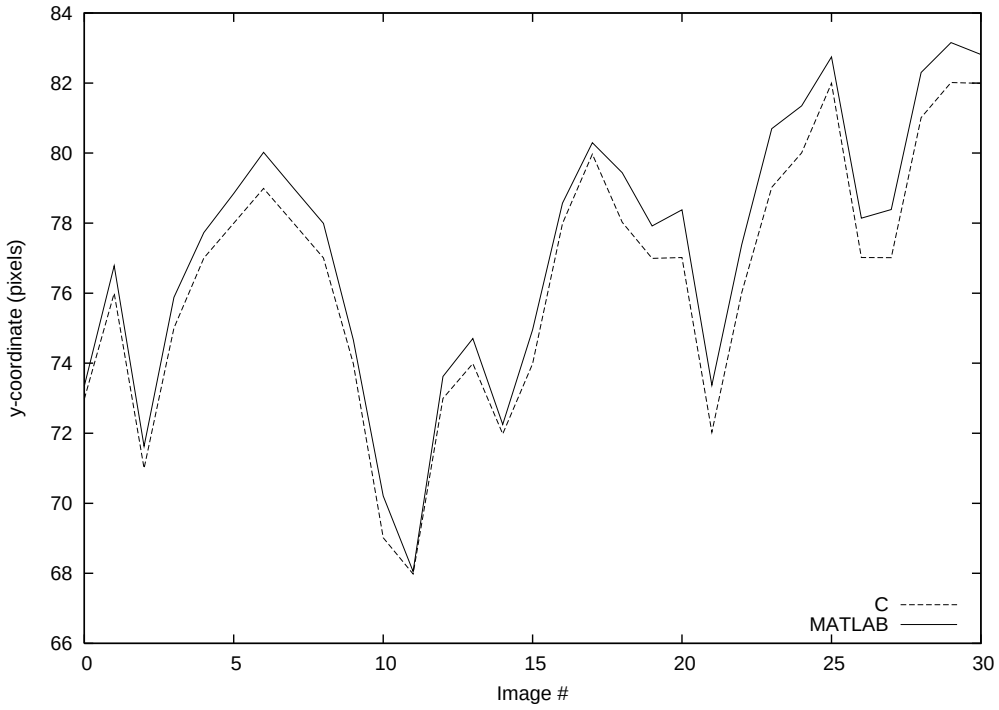

Figure A.2. A plot of the thirty y-coordinates listed in Table A.1.

Appendix B

CODE FOR ANALYSIS

stats.c

This is the program used to calculate the mean square displacements, $\langle x^2 \rangle$ and $\langle y^2 \rangle$, the probability distributions, $P(x)$ and $P(y)$, and the potentials, $U(x)$ and $U(y)$.

```
#include<stdio.h>
#include<math.h>
#include<stdlib.h>
#include"struct.h"


int main(int  argc,char *argv[])
{
typedef struct {
   double x;
   double y;
}exact_coord;

FILE *fp,*fpx,*fpy,*fpr,*fopen();
float x_avg=0,y_avg=0,y_tmp=0,x_tmp=0;
double x_min=0,x_max=0,y_min=0,y_max=0,
tmp,rsq_avg=0,r_avg=0,theta,theta_avg=0;
int frames=15000,j=0,bin=0;

double bin_w=0.5;
int bins=16;
int half_bins=(int)bins/2;

float x_dat[frames],y_dat[frames];
int n_x[bins],n_y[bins];
double P_x[bins],P_y[bins];
exact_coord *dev=NULL,dev_avg_s,dev_avg,*ukt=NULL;
```

```
printf("bins/2=%d\n",half_bins);

#if 1
/*******initialize**********/
dev=(exact_coord *) malloc(sizeof(exact_coord)*frames);
dev_avg_s.x=0;dev_avg_s.y=0;dev_avg.x=0;dev_avg.y=0;
for(j=0;j<bins;++j)
{
        P_x[j]=0;
        P_y[j]=0;
        n_x[j]=0;
        n_y[j]=0;
        }
#endif
/***calculate totals & averages***/
fpx=fopen("testing.txt","r");
//fpy=fopen("ydat_avg.txt","r");
for(j=0;j<frames;++j)
{
fscanf(fpx,"%f",&x_dat[j]);
x_tmp+=x_dat[j];
//printf("%f\n",x_dat[j]);
fscanf(fpx,"%f",&y_dat[j]);
y_tmp+=y_dat[j];
r_avg+=sqrt(pow(x_dat[j],2)+pow(y_dat[j],2));
theta_avg+=atan2(y_dat[j],x_dat[j]);
//printf("%f\n",y_dat[j]);
}
fclose(fpx);
//x_avg=x_tmp/frames;
//y_avg=y_tmp/frames;
r_avg/=frames;
theta_avg/=frames;
x_avg=r_avg*cos(theta_avg);
y_avg=r_avg*sin(theta_avg);
printf("<x>=%f\t<y>=%f\n",x_avg,y_avg);
printf("<r>=%f\t<theta>=%f\t<x>=%f\t<y>=%f\n",
r_avg,theta_avg,
r_avg*cos(theta_avg),
r_avg*sin(theta_avg));
        for(j=0;j<frames;++j)
{
dev[j].x=x_dat[j]-x_avg;
dev[j].y=y_dat[j]-y_avg;
```

```
dev_avg.x+=dev[j].x;
dev_avg.y+=dev[j].y;
dev_avg_s.x+=pow(dev[j].x,2);
dev_avg_s.y+=pow(dev[j].y,2);
rsq_avg+=pow(dev[j].x,2)+pow(dev[j].y,2);

for(bin=-half_bins;bin<=half_bins;++bin)
{
x_min=(double)((double)2*bin*bin_w-bin_w);
x_max=(double)((double)2*bin*bin_w+bin_w);
y_min=(double)(2*bin*bin_w-bin_w);
y_max=(double)(2*bin*bin_w+bin_w);
if(dev[j].x>x_min && dev[j].x<x_max)
n_x[bin+half_bins]+=1;
if(dev[j].y>y_min && dev[j].y<y_max)
n_y[bin+half_bins]+=1;
}
}
dev_avg_s.x/=frames;
dev_avg_s.y/=frames;
dev_avg.x/=frames;
dev_avg.y/=frames;
rsq_avg/=frames;

printf("<x^2>=%f\t<y^2>=%f\t<r^2>=%f\n",
dev_avg_s.x,dev_avg_s.y,rsq_avg);
printf("del_x=%f\tdel_y=%f\n",dev_avg.x,dev_avg.y);
for(j=0;j<bins;++j)
{
P_x[j]=(double)n_x[j]/frames;
P_y[j]=(double)n_y[j]/frames;
}
fp=fopen("probsavg.txt","w");
for(j=0;j<bins;++j)
fprintf(fp,"%d %f  %f  %f\t%f\t%f\n",
j-half_bins,
(j-half_bins)*2*44*bin_w,
P_x[j],P_y[j],
-log(P_x[j]),
-log(P_y[j]));
fclose(fp);
   fp=fopen("coords.txt","w");
fpx=fopen("xdat.txt","w");
fpy=fopen("ydat.txt","w");
```

```
fpr=fopen("rdat.txt","w");
for(j=0;j<frames;j++)
{
  tmp=sqrt(pow(dev[j].x*.05,2)+pow(dev[j].y*.05,2));
  theta=atan(dev[j].y/dev[j].x);
  fprintf(fp,"%f  %f  %f  %f\n",
dev[j].x*.05,dev[j].y*.05,tmp,theta);
  fprintf(fpx,"%f\n",x_dat[j]);
  fprintf(fpy,"%f\n",y_dat[j]);
  tmp=pow(x_dat[j]-x_avg,2)+pow(y_dat[j]-y_avg,2);
  tmp=sqrt(tmp);
  fprintf(fpr,"%f\n",tmp);
}
fclose(fpx);
fclose(fpy);
fclose(fp);
}
```

multi_implay.c

This is the program that was written to find the center of a round object within a specified number of sequentially ordered .png images.

```
//R. Brody.....................8/08/04.
//this will find the highest valued pixel, and ouput the
//value of each pixel in that row to a file.  call this program with
//arguments, first the .png file to be analyzed and the file
//for the data second.

#include<stdio.h>
#include<math.h>
#include"png_io.h"
#include"utils.h"
#include"struct.h"

int main(int argc,char *argv[])
{
int *data=NULL,*mod=NULL;
int width,height,i,j,n,nbright=0,row_max=0;
char file[100],filename[100];
FILE *fp,*fopen();
```

```
int l=4,approx; //approx location of the particle center
double percent=0.98;
point max,*bright;
coord app_center,*app_centers=NULL;
    exact_coord center,*centers=NULL;
int index,frames;

int test=5000;
int x_tmp=0.,y_tmp=0.;
double x_avg=0.,y_avg=0.;
int n_x[21],n_y[21];
double P_x[21],P_y[21];
coord *dev=NULL;

printf("number of frames should be:\n");
scanf("%d",&frames);
app_centers=(coord *) malloc(sizeof(coord)*frames);
centers=(exact_coord *) malloc(sizeof(exact_coord)*frames);
dev=(coord *) malloc(sizeof(coord)*frames);
for(index=0;index<frames;++index)
{
sprintf(filename,"%s%05d.png",argv[1],index);
png2data(&data,&width,&height,filename);
mod = (int *) malloc(sizeof(int)*height*width);

max=find_max(data,width,height);
printf("\nthe max pixel is at %d  with value %d\n",max.i,max.val);
nbright=n_bright(data,percent,max,width*height);
bright=frac_max(data,percent,max,width*height,nbright);

approx=approx_center(l,width,height,nbright,bright,data);
app_center=coordinates(approx,width,height);
app_centers[index]=app_center;
center=com_center(l,app_center,data,width);
centers[index]=center;
printf("center %d at x=%f  y=%f\n",
index,centers[index].x,centers[index].y);

/* the following will write file w/data from row w/highest
valued pixel in image*/
       row_max=(int)(max.i/width);
sprintf(file,"%s.txt",filename);
fp=fopen(file,"w");
for(i=row_max*width;i<width*(row_max+1);++i)
```

```
fprintf(fp,"%d  %d\n",i-row_max*width,data[i]);
fclose(fp);
}
sprintf(file,"centers.txt");
fp=fopen(file,"w");
for(j=0;j<frames;++j)
fprintf(fp,"%d\t%d\t%d\t%d\t%d\n",
j,app_centers[j].i,app_centers[j].x,
app_centers[j].y,data[app_centers[j].i]);
fclose(fp);

        sprintf(file,"x.txt");
fp=fopen(file,"w");
for(j=0;j<frames;++j)
fprintf(fp,"%f\n",centers[j].x);
fclose(fp);

        sprintf(file,"y.txt");
fp=fopen(file,"w");
for(j=0;j<frames;++j)
fprintf(fp,"%f\n",centers[j].y);
fclose(fp);
}
```

struct.h

This code contains a variety of functions that are used in both `stats.c` and `multi_implay.c`.

```
typedef struct {
int i;
int x;
int y;
} coord;

coord coordinates(int p,int width,int height)
{
coord location;

location.y = (int)p/width;
location.x = (int)p%width;//p-location.y*width;
```

```
location.i = p;

return(location);
}

typedef struct {
double x;
double y;
} exact_coord;

typedef struct {
int val;
int i;
} point;

point find_max(int *data,int width,int height)
{
int j;
point p;
p.i=0;
p.val=0;
for(j=0;j<width*height;++j)
{
   p.val=p.val>data[j]?p.val:data[j];
p.i  =p.val>data[j]?p.i:j;
}
//printf("%d   %d\n",p.i,p.val);
return(p);
}


////*calculates the number of points w/min acceptable value*////

int n_bright(int width,int height,int *data,double frac,
point max,int dim)
{
int j,n=0;
int x,y;
int min;
min=(int)(frac*max.val);
for(j=0;j<dim;++j)
{
if(data[j]>=min)
  {
```

```
    x=j%width;
 y=j/width;
 //if(x>15 && x<(width-15) && y>15 && y<(height-15)) ++n;
 ++n;
  }
}

return(n);
}

////*puts nbright into an array of point structures*////

point *frac_max(int width,int height,int *data,double percent,
point max,int dim,int nbright)
{
point *keep=NULL;
int j,min,n=0;
int x,y;
keep=(point *) malloc(sizeof(point)*nbright);
min=(int)(percent*max.val);
//printf("\nmin %d\n",min);
for(j=0;j<dim;++j)
{
if(data[j]>=min)
{
x=j%width;
y=j/width;
//if(x>15 && x<(width-15) && y>15 && y<(height-15))
//{
  keep[n].i=j;
  keep[n].val=data[j];
  ++n;
//}
}
}
return(keep);
   //free(keep);
}

int approx_center(int dim, int w,int h,int num_b,point *loc_b,
int *val)
{
int i,j,k;
int max_mass=0,approx_cen=0;
```

```
long int mass=0;
//printf("\nthe box length for mass search %d pixels\n",dim);

   for(i=0;i<num_b;++i)
   {
mass=0;
for(j=loc_b[i].i-dim*w;j<loc_b[i].i+dim*w+w;j+=w)
{
 if(j>dim)
 {
for(k=j-dim;k<j+dim+1;k++)
{
if(k>dim)mass=mass+val[k];
//printf("mass=%d\n",mass);
   //printf("%d pixel\n",k);
}
 }
}
approx_cen=max_mass>mass?approx_cen:loc_b[i].i;
max_mass=max_mass>mass?max_mass:mass;
}
return(approx_cen);
}

exact_coord com_center(int dim,coord approx,int *data,int width)
{
int x=0,y=0,i=0;
int sum_x=0,sum_y=0,sum_m=0;
exact_coord center;

for(y=approx.y-dim;y<=approx.y+dim;y++)
{
for(x=approx.x-dim;x<=approx.x+dim;x++)
{
i=y*width+x;
sum_x+=(x-approx.x)*data[i];
sum_y+=(y-approx.y)*data[i];
sum_m+=data[i];
}
}
center.x=(double)approx.x+(double)sum_x/sum_m;
center.y=(double)approx.y+(double)sum_y/sum_m;
return(center);
}
```

conditionals.c

This program will calculate the conditional probability of a spatial transition during a time interval Δt, it's spatial inverse, and the ratio of these.

```
#include<stdio.h>
#include<math.h>
#include<stdlib.h>
#include"objects.h"


#define NEW

int main(int argc,char *argv[])
{

  int  *sort=NULL;

  char  filename[100];
  int  i,j,c1,c2,
frames=10000,
bins=51,
Ebins=50,
ini,
fin,
delt;
  int  inimax=8,
         finmax=10,
binmax=15/*binmax must be greater than finmax!!!!!!!*/,
maxdelt=10;

  double bin_w=0.2,
pif,pfi,min,max,
tmp,tmp1,tmp2,tmp3,tmp4,
delta_t,delta_x,
scale,
kt=4.114/*pNnm*/,
gamma=0.0001628/*g/s*/;

  float  raw[frames],
   *window=NULL,
   *probs=NULL,
   *vx=NULL,
```

```
   *vy=NULL,
   *deltaEx=NULL,
   *deltaEy=NULL,
   *force=NULL,k;

  int   *pdeltaE=NULL;

  FILE  *fp,*fp1,*fopen();

  cond_prob *single_step=NULL,
*rats=NULL,
*s_step_noT=NULL;

  lifetime  *live=NULL;

  //*****************memory allocation!******************
  sort    = (int *)   malloc(sizeof(int)*frames);
  vx      = (float *) malloc(sizeof(float)*(frames-1));
  vy      = (float *) malloc(sizeof(float)*(frames-1));
  deltaEx = (float *) malloc(sizeof(float)*(frames-1));
  deltaEy = (float *) malloc(sizeof(float)*(frames-1));
  force   = (float *) malloc(sizeof(float)*(frames-1));
  probs   = (float *) malloc(sizeof(float)*bins);
  pdeltaE = (int *)   malloc(sizeof(int)*Ebins);
  //*****************************************************

  sprintf(filename,"%s",argv[1]);
  fp=fopen(filename,"r");
  tmp=0;
  for(i=0;i<frames;i++)
    {
   fscanf(fp,"%f",&raw[i]);
tmp+=raw[i];
 }
  tmp/=frames;
  for(i=0;i<frames;i++)raw[i]-=tmp;

  //this section sorts the particles into bins
  //and ouputs the file sorted.txt
  sort=sorting(bins,bin_w,frames,raw);
  sprintf(filename,"sorted.txt");
  fp=fopen(filename,"w");
  for(i=0;i<frames;i++)
  {
```

```
 tmp=2*bin_w*sort[i];
 fprintf(fp,"%d\t%f\n",i,tmp);
  }
  fclose(fp);
  /*****************************************************************/
  //add provision for obtaining average lifetimes: output file
  //lifetimes.txt
  //live=(lifetime *)malloc(sizeof(lifetime)*(2*inimax+1));
  live=lifetimes(sort,frames,bin_w,binmax,argv[1]);
  //for(i=0;i<(2*inimax+1);i++)printf("live[%d].life=%f\n",
  //i-inimax,live[i].life);
  /*****************************************************************/
  //get the time indep. single-step conditional probabilities

  //s_step_noT=(cond_prob *)malloc(sizeof(cond_prob)*(2*inimax+1));
  s_step_noT=step_cp(sort,frames,binmax,argv[1]);

  /*****************************************************************/


  k=0.005104;
  scale=30.3/*nm/pixel*/;
  //scale=1.0;
  //printf("k=%f\n",k);
  //maxdelt=10;
  ini=0;
  fin=1;
  rats=(cond_prob *)malloc(sizeof(cond_prob)*maxdelt);

  /********calc the prob and potential at each bin-center**********/

  probs=get_probs(bins,bin_w,frames,sort);
  for(i=0;i<bins;i++)probs[i]=-log(probs[i]);
//now probs[i] is a potential ie U[i]/kT

  /***********get deltaE's & forces & v's*************************/
  tmp=0;
  for(i=0;i<Ebins;i++)pdeltaE[i]=0;
  sprintf(filename,"deltaE_%s",argv[1]);
  fp=fopen(filename,"w");
  for(i=0;i<frames-1;i++)
  {
    deltaEx[i]=probs[sort[i+1]+bins/2]-probs[sort[i]+bins/2];
//deltaE/kT
```

```
 pdeltaE[(int)(2*deltaEx[i]+Ebins/2)]++;
 fprintf(fp,"%f\n",deltaEx[i]);
 if(sort[i+1]-sort[i]==0)force[i]=0;
 else force[i]=-kt*scale*deltaEx[i]/(sort[i+1]-sort[i]);//pN
 vx[i]=scale*(sort[i+1]-sort[i])*30;//nm/s
    //if(i<10)printf("%f\n",force[i]);
 tmp+=gamma*vx[i]-force[i];
  }
  tmp/=(frames-1);
  //printf("<gamma*v-F>=%f\n",tmp);
  fclose(fp);

  sprintf(filename,"deltaEprobs_%s",argv[1]);
  fp=fopen(filename,"w");
  for(i=0;i<Ebins/2;i++)
  {
    //printf("%d\n",i);
 if(pdeltaE[Ebins/2+i]==0 || pdeltaE[Ebins/2-(i+1)]==0)tmp=0;
 else tmp=(double)pdeltaE[Ebins/2+i]/pdeltaE[Ebins/2-(i+1)];
 //printf("+=%d\t-=%d\n",Ebins/2+i,Ebins/2-(i+1));
 fprintf(fp,"%f\t%f\t%f\n",0.5*i,exp(-0.5*i),tmp);
    //printf("%f\t%f\t%f\n",i/2,exp(-i/2),tmp);
  }
  fclose(fp);



/*************************************************************/
/*******************MAIN LOOP*********************************/
/*************************************************************/

  single_step=(cond_prob *)malloc(sizeof(cond_prob)
*maxdelt*(2*inimax+1));

  sprintf(filename,"cond_prob_%d_%s",maxdelt,argv[1]);
  fp=fopen(filename,"w");
  fprintf(fp,"#k=%f pN/nm\tmax_delta_t=%d frames\n",k,maxdelt);
  sprintf(filename,"cond_prob_all_ts_%d_%s",maxdelt,argv[1]);
  fp1=fopen(filename,"w");
  fprintf(fp1,"#k=%f pN/nm\tmax_delta_t=%d frames\n",k,maxdelt);
  for(ini=0;ini<inimax;ini++)
  {
   for(fin=ini+1;fin<=finmax;fin++)
{
```

```
delta_x=sqrt(pow(fin-ini,2))*scale;

rats=calc_ratio(sort,frames,ini,fin,maxdelt);
   tmp=0;
   for(i=0;i<maxdelt;i++)tmp+=rats[ini+inimax+i].ratio;
   tmp/=maxdelt;
   tmp1=0;
   for(i=0;i<maxdelt;i++)
tmp1+=sqrt(pow(rats[ini+inimax+i].ratio-tmp,2));
   tmp1/=maxdelt;
   //printf("av_pif/pfi=%f +/- %f\n",tmp,tmp1);
   tmp2=.5*k*(pow(2*fin*scale*bin_w,2)-pow(2*ini*scale*bin_w,2));
         tmp3=probs[bins/2+fin]-probs[bins/2+ini];
   tmp4=
//tmp3=exp(-tmp2/kt);
   //printf("theory=%f\n",tmp2);
   fprintf(fp,"%d\t%d\t%f\t%f\t%f\t%f\t%f\t%f\t%f\n",ini,fin,delta_x,
tmp2,tmp2/kt,tmp,tmp1,tmp3,exp(-tmp3));
         fprintf(fp1,"%d\t%d\t%f\t%f\t%f\t%f\t",
ini,
fin,
tmp2,
tmp2/kt,
tmp,
tmp1);
   for(i=1;i<=maxdelt;i++)
         {
          fprintf(fp1,"%f\t",rats[ini+inimax+i].ratio);
            if(i==(maxdelt-1))fprintf(fp1,"\n");
#ifdef NEW
if(ini+1==fin)
{
single_step[maxdelt*(ini+inimax)+i-1].bin=ini;
//records bin number
single_step[maxdelt*(ini+inimax)+i-1].time=i;
//records delta_t bin#, maxdelt per bin
single_step[maxdelt*(ini+inimax)+i-1].up=rats[i].up;
//P(i+1,t+delta_t|i,t)
//next line puts the fin->ini cond. prob. in
//correct place for the P(i-1|i)
if((ini+1+inimax)<(2*inimax+1))
               {single_step[maxdelt*(ini+1+inimax)+i-1].down=
rats[i].down;}
//printf("%d\t%d\t%d\t%f\t%f\n",
```

```
//ini,fin,i,rats[i].up,
//single_step[maxdelt*(ini+inimax)+i-1].up);
}
#endif
         }

   ini=-ini;fin=-fin;

rats=calc_ratio(sort,frames,ini,fin,maxdelt);
tmp=0;
for(i=0;i<maxdelt;i++)tmp+=rats[i+inimax+i].ratio;
tmp/=maxdelt;
tmp1=0;
for(i=0;i<maxdelt;i++)tmp1+=sqrt(pow(rats[i+inimax+i].ratio-tmp,2));
tmp1/=maxdelt;
tmp2=.5*k*(pow(2*fin*scale*bin_w,2)-pow(2*ini*scale*bin_w,2));
tmp3=probs[bins/2+fin]-probs[bins/2+ini];
//tmp3=exp(-tmp2/kt);
fprintf(fp,"%d\t%d\t%f\t%f\t%f\t%f\t%f\t%f\t%f\n",ini,fin,delta_x,
tmp2,tmp2/kt,tmp,tmp1,tmp3,exp(-tmp3));
   fprintf(fp1,"%d\t%d\t%f\t%f\t%f\t%f\t",
ini,
fin,
tmp2,
tmp2/kt,
tmp,
tmp1);
         for(i=1;i<=maxdelt;i++)
         {
            fprintf(fp1,"%f\t",rats[i+inimax+1].ratio);
            if(i==(maxdelt-1))fprintf(fp1,"\n");

#ifdef NEW
            if(ini-1==fin)
            {
               single_step[maxdelt*(ini+inimax)+i-1].bin=ini;
               single_step[maxdelt*(ini+inimax)+i-1].time=i;
               single_step[maxdelt*(ini+inimax)+i-1].up=
(double)rats[i].up;
               if((ini-1+inimax)>0)
{single_step[maxdelt*(ini-1+inimax)+i-1].down=
(double)rats[i].down;}
            }
#endif
```

```
        }
ini=-ini;fin=-fin;
}
  }
  //printf("out of main loop\n");
  fclose(fp);
  fclose(fp1);
  //printf("out of main loop\n");

#ifdef NEW
sprintf(filename,"cond_prob_singlesteps_%s",argv[1]);
fp=fopen(filename,"w");
fprintf(fp,"data from %s\n",argv[1]);
   //printf("starting to write last file\n");
for(i=0;i<(2*inimax+1);i++)
{
//printf("%d\n",i);
for(j=0;j<maxdelt;j++)
{
//printf("i=%d\tj=%d\n",i-inimax,j+1);
if(single_step[maxdelt*i+j].up!=0 ||
single_step[maxdelt*i+j].down!=0)
{
tmp=1/live[i].lifeu+1/live[i].lifed;
tmp=1/tmp;
fprintf(fp,"%d\t%d\t%f\t%f\t%f\t%f\t
%f\t%f\t%f\t%f\t%f\t%f\t%f\t%f\n",
i-inimax,
j+1,
single_step[maxdelt*i+j].up,
single_step[maxdelt*i+j].down,
s_step_noT[i].up,
s_step_noT[i].down,
live[i].lifeu,
live[i].lifed,
s_step_noT[i].up*exp(-(j+1)/live[i].lifeu),
s_step_noT[i].down*exp(-(j+1)/live[i].lifed),
tmp,
s_step_noT[i].up*exp(-(j+1)/tmp),
s_step_noT[i].down*exp(-(j+1)/tmp));
}
}
}
```

```
  fclose(fp);
#endif
  //free(live);
  //free(sort);
  free(vx);
  free(vy);
  free(deltaEx);
  free(deltaEy);
  free(force);
  free(probs);
  free(pdeltaE);
  free(rats);
  free(single_step);
}
```

passage_times.c

This program calculates the following transition times, mean first passage times, MFPTs, last-touch-first-touch times, LTFTTs, and mean wiggling times, MWTs.

```
#include<stdio.h>
#include<math.h>
#include<stdlib.h>
#include"objects.h"

int main(int argc,int *argv[])
{

  int *sort=NULL;

  char filename[100];

  int i,j,l,c1,c2,
frames=15000,bins=25,
ini,fin,delt,
maxdelt,counters[4],
      sortmax=0,
inimax=0,
finmax=0;
```

```
  double pif,pfi,
min,max,
bin_w=0.25,
tmp,tmp1,tmp2,tmp3,
delta_t,delta_x,
scale,
kt=4.114,
gamma=0.0001628/*g/s*/;

  float raw[frames],
  *window=NULL,
  *rats=NULL,k,
  *mfpt,
  *P=NULL,
        pass[4],
  boltzi,
  boltzf;


  /*i intend that pass[0] will represent mfpt from
x_i to x_f, pass[1] is the reverse,
  and pass[2] and pass[3] will be ftlt and ltft*/

  FILE *fp,
 *fp1,
 *fopen();

  printf("frames=???\tk=???\n");
  scanf("%d%f",&frames,&k);
  sort=(int *) malloc(sizeof(int)*frames);

  sprintf(filename,"%s",argv[1]);
  fp=fopen(filename,"r");

  tmp=0;
  for(i=0;i<frames;i++)
    {
   fscanf(fp,"%f",&raw[i]);
tmp+=raw[i];
 }
  fclose(fp);

  tmp/=frames;
  for(i=0;i<frames;i++)raw[i]-=tmp;
```

```
  sort=sorting(bins,bin_w,frames,raw);
  sprintf(filename,"sorted.txt");
  fp=fopen(filename,"w");

  P=(float *)malloc(sizeof(float)*bins);
  P=get_probs(bins,bin_w,frames,sort);

  for(i=0;i<frames;i++)
  {
 tmp=2*bin_w*sort[i];
 fprintf(fp,"%d\t%f\n",i,tmp);
  }
  fclose(fp);
  for(i=0;i<frames;i++)sortmax=sortmax>sqrt(pow(sort[i],2))?
sortmax:sort[i];
  printf("sortmax=%d\t",sortmax);
  for(l=sortmax;l>0;l--)
  {
 tmp=0;
 for(i=0;i<frames;i++){if(sort[i]==sortmax)tmp++;}
 if(tmp<0.03*frames)sortmax--;
 if(tmp>=0.03*frames)l=0;
  }
  printf("now sortmax=%d\n",sortmax);
  inimax=sortmax;
  finmax=sortmax;
  mfpt=(float *)malloc(sizeof(float)*(2*inimax+1));
  for(i=0;i<=2*inimax;i++)mfpt[i]=0;
  scale=30.3/*nm/pixel*/;
  printf("k=%f\n",k);
  maxdelt=2;
  ini=0;
  fin=1;
  rats=(float *)malloc(sizeof(float)*maxdelt);

  sprintf(filename,"passage_ratios_%s",argv[1]);
  fp1=fopen(filename,"w");
  sprintf(filename,"passage_times_%s",argv[1]);
  fp=fopen(filename,"w");
  fprintf(fp,"#k=%f pN/nm\t%s\n",k,argv[1]);
  fprintf(fp1,"#k=%f pN/nm\t%s\n",k,argv[1]);
  for(ini=-finmax;ini<=inimax;ini++)
  {
```

```
for(fin=ini+1;fin<=finmax;fin++)
{
delta_x=sqrt(pow(fin-ini,2))*scale;
for(i=0;i<4;i++){pass[i]=0;counters[i]=0;}
      for(j=0;j<frames;j++)
{
   if(sort[j]==ini)
   {
 tmp1=0;
 for(l=j+1;l<frames-1;l++)
 {
if(sort[l]<=ini)tmp1=l;
if(sort[l]>=fin)
{
  counters[0]+=1;
                counters[2]+=1;
  pass[0]+=(l-j);
  if(tmp1>0)pass[2]+=(l-tmp1);
  else pass[2]+=(l-j);
  l+=frames;
}
 }
        }
#if 1
else if(sort[j]==fin)
{
    tmp2=0;
for(l=j+1;l<frames-1;l++)
{
if(sort[l]>=fin)tmp2=l;
if(sort[l]<=ini)
{
counters[1]++;
counters[3]++;
   pass[1]+=l-j;
   if(tmp2>0)pass[3]+=(l-tmp2);
else pass[3]+=(l-j);
l+=frames;
}
}
}
#endif
}
for(i=0;i<4;i++)
```

```
      {
       pass[i]/=30;
pass[i]=(float)(pass[i]/counters[i]);
         if(counters[i]==0)pass[i]=0;
      }
if(pass[0]==0)tmp1=0;
else tmp1=pass[1]/pass[0];
if(pass[2]==0)tmp2=0;
else tmp2=pass[3]/pass[2];
tmp=.5*k*(pow(2*fin*scale*bin_w,2)-
pow(2*ini*scale*bin_w,2))/kt;
boltzf=exp(-0.5*k*pow(fin,2)/kt);
boltzi=exp(-0.5*k*pow(ini,2)/kt);
mfpt[ini+inimax]+=pass[0]*boltzf;
mfpt[fin+inimax]+=pass[1]*boltzi;
tmp3=gamma*pow(scale*(ini-fin),2);
//tmp3/=(kt*pow(log(P[ini+bins/2]/P[fin+bins/2]),2));
if(tmp==0)tmp3=0;
else tmp3/=(kt*pow(tmp,2));
fprintf(fp,"%d\t%d\t%f\t%f\t%f\t%f\t%f\t%f\t
%f\t%f\t%d\t%f\t%f\n",
ini,fin,ini*scale,fin*scale,tmp,
pass[0]*boltzf,pass[1]*boltzi,
pass[2],pass[3],tmp3,ini+bins/2,
P[ini+bins/2],P[fin+bins/2]);
         fprintf(fp1,"%d\t%d\t%f\t%f\t%f\t%f\t%f\n",
ini,fin,ini*scale,fin*scale,
tmp,tmp1,tmp2);
}
  }
  fclose(fp);
  fclose(fp1);
  sprintf(filename,"passage_weights_%s",argv[1]);
  fp=fopen(filename,"w");
  fprintf(fp,"#k=%f pn/NM\t%s\n",k,argv[1]);
  for(i=0;i<=2*inimax;i++)
  {
 mfpt[i]/=2*inimax;
 fprintf(fp,"%d\t%f\t%f\n",
i-inimax,scale*(i-inimax),mfpt[i]);
  }

  fclose(fp);
}
```

objects.h

This is a suite of functions and definitions used by `conditionals.c` and `ltft.c`.

```
typedef struct{
   int cnt;
   int u;
   int d;
}counters;

typedef struct{
   int bin;
   int time;
   float ratio;
   double up;
   double down;
}cond_prob;

typedef struct{
   int   pos;//position
   float u1;//cond. prob to move up 1 bin, x->x+1
   float u;//cond. prob to move such that x_i<x_f
   float d1;//down 1
   float d;//down
}extra_cond_prob;

typedef struct{
   extra_cond_prob x;
   extra_cond_prob y;
   float t;
}condprob2d;


typedef struct{
   float life;
   float lifeu;
   float lifed;
   float u;
   float d;
}lifetime;
```

```
typedef struct{
int x;
int y;
}pstep;

int *sorting(int bins,double bin_w,int frames,float *data)
{
  int i,j,*sort=NULL;
  double min,max,tmp;

  sort=(int *)malloc(sizeof(int)*frames);
  for(i=-bins/2;i<=bins/2;i++)
  {
 min=(double)((double)2*i*bin_w-bin_w);
 max=(double)((double)2*i*bin_w+bin_w);
 tmp=0;
 for(j=0;j<frames;j++)
 {
if (data[j]>=min && data[j]<max)
{
sort[j]=i;
tmp++;
}
 }
 if(tmp>0)printf("%f in bin %d\n",tmp,i);
}
printf("returning sort now\n");
return(sort);
}

float *get_probs(int bins,double bin_w,int frames,int *sort)
{
int i,j,k;
   float tmp;
   float *probs=NULL;

probs=(float *)malloc(sizeof(float)*(bins+1));

for(i=-bins/2;i<=bins/2;i++)
{
k=i+bins/2;
probs[k]=0;
tmp=0;
```

```
for(j=0;j<frames;j++)if(sort[j]==i)probs[k]++;
probs[k]/=frames;
}
return(probs);
}

int *get_counts2d(int bins,int frames,int *sortx,int *sorty)
{
int i,j,k;
int tmp,tmp1,tmp2;

int *count2d=NULL;
   char file[100];
FILE *fpc=NULL,*fopen();

count2d = (int *)malloc(sizeof(int)*bins*bins);
sprintf(file,"counts2d.txt");
fpc=fopen("counts2d.txt","w");
printf("allocated memory and filename\n");
for(i=0;i<frames;i++)count2d[i]=0;
for(i=0;i<frames;i++)
{
tmp=(int)((sorty[i]+bins/2)*bins+(sortx[i]+bins/2));
count2d[tmp]++;
}
for(i=0;i<bins*bins;i++)
{
fprintf(fpc,"%d\t",count2d[i]);
if((i+1)%(bins)==0 && i>0)fprintf(fpc,"\n");
}
printf("now will try to close the file");
fclose(fpc);
return(count2d);
}

condprob2d xycondprobs(int bin,int bins,
int *sortx,int *sorty,int frames)
{
int i,
 j,
 k,
 x,
    y;
```

```
counters cx,
cy;

condprob2d xy2d;

x=bin%bins-bins/2;
xy2d.x.pos=x;
y=bin/bins-bins/2;
xy2d.y.pos=y;

for(i=0;i<frames-1;i++)
{
     if(i==0 || sortx[i-1]!=sortx[i] || sorty[i-1]!=sorty[i])
{
cx.cnt++;//number of unique times in the bin
cy.cnt++;//same as above
k=0;
while(sortx[i+k]==x && sorty[i+k]==y)
{
k++;//number of times in the bin this time around
if(x>0)
{
if(sortx[i+k]==x+1 && sorty[i+k]==y  )xy2d.x.u1++;
if(sortx[i+k]>x    && sorty[i+k]==y  )xy2d.x.u++;
         if(sortx[i+k]==x-1 && sorty[i+k]==y  )xy2d.x.d1++;
         if(sortx[i+k]<x    && sorty[i+k]==y  )xy2d.x.d++;
}
if(y>0)
{
if(sortx[i+k]==x   && sorty[i+k]==y+1)xy2d.y.u1++;
if(sortx[i+k]==x   && sorty[i+k]>y   )xy2d.y.u++;
         if(sortx[i+k]==x   && sorty[i+k]==y-1)xy2d.y.d1++;
         if(sortx[i+k]==x   && sorty[i+k]<y   )xy2d.y.d++;
}
         if(x<0)
         {
            if(sortx[i+k]==x-1 && sorty[i+k]==y  )xy2d.x.u1++;
            if(sortx[i+k]<x    && sorty[i+k]==y  )xy2d.x.u++;
            if(sortx[i+k]==x+1 && sorty[i+k]==y  )xy2d.x.d1++;
            if(sortx[i+k]>x    && sorty[i+k]==y  )xy2d.x.d++;
         }
         if(y<0)
         {
            if(sortx[i+k]==x   && sorty[i+k]==y+1)xy2d.y.u1++;
```

```
            if(sortx[i+k]==x   && sorty[i+k]<y   )xy2d.y.u++;
            if(sortx[i+k]==x   && sorty[i+k]==y+1)xy2d.y.d1++;
            if(sortx[i+k]==x   && sorty[i+k]>y   )xy2d.y.d++;
         }
}
xy2d.t+=k+1;//to get the average time spent in the bin
}
}
if(cx.cnt>0)
{
xy2d.x.u1/=cx.cnt;
xy2d.x.u /=cx.cnt;
xy2d.x.d1/=cx.cnt;
xy2d.x.d /=cx.cnt;
   xy2d.y.u1/=cy.cnt;
   xy2d.y.u /=cy.cnt;
   xy2d.y.d1/=cy.cnt;
   xy2d.y.d /=cy.cnt;
xy2d.t/=cx.cnt;
}
return(xy2d);
}

cond_prob *calc_ratio
(int *sort,int frames,int ini,int fin,int max_delta_t)
{
int i,delta_t,c1=0,c2=0;
float pif,pfi;

cond_prob *ratios=NULL;

ratios=(cond_prob *)malloc(sizeof(cond_prob)*max_delta_t);
for(delta_t=1;delta_t<=max_delta_t;delta_t++)
{
c1=0;c2=0;pif=0;pfi=0;
for(i=0;i<frames-delta_t;i++)
{
if(sort[i]==ini)
{
c1++;
if(sort[i+delta_t]==fin)pif++;
}
if(sort[i]==fin)
{
```

```
c2++;
if(sort[i+delta_t]==ini)pfi++;
}
}
//printf("goodbye\n");
pif/=c1;
pfi/=c2;
ratios[delta_t-1].up=pif;
ratios[delta_t-1].down=pfi;
//printf("%f %f\n",pif,pfi);
if(pfi==0)ratios[delta_t-1].ratio=0;
else ratios[delta_t-1].ratio=pif/pfi;
//printf("%f\n",ratio);
}
return(ratios);
}


float *calc_ratio_2d(int *sortx,int *sorty,int frames,
int xini,int xfin,int yini,int yfin,
int max_delta_t)
{
  int i,delta_t,c1=0,c2=0;
  float *ratios=NULL,pif,pfi;
  ratios=(float *)malloc(sizeof(float)*max_delta_t);
  for(delta_t=1;delta_t<=max_delta_t;delta_t++)
  {
    c1=0;c2=0;pif=0;pfi=0;
    for(i=0;i<frames-delta_t;i++)
    {
      if(sortx[i]==xini && sorty[i]==yini)
      {
        c1++;
        if(sortx[i+delta_t]==xfin && sorty[i+delta_t]==yfin)pif++;
      }
      if(sortx[i]==xfin && sorty[i]==yfin)
      {
        c2++;
        if(sortx[i+delta_t]==xini && sorty[i+delta_t]==yini)pfi++;
      }
  }
  //printf("goodbye\n");
  pif/=c1;
  pfi/=c2;
```

```
  //printf("%f%f\n",pif,pfi);
  if(pfi==0)ratios[delta_t-1]=0;
  else ratios[delta_t-1]=pif/pfi;
  //printf("%f\n",ratio);
  }
  return(ratios);
}


float getcondprob2d(int bins,int i,int j,int delta_t,
int *sortx,int *sorty,int frames)
{
int k,l,m;
int c1=0,
    c2=0,
 xi,xj,yi,yj;

float tmp,condprob;

xi=i%bins-bins/2;
xj=j%bins-bins/2;
yi=i/bins-bins/2;
yj=j/bins-bins/2;

for(k=0;k<frames;k++)
{
if(sortx[k]==xi && sorty[k]==yi)
{
c1++;
if(sortx[k+delta_t]==xj && sorty[k+delta_t]==yj)c2++;
}
}
if(c1==0)tmp=0;
else tmp=c2/c1;
return(tmp);
}


int bin_find(double delta_u,int bins,double bin_width)
{
  int i,j,b,n;
  for(i=0;i<bins;i++)
  {
 b=(int)(i-bins/2);
```

```
 if (delta_u>(bin_width*(b-0.5)) && delta_u<=(bin_width*(b+0.5)))
 {
//printf("\t\t%d\n",i);
   n=i;
i=bins;
 }
  }
  return(n);
}

lifetime *lifetimes(int *sort,int frames,
double bin_w,int binmax,char *txt)
{
int i,j,k,tmp,t_bin;
int *times=NULL;//number of time-steps in a given bin
int *counts=NULL;//number of unique time-spans in a given bin
lifetime *lives=NULL; //avg. lifetime in a given bin

char filename[100];
FILE *fp,*fopen();

/**allocate memory & initialize to zero**********************/
counts = (int *)malloc(sizeof(int)* (2*binmax+1));
times  = (int *)malloc(sizeof(int)* (2*binmax+1));
   lives  = (lifetime *) malloc(sizeof(lifetime) * (2*binmax+1));

for(i=0;i<2*binmax+1;i++)
{
counts[i]=0;
times[i]=0;
lives[i].life=0;
lives[i].lifeu=0;
lives[i].lifed=0;
lives[i].u=0;
lives[i].d=0;
}
/**main loop, avoids double counting by checking that pos. in */
/**prev. bin is not same as ith*******************************/
for(i=0;i<frames;i++)
{
    if((sort[i]+binmax)<(2*binmax+1))
     {
if(i==0 || sort[i]!=sort[i-1])
{
```

```
t_bin=0;
counts[sort[i]+binmax]++;
//add 1 for each unique trip to a bin
t_bin++;
//#t_steps in bin bef transition

j=1;
/*following accounts for bin lifetime of 1 t_step*************/
if(sort[i]!=sort[i+j])
{
if(sort[i]<0)
   {
  if(sort[i]>sort[i+j+1])lives[sort[i]+binmax].u++;
//# of up trans.
   if(sort[i]<sort[i+j+1])lives[sort[i]+binmax].d++;
//# of down trans.
   if(sort[i]>sort[i+j+1])lives[sort[i]+binmax].lifeu+=t_bin;
//t bf u trans
   if(sort[i]<sort[i+j+1])lives[sort[i]+binmax].lifed+=t_bin;
//t bf d trans
   }

   if(sort[i]>0)
   {
  if(sort[i]>sort[i+j+1])lives[sort[i]+binmax].d++;
//# of down trans
   if(sort[i]<sort[i+j+1])lives[sort[i]+binmax].u++;
//# of up trans
   if(sort[i]>sort[i+j+1])lives[sort[i]+binmax].lifed+=t_bin;
//t bf d trans
   if(sort[i]<sort[i+j+1])lives[sort[i]+binmax].lifeu+=t_bin;
//t bf u trans
    }
}
/***********************************************/
/*if lifetime is greater than one t_step*******************/
while(sort[i]==sort[i+j])
{
t_bin++;
   j++;
}
/*****************************************************/
    if(t_bin!=j)printf("t_bin=%d\t&\tj=%d\n",t_bin,j);
    if(sort[i]<0)
```

```
    {
      if(sort[i]>sort[i+j])lives[sort[i]+binmax].u++;
//# of up trans.
      if(sort[i]<sort[i+j])lives[sort[i]+binmax].d++;
//# of down trans.
      if(sort[i]>sort[i+j])lives[sort[i]+binmax].lifeu+=j;
//t spent bef. u trans.
      if(sort[i]<sort[i+j])lives[sort[i]+binmax].lifed+=j;
//t spent bef. d trans.
     }

    if(sort[i]>0)
    {
   if(sort[i]>sort[i+j])lives[sort[i]+binmax].d++;
//# of down trans.
    if(sort[i]<sort[i+j])lives[sort[i]+binmax].u++;
//# of up trans.
    if(sort[i]>sort[i+j])lives[sort[i]+binmax].lifed+=j;
//t spent bef. d trans.
    if(sort[i]<sort[i+j])lives[sort[i]+binmax].lifeu+=j;
//t spent bef. u trans.
    }

times[sort[i]+binmax]+=j;
}
     }
}

/*calculate avg lifetimes & prob's of each transition*/
/*****and write to file******/
sprintf(filename,"lifetimes_%.2f-%s",bin_w,txt);
fp=fopen(filename,"w");
for(i=0;i<(2*binmax+1);i++)
{if(i!=binmax){
lives[i].life=counts[i]>0  ? (float)times[i]/counts[i]:0;
lives[i].lifeu=lives[i].u>0 ? lives[i].lifeu/lives[i].u:0;
lives[i].lifed=lives[i].d>0 ? lives[i].lifed/lives[i].d:0;
lives[i].u  =counts[i]>0  ? (float)lives[i].u/counts[i]:0;
      lives[i].d  =counts[i]>0  ? (float)lives[i].d/counts[i]:0;
//printf("times[%d]=%d\n",i-binmax,times[i]);
fprintf(fp,"%d\t%f\t%f\t%f\t%f\t%f\t%d\t%d\n",
i-binmax,
lives[i].life,
lives[i].lifeu,
```

```
lives[i].lifed,
lives[i].u,
lives[i].d,
times[i],
counts[i]);
}
}
/*****************************************************************/
fclose(fp);

return(lives);
}


#if 1

cond_prob *step_cp(int *data,int frames,int max,char *out)
{
int i,j,k;
double tmp1,tmp2;

counters *c=NULL;
cond_prob *P=NULL;


char file[100];
FILE *fp,*fopen();
c=(counters *)malloc(sizeof(counters)*(2*max+1));
P=(cond_prob *)malloc(sizeof(cond_prob)*(2*max+1));

for(i=0;i<2*max+1;i++)
{
P[i].up=0;
P[i].down=0;
c[i].cnt=0;
c[i].u=0;
c[i].d=0;
}

for(i=0;i<frames-1;i++)
{
tmp1=sqrt(pow(data[i],2));
if(tmp1<=max)
{
```

```
if(i==0 || data[i]!=data[i-1])
{
c[data[i]+max].cnt++;
j=0;
while(data[i]==data[i+j])
{
j++;
if(data[i]<0 && data[i]>data[i+j])c[data[i]+max].u++;
if(data[i]<0 && data[i]<data[i+j])c[data[i]+max].d++;
if(data[i]>0 && data[i]>data[i+j])c[data[i]+max].d++;
if(data[i]>0 && data[i]<data[i+j])c[data[i]+max].u++;
}
}
}
}
//for(i=0;i<2*max+1;i++)printf("%d\t%d\n",i-max,c[i].cnt);
sprintf(file,"one-step_cond_prob-%s",out);
fp=fopen(file,"w");
for(i=0;i<(2*max+1);i++)
{if(i!=max){
P[i].up=c[i].cnt>0?(double)c[i].u/c[i].cnt:0;
P[i].down=c[i].cnt>0?(double)c[i].d/c[i].cnt:0;
fprintf(fp,"%d\t%f\t%f\n",i-max,P[i].up,P[i].down);
}}
fclose(fp);
return(P);
}
#endif


double p2pLTFT(int ini,int fin,int *data,int frames)
{
int i,j,k,
cnt=0,
tmp;

double ltft=0;
   //printf("starting p2p loop\n");
for(i=0;i<frames-1;i++)
{
if(data[i]==ini)
{
cnt++;
//printf("cnt=%d\n",cnt);
```

```
j=1;
while(i+j<frames && data[i+j]!=fin)
{
if(i+j<frames && data[i+j]==ini)
{
i+=j;
j=1;
}
else j++;
}
}
//printf("j=%d\t",j);
ltft+=(double)j;
i+=j;
}
//printf("j=%d\tltft=%f\tcnt=%d\n",j,ltft,cnt);
ltft=cnt==0?0:ltft/cnt;
return(ltft);
}

int psteps(int *sortx,int *sorty,int frames)
{
int i,j,k;
int px[4],py[4],
 stepx,stepy,abs_stepx,abs_stepy;

pstep *Pstep=NULL;

Pstep=(pstep *)malloc(sizeof(pstep)*frames);
for(i=0;i<frames;i++)
{
Pstep[i].x=0;
Pstep[i].y=0;
}
printf("frames=%d\n",frames);
for(i=0;i<4;i++)
{
px[i]=0;
py[i]=0;
}

for(i=0;i<frames-1;i++)
{
stepx=sortx[i+1]-sortx[i];
```

```
abs_stepx=stepx<0?-stepx:stepx;
stepy=sorty[i+1]-sorty[i];
abs_stepy=stepy<0?-stepy:stepy;
Pstep[i].x=stepx;
Pstep[i].y=stepy;
if(abs_stepx<4)px[abs_stepx]++;
if(abs_stepy<4)py[abs_stepy]++;
}
for(i=0;i<4;i++)
{
printf("px[%d]=%d\tpy[%d]=%d\n",i,px[i],i,py[i]);
}

return(0);

}
```

brownies.c

This is our code written to simulate our experiments.

```
#include<stdio.h>
#include<stdlib.h>
#include<math.h>
#include<time.h>

#define pi M_PI

int main()
{
FILE *fp0,*fp,*fp1,*fp1a,*fp2,*fopen();
char file0[100],file[100],file1[100],file1a[100],file2[100];
int tmax=2000000,i,srate,fskip;

double xo,vo,delta_t,r1,r2,s,theta,n1,n2;
double tau,c,k,kb,T,kt,gamma,m,mu,sigma;
double tmp,tmp1,tmp2,delta_t_s;
double *v=NULL,*x=NULL,*y=NULL;

   time_t t1;
struct tm *tiempo;
```

```
gamma=0.00000817;//.000008904;
kt=4.114;
k=0.005;

delta_t=0.00001;
srate=30;
delta_t_s=(double)1/srate;
fskip=delta_t_s/delta_t;
printf("delta_t_s=%f\tframeskip=%d\n",delta_t_s,fskip);
xo=0;
vo=0;
tau=1;
c=1;//pow(10,6);
v=(double *)malloc(sizeof(double)*tmax);
x=(double *)malloc(sizeof(double)*tmax);
   y=(double *)malloc(sizeof(double)*tmax);
time(&t1);
tiempo=localtime(&t1);
srandom(tiempo->tm_sec*tiempo->tm_min);

mu=exp(-delta_t/tau);
sigma=sqrt(c*tau/2*(1-exp(-2*delta_t/tau)));

x[0]=0;
y[0]=0;
v[0]=0;
tmp1=sqrt(2*kt/gamma);
tmp2=sqrt(delta_t);
#if 1
for(i=1;i<tmax;i++)
{
r1=(double)rand()/RAND_MAX;
r2=(double)rand()/RAND_MAX;
s=sqrt(2*log(1/r1));
theta=2*pi*r2;
n1=s*cos(theta);
//r1=(double)rand()/RAND_MAX;r2=(double)rand()/RAND_MAX;
//s=sqrt(2*log(1/r1));
//theta=2*pi*r2;
n2=s*sin(theta);
//x[i]=x[i-1]+v[i-1]*delta_t;
//v[i]=mu*v[i-1]+sigma*n1;
x[i]=x[i-1]-k/gamma*x[i-1]*delta_t+tmp1*n2*tmp2;
y[i]=y[i-1]-k/gamma*y[i-1]*delta_t+tmp1*n1*tmp2;
```

```
}
   sprintf(file0,"x_all.txt");
fp0=fopen(file0,"w");
sprintf(file,"xdat.txt");
   fp=fopen(file,"w");
sprintf(file1,"x.txt");
fp1=fopen(file1,"w");
sprintf(file1a,"x_compare.txt");
fp1a=fopen(file1a,"w");
sprintf(file2,"y.txt");
fp2=fopen(file2,"w");
//printf("got here\n");
for(i=0;i<tmax;i++)
{
fprintf(fp0,"%f\n",x[i]);
   fprintf(fp,"%f\t%f\t%f\n",delta_t*i,v[i],x[i]);
      tmp1=i%fskip;
if(tmp1==0)
{
fprintf(fp1,"%f\n",x[i]);
fprintf(fp1a,"%f\t%f\n",delta_t*i,x[i]);
fprintf(fp2,"%f\n",y[i]);
}
}
fclose(fp0);fclose(fp);fclose(fp1);fclose(fp2);
#endif
   tmp=0;
tmp1=0;
for(i=0;i<tmax;i++/*=fskip*/){
r1=(double)rand()/RAND_MAX;r2=(double)rand()/RAND_MAX;
s=sqrt(2*log(1/r1));
theta=2*pi*r2;
n1=s*cos(theta);
n2=s*sin(theta);
tmp+=n1;
tmp1+=n1*n1;}
tmp/=(tmax/*/25*/);
tmp1/=(tmax/*/25*/);
printf("<n1^2>=%f\t<n1>=%f\tstdev{n1}=%f\n",tmp1,tmp,tmp1-tmp*tmp);}
```

BIOGRAPHY OF THE AUTHOR

On the afternoon of November 6, 1979 Ross Paul Brody was born to Sue Jean Brody and Eliott Stuart Brody in Portland, Oregon. It is said that from an early age he maintained a special fondness for mashed potatoes and most things round. He is also perpetually amazed by snowflakes and enamored of toys.

A bonafide web-foot, Ross grew up in the Willamette valley and graduated from Silverton Union High School in 1997. He relocated to Corvallis to attend Oregon State University and in June 2001 he obtained the degree of Bachelor of Science in Physics. He then obtained a driver's license, bought a truck, insured it, and drove to Orono, Maine to attend The University of Maine. Upon completion of his degree, Ross will pedal a bicycle across the United States of America. Ross is a candidate for the Doctor of Philosophy degree in Physics from The University of Maine in May 2008.