Connecting Brownian Motion and Partial Differential Equations with Applications in Statistical and Quantum Mechanics

By: Zachary Alan Selk

Advisor: Dirk Deckert, PhD

Senior thesis

Submitted in partial satisfaction of the requirements for Highest Honors for the degree of

BACHELOR OF SCIENCE

in

MATHEMATICS

in the

COLLEGE OF LETTERS AND SCIENCE

of the

UNIVERSITY OF CALIFORNIA,

DAVIS

2014

Abstract

In this paper, I review the link between stochastic processes and partial differential equations. In particular, I demonstrate how the heat and Schrödinger equations can be understood in terms of Brownian motion. This connection is demonstrated through Feynman-Kac formulas. I give a physical intuition why one should expect the heat equation should be understood in terms of Brownian motion by arguments given by Einstein and Smoluchowski.

Acknowledgment: I would like to thank Dirk Deckert for all of his encouragement, knowledge and advice.

Contents

Chapter 1

Basics from probability theory

Brownian Motion was first studied by Robert Brown in 1828 [1]. He encountered it while studying the erratic motion of pollen suspended in water. He didn't know why the pollen moved like it did. It wasn't until 1905 that Albert Einstein proposed a solution. He proposed that Brownian motion was the motion of the pollen particle as it randomly collided with water molecules. In his paper, he writes "In this paper it will be shown that according to the molecular-kinetic theory of heat, bodies of microscopically-visible size suspended in a liquid will perform movements of such magnitude that they can be easily observed in a microscope, on account of the molecular motions of heat" [5].

However before we start, we must first define Brownian Motion rigorously. To do this, we briefly review the basics of probability theory. This includes a lot of definitions, but if the reader can get past the formality, the picture is quite nice and unified.

Definition 1. *A* ***Sigma Algebra****,* Σ *of some universal set,* Ω *is a collection of subsets of* Ω *satisfying:*

1. $A \in \Sigma \implies A^c \in \Sigma$
2. $A_i \in \Sigma \implies \bigcup_{i=1}^{\infty} A_i \in \Sigma$ *for any countable collection of subsets*
3. $\Omega \in \Sigma$

Definition 2. *The* ***Borel Sets*** *of the Real Numbers,* $\mathscr{B}(\mathbb{R})$ *is the smallest* Σ*-algebra that contain all the open subsets of* $\mathbb{R}$*.*

More generally, given any topological space, $(X, \mathscr{T})$, we may define the Borel Sets in a similar fashion. The Borel Sets are nice because it gives us a connection between Probability Theory and Topology.

Definition 3. *A **Measurable Space** is some set,* Ω *along with a* Σ*-algebra of* Ω*, written in an ordered pair:* (Ω, Σ)*.*

Definition 4. *Given a measurable space,* (Ω, Σ)*, a **Measure** is a function* $\mu\colon \Sigma \to [0, \infty]$ *satisfying:*

1. $\mu(\emptyset) = 0$

2. $\mu(\bigcup\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} \mu(A_i)$ *if* A_i *are disjoint and countable*

Definition 5. *A **measure space** is an ordered triple* (Ω, Σ, μ)*:*

1. Ω *is some set*

2. Σ *is a* Σ*-algebra*

3. μ *is a measure*

Note, the reason why we need to consider the Borel sets rather than just "all subsets" is that some subsets of the Real Numbers turn out to be not measurable (with respect to the Lebesgue, or standard, measure)! An example of non measurable sets are the **Vitali Sets**. For more information, see [1].

Definition 6. *Given a measurable space,* (Ω, Σ)*, a **Probability Measure** is a measure,* $P\colon \Sigma \to [0, 1]$ *with the additional property that* $P(\Omega) = 1$*.*

Definition 7. *A **Probability Space** is a measure space,* (Ω, Σ, P) *where* P *is a probability measure. In this case, we shall refer to:*

1. Ω *as the set of outcomes*

2. Σ *as the set of all events*

3. $P(A)$ *is the probability of event* A

Definition 8. *Given a measure space,* (Ω, Σ, μ)*, we say a property,* $Q = Q(\omega)$*, for* $\omega \in \Omega$*, holds **Almost Everywhere** if* $\mu(S) = 0$*, where* $S = \{\omega \in \Omega | Q(\omega) \text{is not true}\}$*, should* $S \in \Sigma$*.*

If (Ω, Σ, P) is a probability space, then we say Q holds **Almost Surely**

Definition 9. *Given a probability space,* (Ω, Σ, μ) *and a measurable space,* (E, F)*, a **random variable*** $X(\omega)$ *is a function* $X : \Omega \to E$ *such that for every set* $B \in F$ *, we have that* $X^{-1}(B) = \{\omega \in \Omega | X(\omega) \in B\} \in \Sigma$*.*

Definition 10. *If we have a probability space,* (Ω, Σ, μ)*, a measurable space,* (E, F)*, and a random variable* $X : \Omega \to E$*, we may define the **Pushforward Measure** on* (E, F) *as* $P_X(B) = \mu(X^{-1}(B))$*, for all* $B \in F$*. This makes a probability space,* (E, F, P)*.*

Definition 11. *Given a probability space, (Ω, Σ, μ), and a measurable space, (E, F), a **Stochastic Process**, X_t is a collection of random variables $X_t \colon \Omega \to E$ indexed by a totally ordered indexing set, T.*

Definition 12. *Given a probability space, (Ω, Σ, μ), the measurable space, $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, and a random variable, $X \colon \Omega \to \mathbb{R}$, we may define the **Distribution Function** as:*

$$F_X(x) = \mu(X \leq x) = \mu\left(\{\omega \in \Omega \colon X(\omega) \leq x\}\right)$$

Definition 13. *If X has a distribution function that is of the form:*

$$F_X(x) = \int_{-\infty}^{x} \rho(t)dt$$

*then we call $\rho(t)$ the **Probability Density Function**, provided such a $\rho(t)$ exists.*

The integral in this case is the **Lebesgue Integral** [1]

Definition 14. *Given a probability space, (Ω, Σ, μ) and a random variable, $X(\omega) \colon \Omega \to \mathbb{R}$ is a $(\Omega, \Sigma) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ measurable random variable, we define the **Expectation Value**, $\mathbb{E}(X(\omega))$ as $\mathbb{E}(X) = \int_\Omega X(\omega)d\mu(\omega)$, where $d\mu(\omega)$ denotes integrating with respect to the probability measure, μ.*

Note, the expectation value can sometimes be referred to as the **Mean**.

Definition 15. *Let (Ω, Σ) a measurable space with two measures, μ and ν. We say μ is **Absolutely Continuous** with respect to ν, denoted $\mu \ll \nu$ if for all $E \in \Sigma$, $\nu(E) = 0$ implies $\mu(E) = 0$*

Definition 16. *A measure space, (Ω, Σ, μ) is σ-finite if there is a sequence $(A_i)_{i \in \mathbb{N}}$ with $A_i \in \Sigma$ for $i \in \mathbb{N}$ such that:*

$$\Omega = \bigcup_{n=1}^{\infty} A_i$$

and $\mu(A_i) < \infty$

Theorem 1 (Radon Nikodym)**.** *Let (Ω, Σ) be a measurable space with two measures, μ and ν such that $\mu \ll \nu$, (Ω, Σ, μ) is σ-finite and (Ω, Σ, ν) is σ-finite. Then there exists a $(\Omega, \Sigma) \to (\mathbb{R}, B(\mathbb{R}))$ measurable function, $f \colon \Omega \to \mathbb{R}$ such that for all $E \in \Sigma$:*

$$\mu(E) = \int_E f d\nu$$

*f is called the **Radon Nikodym Derivative**:*

$$\frac{d\mu}{d\nu} = f$$

Proof. see [1] $\square$

Why this is important is that it allows us to easily calculate expectation value if there is a probability density function.
Let (Ω, Σ, μ) be a probability space. Let $X(\omega) \colon \Omega \to \mathbb{R}$ be a $(\Omega, \Sigma) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ measurable random variable with some probability distribution function, $\rho(x)$. Then there is a pushforward measure on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$, P such that:

$$P(E) = \mu(X^{-1}(E)) = \int_E \rho(x)dx$$

In this way, $\rho(x)$ is the Radon Nikodym derivative,

$$\frac{d\mu}{dx} = \rho(x)$$

From definition 14:

$$\mathbb{E}(X(\omega)) = \int_\Omega X(\omega)d\mu(\omega) \tag{1.1}$$

Morally we may think of $y = X(\omega)$, $d\mu = \rho(y)dy$ in an analogous way to u-substitution from elementary calculus

$$\mathbb{E}(X(\omega)) = \int_\mathbb{R} y\rho(y)dy \tag{1.2}$$

provided this integral exists.

More generally, if $\psi(y)$ is a measurable function: Then we may arrive at:

$$\mathbb{E}(\psi(X(\omega))) = \int_\mathbb{R} \psi(y)\rho(y)dy \tag{1.3}$$

provided this integral exists.

Definition 17. *Given a random variable X, we define the* ***variance*** *of X as* $Var(X) = \mathbb{E}((\mathbb{E}(X) - X)^2)$*, provided it exists.*

By simply expanding and noting that expectation is linear, we get that $Var(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ which is the most commonly used form.

Definition 18. ***Brownian Motion*** *Brownian Motion is a real valued stochastic process, B_t, $t \geq 0$ satisfying:*

1. *if* $t_1 > s_1 > t_2 > s_2$*, then* $\mathbb{E}((B_{t_2} - B_{s_2})(B_{t_1} - B_{s_1})) = \mathbb{E}(B_{t_2} - B_{s_2})\mathbb{E}(B_{t_1} - B_{s_1})$
2. $B_t - B_s$ *have a normal distribution with mean 0 and variance t, that is for* $t > s \geq 0$*, and* $A \in B(\mathbb{R})$*,* $P((B_t - B_s) \in A) = \int_A \frac{1}{\sqrt{2\pi t}} \exp\{\frac{-x^2}{2t}\}dx$
3. $B_0 = 0$
4. *almost surely, the function* $t \to B_t$ *is continuous*

Brownian motion is a map such that for all $t \geq 0$, $B_t \colon \Omega \to \mathbb{R}$ and $\omega \mapsto B_t(\omega)$. Then almost surely, $t \mapsto B_t(\omega) \in C(\mathbb{R}_{\geq 0}, \mathbb{R})$ where Ω is the set of all Brownian paths. Then we may define $t \mapsto B_t(\omega)$ as one particular parameterized Brownian path, with "time" parameter t. That is, $t \mapsto B_t(\omega) \colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ and $B_t(\omega)$ is almost surely continuous.

We may calculate the expectation value of Brownian Motion:

$$\mathbb{E}(B_t) = \int \frac{y}{\sqrt{2\pi t}} \exp\{\frac{-y^2}{2t}\} dy \tag{1.4}$$

Using (1.2). Since we are integrating over all space and the integrand is odd, we have that:

$$\mathbb{E}(B_t) = 0 \tag{1.5}$$

Recall from definition 14:

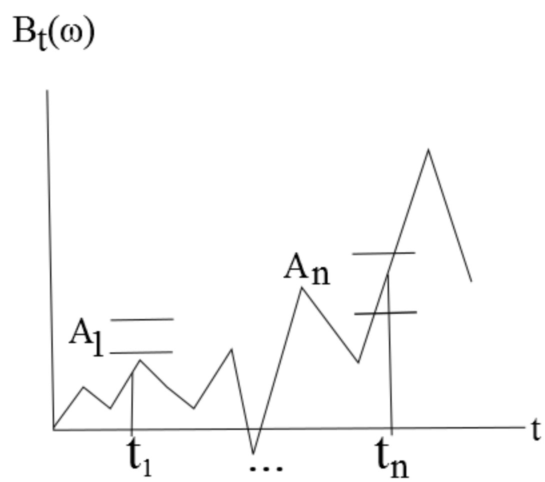
$$\mathbb{E}((B_t)) = \int_\Omega B_t(\omega) d\mu(\omega) \tag{1.6}$$

This is a sum over all Brownian paths $B_t(\omega)$ $\omega \in \Omega$, each with a weight factor of $d\mu(\omega)$. What is the weight for this summation? What is $d\mu(\omega)$?
To give a heuristic picture, we introduce the idea of "gates". That is, we consider times:

$$t_1, ..., t_n \in \mathbb{R}_{\geq 0}$$

and elements of $\mathscr{B}(\mathbb{R})$:

$$A_1, ...A_n$$



We want to find the weight for the Brownian particles going through these n gates:

$$\mu(B_{t_1}^{-1} \in A_1, ..., B_{t_n}^{-1} \in A_n)$$

using the pushfoward measure gives:

$$\mu(B_{t_1}^{-1}(A_1), ..., B_{t_n}^{-1}(A_n)) = P(B_{t_1} \in A_1, ..., B_{t_n} \in A_n) \tag{1.7}$$

By definitions 12 and 13:

$$P(B_{t_i} \in A_i) = \int_{A_i} \rho(y)dy$$

By independence:

$$P(B_{t_1} \in A_1, ..., B_{t_n} \in A_n) = \int_{A_1} \rho(y_1)dy_1 ... \int_{A_n} \rho(y_n)dy_n = \int_A \rho(y_1), ...\rho(y_n)dy_1...dy_n$$

where $A = A_1 \times ... \times A_n$

Coming back to our question, the weight $d\mu(\omega)$ can heuristically be thought of as the infinitesimal version of (1.7) as $n \to \infty$ and as $\Delta t = t_i - t_{i-1} \to 0$. This gave rise to Feynman Path integration [2].

Chapter 2

Heat Equation

2.1 A function

The connection between partial differential equations and Brownian Motion can be shown by using nothing more than basic differentiation. Let B_t be a generalized Brownian motion. We introduce a scaled time parameter, $t \mapsto Dt$, $D > 0$. We define $u(x,t) = \mathbb{E}(\psi(x - B_t))$ as a function, where $\psi(x) \in C_c^\infty(\mathbb{R})$, the compactly supported smooth functions.

$$u(x,t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi Dt}} \exp\left(\frac{-z^2}{2Dt}\right) \psi(x-z)dz \tag{2.1}$$

Then letting:

$$y = x - z, dy = dz$$

$$u(x,t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi Dt}} \exp\left(\frac{-(x-y)^2}{2Dt}\right) \psi(y)dy \tag{2.2}$$

We may differentiate this function with respect to t:

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial t} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi Dt}} \exp\left(\frac{-(x-y)^2}{2Dt}\right) \psi(y)dy \tag{2.3}$$

We would like to pass the derivative inside the integral now, in order to do this we use:

Theorem 2 (Lebesgue Dominated Convergence Theorem)**.** *Let $\{f_n\}$ be a sequence of real valued measurable functions on a measure space, (Ω, Σ, μ) such that f_n converges pointwise to some function $f(x)$. Assume that there exists some integrable function $g(x)$ such that:*

$$|f_n(x)| \leq g(x)$$

Then:

$$\lim_{n\to\infty} \int_\Omega f_n(x)d\mu = \int_\Omega f(x)d\mu$$

Letting $v = \sqrt{t}y$ and $dv = \sqrt{t}dy$, we may write $u(x,t)$ as:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi D}} \exp\left(\frac{-v^2}{2D}\right) \psi(x - \sqrt{t}v)dv \tag{2.4}$$

Now fix x and $t > 0$.

$$\frac{\partial u}{\partial t} = \frac{1}{\sqrt{2\pi D}} \lim_{h\to 0} \int_{-\infty}^{\infty} \exp\left(\frac{-v^2}{2D}\right) \frac{1}{h} \left(\psi(x - \sqrt{t+h}v) - \psi(x - \sqrt{t}v)\right) dv \tag{2.5}$$

By the intermediate value theorem, for an arbitrarily small h there is a $c \in [t, t+h]$ where $[t, t+h] \cap \mathbb{R}_0^- = \emptyset$ such that

$$\frac{1}{h} \left(\psi(x - \sqrt{t+h}v) - \psi(x - \sqrt{t}v)\right) = v\frac{-1}{2\sqrt{c}}\psi'(x - \sqrt{c}v) \tag{2.6}$$

Since $\psi(x - \sqrt{s}v)$ is smooth, it is bounded for all $s \in [t, t+h]$. Since $t > 0$, $\frac{-1}{2\sqrt{s}}$ is bounded on the same interval. Define:

$$g(x,v) := \sup_{s\in[t,t+h]} \left| v\frac{-1}{2\sqrt{s}}\psi'(x - \sqrt{s}v)\right| \tag{2.7}$$

For arbitrarily small h

$$\left|\exp\left(\frac{-v^2}{2D}\right) \frac{1}{h} \left(\psi(x - \sqrt{t+h}v) - \psi(x - \sqrt{t}v)\right)\right| \leq \left|\exp\left(\frac{-v^2}{2D}\right) g(x,v)\right| \tag{2.8}$$

Clearly,

$$\exp\left(\frac{-v^2}{2D}\right) g(x,v)$$

is integrable, so using this dominating function, we may employ Lebesgue's dominated convergence theorem:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \int_{-\infty}^{\infty} \frac{\partial}{\partial t} \frac{1}{\sqrt{2\pi Dt}} \exp\left(\frac{-(x-y)^2}{2Dt}\right) \psi(y)dy \\ &= \int_{-\infty}^{\infty} \psi(y) \exp\left(\frac{-(x-y)^2}{2Dt}\right) \frac{1}{\sqrt{2\pi Dt}} \left(\frac{(x-y)^2}{2Dt^2} - \frac{1}{2t}\right) dy \end{aligned} \tag{2.9}$$

Analogously, we can use Dominated Convergence with respect to x. Differentiating with respect to x twice gives:

$$\frac{\partial u}{\partial x} = \int_{-\infty}^{\infty} \psi(y) \exp\left(\frac{-(x-y)^2}{2Dt}\right) \frac{1}{\sqrt{2\pi Dt}} \frac{-(x-y)}{Dt} dy \tag{2.10}$$

$$\frac{\partial^2 u}{\partial x^2} = \int_{-\infty}^{\infty} \psi(y) \exp\left(\frac{-(x-y)^2}{2Dt}\right) \frac{1}{\sqrt{2\pi Dt}} \left(\frac{(x-y)^2}{D^2t^2} - \frac{1}{Dt}\right) dy \tag{2.11}$$

Now comparing (2.9) and (2.11) leads to the result:

$$\frac{\partial u}{\partial t} = \frac{1}{2} D \frac{\partial^2 u}{\partial x^2} \tag{2.12}$$

That is, u solves the heat equation! This is a remarkable result, because on one hand we have something seemingly deterministic, that's tied into something completely stochastic! Also, since ψ is never defined, we may vary ψ to get *ALL* solutions to the heat equation. Also note that at time $t = 0$, $u(x,t) = \mathbb{E}(\psi(x-0)) = \psi(x)$. We must check continuity at $t = 0$.

$$\lim_{t\to 0^+} u(x,t) = \lim_{t\to 0^+} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi Dt}} \exp\left(\frac{-z^2}{2Dt}\right) \psi(x-z)dz \tag{2.13}$$

Again, let $v = \sqrt{t}z$ and $dv = \sqrt{t}dz$ to get to the form in (2.4)

$$\lim_{t\to 0^+} u(x,t) = \lim_{t\to 0^+} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi D}} \exp\left(\frac{-v^2}{2D}\right) \psi(x-\sqrt{t}v)dv \tag{2.14}$$

Here we may yet again apply Lebesgue's dominated convergence theorem, also noting that $\psi(x)$ is continuous:

$$\begin{aligned}\lim_{t\to 0^+} u(x,t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi D}} \exp\left(\frac{-v^2}{2D}\right) \psi(x)dv \\ &= \psi(x) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi D}} \exp\left(\frac{-v^2}{2D}\right) dv \\ &= \psi(x)\end{aligned} \tag{2.15}$$

This leads us to the Feynman-Kac formula:

Theorem 3 (Feynman-Kac)**.** *Given the partial differential equation:*

$$\frac{\partial u}{\partial t} = \frac{1}{2} D \frac{\partial^2 u}{\partial x^2}$$

and an initial condition:

$$u(x,0) = \psi(x)$$

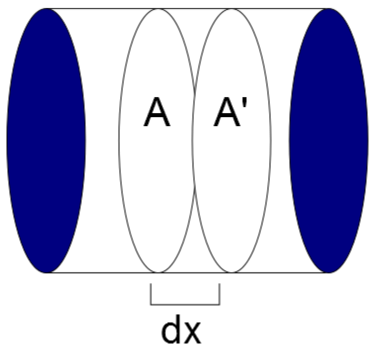
the solution is

$$u(x,t) = \mathbb{E}(\psi(x - B_t))$$

2.2 Einstein-Smoluchowski

In his 1905 paper, Einstein derived the heat equation on a macroscopic scale using thermodynamic arguments [2]. This explicitly shows the relation between Brownian Motion and the Heat Equation and also demonstrates the validity of atomism. It shows how macroscopic phenomenon can be constructed by microscopic principles.

Consider a cylinder, with cross-sectional area A, full of a liquid, and Brownian particles suspended in it, as seen in picture:



We assume that the particles don't interact. We additionally assume that they all have the same mass. The aim is to find $\rho(x,t) = \frac{n(x)}{V}$, the density of particles where $n(x)$ is the number of particles and V is volume. First, note that $\rho(x,t)$ must satisfy the continuity equation:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial j}{\partial x} \tag{2.16}$$

where j is the current. We also know the ideal gas law:

$$PV = n(x)kT$$

where $n(x)$ is the number of particles in $V = dV$, k, and T are the Boltzmann constant and temperature.
We divide by V and arrive at:

$$P = \rho kT \tag{2.17}$$

We may then differentiate with respect to x:

$$\frac{dP}{dx} = \frac{d\rho}{dx}kT \tag{2.18}$$

Multiplying by $\frac{A}{n(x)}$

$$\frac{A}{n(x)}\frac{dP}{dx} = \frac{d\rho}{dx}kT\frac{A}{n(x)} \tag{2.19}$$

$$A\frac{dP}{n(x)} = \frac{d\rho}{dx}kT\frac{dV}{n(x)} \tag{2.20}$$

Noting that area times pressure is force, and $\frac{1}{\rho} = \frac{dV}{n(x)}$, we get that the force per particle due to pressure is:

$$\frac{dF}{n(x)} = \frac{1}{\rho}\frac{d\rho}{dx}kT \tag{2.21}$$

However, not only do particles experience a force from pressure, there is also a friction force, which we may model as $f = -\gamma v$ where v is the velocity of the particle and γ is the friction coefficient. In equilibrium these are equal. We have:

$$-\gamma v = \frac{1}{\rho}\frac{d\rho}{dx}kT \tag{2.22}$$

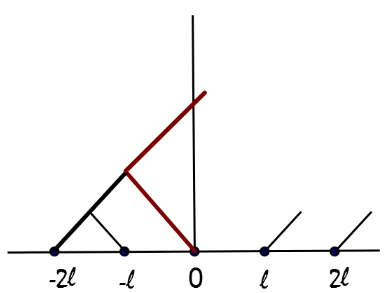
$$-\rho v = \frac{kT}{\gamma}\frac{d\rho}{dx} \tag{2.23}$$

But $\rho v = j$, the current! Plugging (2.23) into (2.16) gives that:

$$\frac{\partial \rho}{\partial t} = \frac{kT}{\gamma}\frac{\partial^2 \rho}{\partial x^2} \tag{2.24}$$

That is, we have derived the heat equation! This equations governs the motion of density of the "Brownian" particle. But Einstein never even defined what "Brownian Motion" is! Without a microscopic picture, Einstein predicts the constant D! This constant has been verified experimentally.
Now that we know that the solution is a probability distribution, it makes (2.1) seem reasonable. What exactly IS Brownian motion? Why should we expect Brownian motion? Why should we expect a normal distribution to solve the heat equation? To answer these questions, Smulochowski offers another derivation.

We now turn away from the macroscopic principles of Einstein's arguments, and turn to a microscopic argument.
To start, we consider infinitely many particles of the same mass on a lattice with spacing l. We look at one particle, and center it at "0". We give each particle a random starting velocity, $\pm v$ each with probability $\frac{1}{2}$. We are interested in finding X_t, the position of the "middle" particle at a time t. Define $\Delta t = \frac{l}{2v}$. Δt is the time step. That is, every Δt seconds, we have a collision (or lack of). This is essentially a random walk. Every time step the particle goes to the left with probability a half, right with probability a half. For example, one situation might look like the following picture:



We conclude that:

$$X_t = \sum_{k=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \Delta x_k \tag{2.25}$$

Where $\lfloor \frac{t}{\Delta t} \rfloor$ is the floor function, that is the greatest integer $n \leq \Delta t$. What this sum is saying is that every Δt seconds, there is a change in position, Δx_k. The sum of these changes is the position at time t. Δx_k is $\frac{l}{2}$ with probability $\frac{1}{2}$ and $\frac{-l}{2}$ with probability $\frac{1}{2}$. In addition, we assume that Δx_i and Δx_j are independent if $i \neq j$.

We may now compute expectation values:

$$\mathbb{E}(X_t) = \mathbb{E}\left(\sum_{k=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \Delta x_k \right) \tag{2.26}$$

Then by linearity, we can get:

$$\begin{aligned} \mathbb{E}(X_t) &= \sum_{k=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \mathbb{E}\Delta x_k \\ &= \left\lfloor \frac{t}{\Delta t} \right\rfloor \left(\frac{-l}{2}\frac{1}{2} + \frac{l}{2}\frac{1}{2} \right) \\ &= 0 \end{aligned} \tag{2.27}$$

We have expectation of 0, which is characteristic of Brownian motion, now lets calculate:

$$\mathbb{E}(X_t^2) = \mathbb{E}\left(\sum_{k=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \Delta x_k \right)^2 \tag{2.28}$$

Then by expanding:

$$\mathbb{E}(X_t^2) = \sum_{k=1}^{\lfloor \frac{t}{\Delta t} \rfloor} \mathbb{E}(\Delta x_k^2) + \sum \mathbb{E}(c_{i,j} \Delta x_i \Delta x_j) \tag{2.29}$$

The last term vanishes, as linear terms have expectation value zero and cross terms are independent.

$$\begin{aligned} \mathbb{E}(X_t^2) &= \left\lfloor \frac{t}{\Delta t} \right\rfloor \left(\frac{l^2}{4}\frac{1}{2} + \frac{l^2}{4}\frac{1}{2} \right) \\ &= \left\lfloor \frac{t}{\Delta t} \right\rfloor \frac{l^2}{4} \end{aligned} \tag{2.30}$$

For any fixed l or v, we have that X_t^2 is proportional to t, we write this as $X_t^2 \sim t$

This already looks a lot like Brownian Motion. Recall the definition of Brownian motion (def 18). We have independent, increments with mean 0. However here we have discrete time steps rather than continuous time. To introduce continuous time, we scale X_t.

Define $X_t^\epsilon = \epsilon X_{t/\epsilon^2}$. We must scale out, and limit $\Delta t \to 0$. We choose the prefactor and scaled time parameter to be in a certain ratio, as to maintain mean and variance. Again, lets compute expectation values:

$$\mathbb{E}(X_t^\epsilon) = \epsilon 0 = 0 \tag{2.31}$$

$$\mathbb{E}((X_t^\epsilon)^2) = \epsilon^2 \mathbb{E}((X_{t/\epsilon^2})^2) \sim \epsilon^2 \frac{t}{\epsilon^2} = t \tag{2.32}$$

We still have the same properties when we scale, but in order to get Brownian Motion, we must take the $\epsilon \to 0$ limit.

First, we find the probability density function, $p^\epsilon(x,t)$ for X_t^ϵ, then take the limit $\epsilon \to 0$. To do this, we consider the Fourier Transform, $\widehat{p^\epsilon(x,t)}$

$$\sqrt{2\pi}\widehat{p^\epsilon(x,t)}(k) = \mathbb{E}(e^{-ikX_t^\epsilon}) = \int_{-\infty}^{\infty} e^{-ikx} p^\epsilon(x,t) dx \tag{2.33}$$

Then we may expand X_t^ϵ:

$$\mathbb{E}(e^{-ikX_t^\epsilon}) = \mathbb{E}\left(\exp\left\{-ik\epsilon \sum_{j=1}^{\lfloor \frac{t}{\epsilon^2 \Delta t} \rfloor} \Delta x_j\right\}\right) \tag{2.34}$$

$$\mathbb{E}(e^{-ikX_t^\epsilon}) = \mathbb{E}\left(\prod_{j=1}^{\lfloor \frac{t}{\epsilon^2 \Delta t} \rfloor} e^{-i\epsilon k \Delta x_j}\right) \tag{2.35}$$

Then by independence, we may move the expectation value inside:

$$\mathbb{E}(e^{-ikX_t^\epsilon}) = \prod_{j=1}^{\lfloor \frac{t}{\epsilon^2 \Delta t} \rfloor} \mathbb{E}(e^{-i\epsilon k \Delta x_j}) \tag{2.36}$$

Then since all Δx_j are identically distributed, we may turn the product into a power. Also we Taylor expand:

$$\mathbb{E}(e^{-ikX_t^\epsilon}) = (\mathbb{E}(1 - ik\epsilon\Delta x_j - \frac{1}{2}k^2\epsilon^2(\Delta x_j)^2 + O(\epsilon^3))^{\lfloor \frac{t}{\epsilon^2 \Delta t} \rfloor} \tag{2.37}$$

where $O(\epsilon^3)$ means higher ordered terms.

$$\mathbb{E}(e^{-ikX_t^\epsilon}) = \left(1 - \frac{1}{8}k^2\epsilon^2 l^2 + O(\epsilon^4)\right)^{\lfloor \frac{t}{\epsilon^2 \Delta t} \rfloor} \tag{2.38}$$

The higher order terms don't matter. To see this, simply Taylor expand again and take the $\epsilon \to 0$ limit. Define $\epsilon_n^2 = \frac{1}{n\Delta t}$. Then:

$$\mathbb{E}(e^{-ikX_t^{\epsilon_n}}) = (1 - \frac{tl^2k^2}{8n\Delta t} + O(\frac{1}{n^2}))^n \tag{2.39}$$

Taking limits at $n \to \infty$ and $\epsilon \to 0$

$$\lim_{n\to\infty} \sqrt{2\pi}\widehat{p^\epsilon(x,t)}(k) = \lim_{n\to\infty} \left(1 - \frac{tl^2k^2}{8n\Delta t} + O(\frac{1}{n^2})\right)^n = \exp\{\frac{-lvtk^2}{4}\} \tag{2.40}$$

We may inverse Fourier transform to get

$$\begin{aligned} p(x,t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}e^{\frac{-lvtk^2}{4}}dk \\ &= \int_{-\infty}^{\infty} \exp\{ikx - \frac{lvtk^2}{4}\}dk \end{aligned}$$

Let $a = \frac{lvt}{4}$ and $b = ix$ for convenience. This becomes:

$$\begin{aligned} p(x,t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\{-ak^2 + bk\}dk \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\{-a\left(k^2 - \frac{b}{a}k\right)\}dk \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\{-a\left(k^2 - \frac{b}{a}k + \frac{b^2}{4a^2} - \frac{b^2}{4a^2}\right)\}dk \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\{-a\left(k - \frac{b}{2a}\right)^2 - \frac{b^2}{4a}\}dk \\ &= \frac{1}{2\pi}\exp\{-\frac{b^2}{4a}\}\int_{-\infty}^{\infty} \exp\{-a\left(k - \frac{b}{2a}\right)^2\}dk \end{aligned} \tag{2.41}$$

Then let $y = k - \frac{b}{2a}$, $dy = dk$. Then this integral becomes:

$$p(x,t) = \frac{1}{2\pi}\exp\{-\frac{b^2}{4a}\}\int_{-\infty}^{\infty} e^{-ay^2}dy \tag{2.42}$$

Let $u = \sqrt{a}y$, then $du = \sqrt{a}dy$ and $y^2 = \frac{1}{a}u^2$. This makes the integral:

$$p(x,t) = \frac{1}{\sqrt{a}}\frac{1}{2\pi}\exp\{-\frac{b^2}{4a}\}\int_{-\infty}^{\infty} e^{-u^2}du \tag{2.43}$$

This is a well known integral:

$$\int_{-\infty}^{\infty} e^{-u^2}du = \sqrt{2\pi}$$

so we arrive at:

$$p(x,t) = \frac{1}{\sqrt{2\pi a}} \exp\{-\frac{b^2}{4a}\}$$

Plugging in the constants, a and b gives that:

$$p(x,t) = \frac{1}{\sqrt{\pi l v t}} \exp\{-\frac{x^2}{lvt}\} \tag{2.44}$$

To connect these two arguments, we can compare constants to get that:

$$\frac{kT}{\gamma} = \frac{1}{2}D = lv \tag{2.45}$$

We have modeled a Brownian motion as a sequence of particle collisions. We then explained why it should have a Gaussian distribution. By Feynman-Kac, we know that the expectation of this is the solution to the heat equation.

Chapter 3

Other Remarks

3.1 Existence of Brownian Motion

Although morally, Smulochowski's argument demonstrates that Brownian Motion exists, it lacks certain technical considerations. The first problem is that we don't know if there is *ONE* probability space with *ONE* probability measure where the random variables, $B_t, t \geq 0$ live. The second problem is continuity. In the limit $\epsilon \to 0$, intuitively the resulting process should be continuous. But it's not guaranteed. To take care of the first consideration one resorts to the Kolmogorov's Extension Theorem. The full proof that Brownian Motion exists can be found in [1]. We only wish to point the difficulties.

Theorem 4 (Kolmogorov)**.** *Define a measure on* $\mathbb{R}$*:*

$$p_{t_m - t_{m-1}}(a, b) \coloneqq \frac{1}{\sqrt{2\pi t}} \exp\{\frac{-(b-a)^2}{2t}\}$$

and a measure on $\mathbb{R}^n$*:*

$$\mu_{t_1,\ldots t_n}(A_1 \times \ldots \times A_n) = \int_{A_1} dx_1 \ldots \int_{A_n} dx_n \prod_{m=1}^{n} p_{t_m - t_{m-1}}(x_m, x_{m-1})$$

Where $A_i \in \mathscr{B}(\mathbb{R})$*,* $t_0 = 0$
In addition, assume for all $\{s_1, \ldots, s_{n-1}\} \subset \{t_1, \ldots, t_n\}$*:*

$$\mu_{s_1,\ldots s_{n-1}}(A_1 \times \ldots \times A_{n-1}) = \mu_{t_1,\ldots t_n}(A_1 \times \ldots \times \mathbb{R} \times A_{j+1} \times \ldots \times A_n)$$

where $s_j \notin \{t_1, \ldots, t_n\}$
Let $\Omega_0 = \{$*functions* $\omega \colon [0, \infty) \to \mathbb{R}\}$ *and* F_0 *be the* Σ*-algebra generated by the finite dimensional sets* $\{\omega \colon \omega(t_i) \in A_i$ *for* $1 \leq i \leq n\}$*, for* $A_i \in B(\mathbb{R})$*. Then there is a unique probability measure,* ν *on* (Ω_0, F_0) *such that* $\nu\{\omega \colon \omega(0) = 0\} = 1$ *and for* $0 < t_1 < \ldots < t_n$

$$\nu(\{\omega \colon \omega(t_i) \in A_i\}) = \mu_{t_1,\ldots t_n}(A_1 \times \ldots \times A_n) \tag{3.1}$$

where μ *is the probability measure*

Kolmogorov's extension theorem gives the existence of a **Gaussian Process**

Definition 19. *Given a probability space,* (Ω, Σ, μ)*, a* ***Gaussian Process****,* $X_t(\omega) \colon \Omega \to \mathbb{R}$ *is a stochastic process with independent, normally distributed increments.*

In closing we have a stochastic process $B_t(\omega) \colon \Omega \to \mathbb{R}$ such that:

1. if $t_1 > s_1 > t_2 > s_2$, then $\mathbb{E}((B_{t_2} - B_{s_2})(B_{t_1} - B_{s_1})) = \mathbb{E}(B_{t_2} - B_{s_2})\mathbb{E}(B_{t_1} - B_{s_1})$
2. $B_t - B_s$ have a normal distribution with mean 0 and variance t, that is for $t > s \geq 0$, and $A \in B(\mathbb{R})$, $P((B_t - B_s) \in A) = \int_A \frac{1}{\sqrt{2\pi t}} \exp\{\frac{-x^2}{2t}\}dx$
3. $B(0) = 0$

However, we are still lacking continuity. This is challenging, because we can't even check that the process Kolmogorov's extension theorem gives us is continuous. Meaning:

$$\{\omega \in \Omega | t \mapsto X_t \text{is continuous}\} \notin F_0$$

This is because the distribution depends on only countably many coordinates.

3.2 Schrödinger Equation, Feynman Path Integrals

In his PhD Thesis, Richard Feynman created a new formulation of quantum mechanics. [4] He called this the *Path Integral Formulation* of quantum mechanics. We give a quick overview of it here, and explain how quantum mechanics is related to statistical mechanics. In his paper, he proposes that all possible paths contribute to the observed path of a particle. That is, each path has a certain weight associated with it, then all these paths are summed over with this weight. This is known as a path integral.
What does this have to do with Brownian Motion? Well, we are integrating over paths with a certain weight. This should remind you of taking expectation values of Brownian Motion! Lets re-frame the question. Take the function (2.1) and perform what's called *Wick Rotation*. That is, map $t \to it$, introducing the idea of imaginary time. This function becomes:

$$u(x,t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi Dit}} \exp\{\frac{-(x-y)^2}{2Dit}\}\psi(y)dy \tag{3.2}$$

We may define $C = iD$. Then the computation in chapter 2 holds exactly the same, and we get that the new, Wick Rotated function solves:

$$\frac{\partial u}{\partial t} = \frac{1}{2}C\frac{\partial^2 u}{\partial x^2} \tag{3.3}$$

Or

$$\frac{\partial u}{\partial t} = \frac{1}{2} i D \frac{\partial^2 u}{\partial x^2} \tag{3.4}$$

That is, u solves the "free" Schrödinger equation:

$$\frac{\partial u}{\partial t} = \frac{i\hbar}{2m} \frac{\partial^2 u}{\partial x^2} \tag{3.5}$$

where $D = \frac{\hbar}{m}$ and $\hbar$ is the reduced Planck's constant and m is the mass of a particle. By "free", we mean that there is no potential function.
In statistical mechanics, solutions to the heat equation are generated by expectation values of functions of Brownian motion running in real time. In quantum mechanics in a loose sense, the solutions to the Schrödinger equation are generated by expectation values of functions of Brownian motion running in imaginary time.

3.3 Closing Remarks

The connection between Brownian motion and partial differential equations is truly remarkable. On one hand, we have something that by its very definition is random, on the other hand we have something that is truly deterministic. They are connected in a deep and maybe at first unexpected way. Looking forward, we may abstract this connection even further through *Stochastic Calculus*. For this, I recommend [3]. Stochastic calculus formalizes this connection through stochastic differential equations (SDEs).

Bibliography

[1] Rick Durrett. *Probability Theory and Examples* 2013.

[2] Detlef Dürr, Stefan Teufel. *The Physics and Mathematics of Quantum Theory* 2009.

[3] Fima C. Klebaner. *Introduction to Stochastic Calculus with Appliciations* 2006.

[4] Sakurai, J. J., San Fu Tuan. *Modern Quantum Mechanics* 1985.

[5] Einstein, Albert, trans. Cowper, A.D. *Investigations on the Theory of the Brownian Movement* 1956.