BROWNIAN MOTION

JUSTIN HARTMANN

ABSTRACT. This paper begins to explore a rigorous introduction to probability theory using ideas from algebra, measure theory, and other areas. We start with a basic explanation of terms and ideas and then move onto an attempt at explaining Brownian motion.

CONTENTS

This paper attempts to explain the phenomenon known as Brownian motion. As a result, a thorough knowledge of probability theory is needed. We will assume the reader has a cursory knowledge of probability terms and ideas. Therefore, we will provide an introduction to probability theory, but by no means a complete one. We begin with an introduction to measure theory.

1. MEASURE THEORY

Before we get to what a measure is, we must have some objects to measure:

Definition 1.1. An *algebra* is essentially a family of sets closed under finitely many set operations. That is, if S is a set then a collection Σ_0 of subsets of S is an algebra if:

(1) $S \in \Sigma_0$
(2) $F \in \Sigma_0 \implies F^c \in \Sigma_0$
(3) $F, G \in \Sigma_0 \implies F \cup G \in \Sigma_0$.

Definition 1.2. A *sigma algebra* Σ is an algebra closed under countably many set operations, i.e. whenever $F_n \in \Sigma$, then

$$\bigcup_n F_n \in \Sigma.$$

Of course $A \cap B = (A^c \cup B^c)^c$, so both algebras and sigma algebras are closed under finite and countable intersections, respectively. Now let us consider what we can do if our set, S, is a topological space.

Date: August 24, 2009.

Definition 1.3. If S is a topological space, then the sigma algebra generated by the family of open sets of S is called the *Borel algebra*. We will write simply $\mathcal{B}$ for the Borel algebra generated by the open sets of $\mathbb{R}$, $\mathcal{B}(\mathbb{R})$.

Now, this definition is a bit hard to digest. For instance, wha does the "sigma algebra generated by the family of open sets" really mean? Let us discuss this. Take a general set X and a subset of that set A. Then let's set ξ to be the collection of all σ -algebras on X that contain A. Then, by the definition, the intersection of all the σ-algebras in ξ is a σ-algebra itself. We'll denote it by $\sigma(A)$. Then, we call $\sigma(A)$ the sigma algebra generated by A. Now, replace A with the family of open sets of S, and we have the Borel algebra.

Now we have the objects we need, so we can define the space that we will use for the measure.

Definition 1.4. A pair (S, Σ) where S is a set and Σ is a sigma algebra is a *measurable space*.

Definitions 1.5. A *measure* μ is a *countably additive* map μ: $\Sigma \rightarrow [0, \infty]$ where Σ is a sigma algebra. By countably additive, we mean that for a sequence $A_1, A_2, \ldots \in \Sigma$, $\mu\left(\cup_{n=1}^{\infty} A_n\right) = \Sigma_{n=1}^{\infty} \mu(A_n)$. We also require that $\mu(\emptyset) = 0$. A triple (S, Σ, μ) is a *measure space*.

Now, to begin to work with probability, we need a specific type of measure. We need a *probability measure* and a corresponding *probability triple.*

Definition 1.6. If $\mu(S) = 1$, then μ is a *probability measure*. Then, the triple (S, Σ, μ) is a *probability triple*. In probability theory it is common to denote S by Ω, Σ by $\mathcal{F}$, and μ by $\mathbb{P}$. Thus we have the triple $(\Omega, \mathcal{F}, \mathbb{P})$.

2. Working with the Probability Triple

Now that we have our triple, $(\Omega, \mathcal{F}, \mathbb{P})$, we must begin to work with our triple in order to perform experiments and find probabilities of outcomes. First, we must introduce some terminology.

Definitions 2.1. We call Ω the *sample space* and a point $\omega \in \Omega$ is called a *sample point*. The $\sigma-$ algebra $\mathcal{F}$ on Ω is called the *family of events*. Then, an *event* F is a member of $\mathcal{F}$, or a $\mathcal{F}-$ measurable subset of Ω.

Intuitively, we can think of Ω as the set of all possible outcomes and ω as one of these outcomes. Then, $\mathbb{P}(F)$ is the probability that $\omega \in F$.

Example 2.2. Consider rolling a fair, six-sided die. Then, $\Omega = \{1, 2, 3, 4, 5, 6\}$ and $\mathcal{F}$ is the set of all subsets of Ω, in other words, the power set. Take the event F that you roll a prime number. Then $F = 2, 3, 5$ and $\mathbb{P}(F) = \frac{3}{6} = \frac{1}{2}$.

Let us notice a few things about this example. First of all, the σ-algebra $\mathcal{F}$ is the power set of Ω. This is true because our example is discrete. If, for example, Ω were uncountable, then $\mathcal{F}$ really could not be the power set. If we did try to set $\mathcal{F}$ to the power set, it would be impossible to give a unique measure to each set in $\mathcal{F}$. It is simply too large. Think of it this way: when Ω is discrete, the size of the power set, and thus of $\mathcal{F}$, is 2^{Ω}. What happens if Ω is uncountable? It is easy to see that we quickly run into problems. In this case we use $\mathcal{B}(\Omega)$ for $\mathcal{F}$.

Another thing to notice here is that the probability measure, $\mathbb{P}$ is uniform. Every set of $\mathcal{F}$ with the same number of elements has the same measure. This is only true becasue we have assumed a "fair" die. If we instead assumed the die was weighted so that a 6 showed up half of the time, then the measure would not be uniform. Every set with a 6 would have at least a probability of $\frac{1}{2}$.

So, we can now calculate the probabilities of events. This is nice and useful, but most often we will not be working with events that have such a clear distinction. Interesting probability theory makes use of random variables, which we will now define. Intuitevely, we shall think of random variables as possible data. They represent what an outcome may or may not be, and the realization of a random variable is an event. So, one may think of the random variable before the experiment being replaced by an event after the experiment is performed. For instance, the number of heads in 5 tosses of a coin is a random variable, and one realization of that random variable might be $\{H, H, T, H, T\}$.

Definition 2.3. Given a measurable space (S, Σ) and a function $h : \Sigma \to \mathbb{R}$ we say h is *Σ − measurable* if $h^{-1} : \mathcal{B} \to \Sigma$. $m\Sigma$ then denotes the class of Σ-measurable functions on S.

Definition 2.4. Then, using our probability triple $(\Omega, \mathcal{F}, \mathbb{P})$, we define a *random variable* to be an element of $m\mathcal{F}$. So, for the random variable X, $X : \Omega \to \mathbb{R}$ and $X^{-1} : \mathcal{B} \to \mathcal{F}$.

Example 2.5. Let us now consider an example of coin tossing. Say we toss a coin N times and record the results. Then our sample space Ω is a collection of N-tuples of all combinations of H, heads, and T, tails: $\Omega = \{H, T\}^N$. Then,

$$\omega = (\omega_1, \omega_2, ..., \omega_N) \text{ where } \omega_n \in \{H, T\},$$

And the random variable $X_n(\omega)$ can be defined as:

$$X_n(\omega) := \begin{cases} 1 & \text{if } \omega_n = H \\ 0 & \text{if } \omega_n = T \end{cases}$$

We then also have the random variable $S_N = X_1 + X_2 + \cdots + X_N =$ number of heads in N tosses of the coin. One must of course check X_n and S_N to make sure they are measurable, which is easy.

So, thus far, we have a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$, with a sample space Ω, a sigma algebra $\mathcal{F}$ generated by the sample space, and a probability measure $\mathbb{P}$ on $\mathcal{F}$. We have also seen how to create events from the sample space and also random variables out of the σ-measurable functions on the sample space. Now, we will see how to put these ideas together to create a probability distribution function.

As we have seen, X^{-1} maps from the Borel algebra, $\mathcal{B}$, to $\sigma(X) \in \mathcal{F}$ and the probability measure, $\mathbb{P}$, maps from $\mathcal{F}$ to $[0, 1]$. By $\sigma(X)$, we mean the σ-algebra generated by the random variable X. That is, $\sigma(X) := X^{-1}(\mathcal{B})$. It is the σ-algebra that is the result of taking the inverse function, X^{-1}, on $\mathcal{B}$. Thus, we have:

$$[0, 1] \xleftarrow{\mathbb{P}} \sigma(X) \xleftarrow{X^{-1}} \mathcal{B}$$

Thus, it is quite natural to think of the composition function of $\mathbb{P}$ and X^{-1}.

Definition 2.6. We will define the *law*, $\mathcal{L}_X$, of X to be: $\mathcal{L}_X = \mathbb{P} \circ X^{-1}$.

Definition 2.7. Now, to work with this, we will want to use a *π-system on* $\mathbb{R}$. *This we will define as a non-empty family subsets of* $\mathbb{R}$ *that is closed under finite intersection.*

Then, $\mathcal{L}_X : \mathcal{B} \to [0,1]$ and, following directly from the definition, $\mathcal{L}_X$ is a probability measure on $(\mathbb{R}, \mathcal{B})$. Now, we can generate $\mathcal{B}$ by using the π-system $\pi(\mathbb{R}) = \{(-\infty, c] : c \in \mathbb{R}\}$. Thus, we can specify the law by using a corresponding function $F_X : \mathbb{R} \to [0,1]$.

Definition 2.8. We will define this function as $F_X(c) := \mathcal{L}_X(-\infty, c]$ and name it the *distribution function* of X.

Intuitively, we can think of this function in probabilistic terms. It is simply the probability that the random variable X is less than or equal to c. Another way to think of it is the probability of the collection of sample points ω such that $X(\omega) \leq c$. Thus we have:

$$F_X(c) = \mathbb{P}(X \leq c) = \mathbb{P}\{\omega : X(\omega) \leq c\}.$$

This is nice to have. Now, given any experiment, we can determine that the probability that a given random variable takes on a certain value is less than or equal to a given c. This allows for many applications itself. To work with the distribution function, we need the following properties:

For a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$, a random variable X, and a distribution function F_X:

(1) $F_X : \mathbb{R} \to [0,1]$.

Proof. The distribution function F_X is a composition function of the law, $\mathcal{L}_X(-\infty, c]$, which is composed of two functions: $\mathbb{P} \circ X^{-1}$. Since the probability measure, $\mathbb{P}$, goes from the σ-algebra to $[0,1]$, the distribution function must go from $\mathbb{R}$ to $[0,1]$. Intuitively, this makes sense because the random variable can take on any value in $\mathbb{R}$ but the probability that it takes on that value can be no less than 0 and no greater than 1. □

(2) F_X is increasing. That is, $x \leq y \Rightarrow F_X(x) \leq F_X(y)$.

Proof. This comes from the definition of the probability measure. From this, we know that if $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$. Well, if $x \leq y$, and we take the events $A = \{X \leq x\}$ and $B = \{X \leq y\}$, then $A \subset B$ and then $\mathbb{P}(A) \leq \mathbb{P}(B)$. Then, from the definition of a distribution function, we know that $F_X(x) \leq F_X(y)$. □

(3) $\lim_{x\to\infty} F_X(x) = 1$ and $\lim_{x\to-\infty} F_X(x) = 0$.

Proof. First, $\lim_{x\to\infty} F_X(x) = \mathbb{P}(X \leq +\infty) = \mathbb{P}(X \in \mathbb{R}) = \mathbb{P}(S) = 1$, using the definition of a probability measure.
Similarly, $\lim_{x\to-\infty} F_X(x) = \mathbb{P}(X \leq -\infty) = \mathbb{P}(\emptyset) = 0$. □

(4) F_X is right-continuous.

Proof. A property of measures is the monotone-convergence property. Basically, it means that if $G_n \in \mathcal{F}$, with $G_n \downarrow G$ and for some k $\mathbb{P}(G_k) < \infty$, then $\mathbb{P}(G_k) \downarrow \mathbb{P}(G)$. Then, we see that $\mathbb{P}(X \leq x + n^{-1}) \downarrow \mathbb{P}(X \leq x)$. So, this, along with (2) give that F_X is right-continuous. □

3. Independence and Expectation

Now we begin a discussion of the topic of independence. We will define what it means for algebras, random variables, and events to be independent. These definitions will allow us great flexibility and problem solving abilities when we discuss the specifics of Brownian motion. For now, let us simply define the terms so we will have them later.

Definition 3.1. Suppose we have our probability triple $(\Omega, \mathcal{F}, \mathbb{P})$. Consider sub-$\sigma$-algebras from $\mathcal{F}$. So, take $F_1, F_2, \ldots$ as sub-σ-algebras of $\mathcal{F}$. Then, we say that $F_1, F_2, \ldots$ are *independent* if for $G_i \in F_i$ and $i \in \mathbb{N}$ we have:

$$\mathbb{P}(G_1 \cap G_2 \cap \cdots \cap G_n) = \prod_1^n \mathbb{P}(G_i)$$

We will now define independence for random variables and events. These will be much easier to define, as they pretty much follow from the independence of σ-algebras.

Definition 3.2. We say that the random variables $X_1, X_2, \ldots$ are *independent* if their corresponding σ-algebras are independent. That is, $X_1, X_2, \ldots$ are independent if $\sigma(X_1), \sigma(X_2), \ldots$ are independent.

Definition 3.3. We call events $E_1, E_2, \ldots$ *independent* if the σ-algebras $\mathcal{E}_1, \mathcal{E}_2, \ldots$ are independent where $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability triple and $\mathcal{E}_i = \{\emptyset, E_i, \Omega \setminus E_i, \Omega\}$.

Definition 3.4. Given a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$, the *expectation* of a random variable X, $E(X)$ is:

$$E(X) := \int_\Omega X \, d\mathbb{P} = \int_\Omega X(\omega) \, \mathbb{P}(d\omega)$$

Definition 3.5. Given two random variables X and Y and their expectations, $E(X) = \mu_x$ and $E(Y) = \mu_Y$, the *covariance of X and Y*, is:

$$\text{Cov}(X, Y) := E[(X - \mu_X)(Y - \mu_Y)]$$

We would like to think of the expectation as the mean, or average, of the random variable. In a sense, it is the most likely value that the random variable will have. The covariance, intuitively, will be a measure for how related two random variables are. So, a positive covariance will mean that there is a direct relationship between how the two variables vary and a negative covariance will mean that there is an indirect relationship between how the two variables vary. We will find both values to be extraordinarily useful. The variance of a random variable, defined below, also has an intuitive meaning. We think of it as a measure of dispersion. That is, it is a measure of how far a random variable is from its mean.

Definition 3.6. The *variance* of a random variable x, $\text{Var}(X)$, is:

$$\text{Var}(X) := E[(X - \mu_X)^2] = \text{Cov}(X, X)$$

4. Brownian Motion

Now we come to the main subject of the paper, that is, Brownian motion. Brownian motion was "discovered" when the English scientist Brown noticed that when he observed pollen particles in water under a microscope slide the pollen was constantly in motion. In fact, pollen slides as old as 100 years were *still moving!* This was a very surprising fact at the time. Over the years numerous mathematicians and statisticians have constructed Brownian motion with little more than a probability distribution, means, and covariances. So, we begin our discussion now with the definition of Brownian motion for one dimension:

Definition 4.1. A stochastic process $\{B(t) : t \geq 0\}$ is called a *Brownian motion* with an initial point $x \in \mathbb{R}$ if:

(1) $B(0) = X$
(2) $B(t)$ has independent increments, meaning that $\forall\, t$ the increments $B(t_n) - B(t_{n-1}), B(t_{n-1}) - B(t_{n-2}), \ldots, B(t_2) - B(t_1)$ are independent random variables.
(3) $\forall\, t \geq 0 \,\text{and}\, h > 0$ the random variables $B(t+h) - B(t)$ have a normal probability distribution with expectation 0 and variance h.
(4) the map $t \to B(t)$ is continuous.

Notice that in the definition we reference the normal probability distribution. Intuitively, the normal distribution describes a random process that centers around a mean. Mathematically, we define it as follows:

Definition 4.2. A random variable X has a *normal distribution* with mean μ and variance σ^2 if:

$$\mathbb{P}(X \leq x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{\frac{-(u-\mu)^2}{2\sigma^2}} du \quad \forall\, x \in \mathbb{R}$$

So, this is nice, but why is this characterization a good model for the random movement discovered by Brown? Well, intuitively, we see that there is a starting point, X, and then that each increment is independent of the other. Later, we will see that this is related to the Markov property. The idea is that if the increments are independent, then the movement of the "particle" has nothing to do with past locations of the particle, only current locations. This is good for randomness. Then, continuity is also good for randomness, since it means that rather than having a definite point at each time t, our "particle" has a probability distribution for being at that point. That probability being the normal distribution is good, because this means that that our "particle" will move more or less randomly. It will be centered around a mean, 0, but vary in each movement with a variation related to its variance, h.

Now that we have a firm definition of Brownian motion, we would like to discover a few properties that make it especially interesting to discover. The first is this, Brownian motion is nowhere differentiable. This statement will of course require some explanation, and a proof. The interesting thing about this statement is this: from the definition we can see that there is some inherent regularity and overall "niceness" to Brownian motion. We need few inputs, really just a starting point, a mean, and the variance, and the increments are not only independent, but normally distributed. On top of that, the function is continuous. Yet, it is nowhere differentiable. That is remarkable.

Theorem 4.3. *For a standard Brownian motion $\{B(t) : t \geq 0\}$, almost surely $B(t)$ is not differentiable at t, and the upper and lower derivatives are $+\infty, -\infty$, respectively.*

Proof. First, we create another process, also a Brownian motion, $\{X(t) : t \geq 0\}$ such that:

$$X(t) = \begin{cases} 0 & \text{if } t = 0, \\ tB(1/t) & \text{if } t > 0 \end{cases}$$

Now, $X(s) = B(t+s) - B(t)$ is a standard Brownian motion with the fact that the derivative of X at 0 is the same as the derivative of B at t. So, we simply need to show that X is not differentiable at 0.

Now, we'll define what we mean by upper and lower derivatives. For the upper derivative, we have:

$$D^t f(t) = \limsup_{h \downarrow 0} \frac{f(t+h) - f(t)}{h}$$

And the lower derivative will be:

$$D_t f(t) = \liminf_{h \downarrow 0} \frac{f(t+h) - f(t)}{h}$$

Now, to find the derivatives of $X(0)$:

$$D^t X(0) \geq \limsup_{n \to \infty} \frac{X(\frac{1}{n}) - X(0)}{\frac{1}{n}} \geq \limsup_{n \to \infty} \sqrt{n} X(\frac{1}{n}) = \limsup_{n \to \infty} \frac{B(n)}{\sqrt{n}}$$

Then, similarly:

$$D_t X(0) = \liminf_{n \to \infty} \frac{B(n)}{\sqrt{n}}$$

Thus, we simply need to show:

$$\limsup_{n \to \infty} \frac{B(n)}{\sqrt{n}} = \infty \quad \text{and} \quad \liminf_{n \to \infty} \frac{B(n)}{\sqrt{n}} = -\infty$$

Proposition 4.4.

$$\limsup_{n \to \infty} \frac{B(n)}{\sqrt{n}} = \infty \quad \textit{and} \quad \liminf_{n \to \infty} \frac{B(n)}{\sqrt{n}} = -\infty$$

Proof. Now, the reason we have this in a separate proposition, is that it is quite complicated. We need a few different ideas. First off, we will use Fatou's Lemma and the Hewitt-Savage 0-1 Law. The proof for these are out of the scope of this paper, but they will be extraordinarily useful to use.

Lemma 4.5 (Fatou's Lemma)**.** *Fatou's Lemma simply states that for non-negative measurable functions, f_n, on a measure space $(\Omega, \mathcal{F}, \mathbb{P})$:*

$$\int_\Omega \limsup f_n d\mathbb{P} \geq \limsup \int_\Omega f_n d\mathbb{P}.$$

Lemma 4.6 (Hewitt-Savage 0-1 Law)**.** *Set* $(X_i)_{i\geq 1}$ *to be a sequence of independent and identically-distributed sequence of random variables. Then, set the* σ*-algebra* ξ *to be all events which depend on* $(X_i)_{i\geq 1}$ *and whose probabilities are unchanged by finite permutations of* i*. Then, for all* $E \in \xi$*,* $\mathbb{P}(E) = 0$ *or* 1*.*

With this, we can now prove the proposition, and thus the theorem. To do this, we want to show that $B(n) > c\sqrt{n}$ with probability one. So:

$$\mathbb{P}(B(n) > c\sqrt{n} \text{ infinitely often}) = \mathbb{P}(\limsup_{n\to\infty} B(n) > c\sqrt{n}) \geq \limsup_{n\to\infty} \mathbb{P}(B(n) > c\sqrt{n})$$

Where c is some positive integer and the last step uses Fatou's Lemma. Then, by the scaling property, $\limsup_{n\to\infty} \mathbb{P}(B(n) > c\sqrt{n})$ becomes $\limsup_{n\to\infty} \mathbb{P}(B(1) > c)$, which is clearly positive, so the conditions for Fatou's Lemma are satisfied.

Now, set $X_n = B(n) - B(n-1)$. Then we have:

$$\mathbb{P}(B(n) > c\sqrt{n} \text{ infinitely often}) = \mathbb{P}(\Sigma_{i=1}^{n} X_i > c\sqrt{n} \text{ infinitely often})$$

Well, this is an exchangeable event, so we can use the Hewitt-Savage 0-1 Law to get that $\mathbb{P}(B(n) > c\sqrt{n}) = 1$. Then, if we take the intersection for all positive integers c, we get the first part of the proposition, i.e.:

$$\limsup_{n\to\infty} \frac{B(n)}{\sqrt{n}} = \infty$$

The second part we prove similarly.

□

And this proves the theorem. □

The other property of Brownian motion that we would like to discuss is the idea that the zero set is uncountable. That is, for a standard linear Brownian motion, $\{B(t) : t \geq 0\}$, the zero set, $\{t \geq 0 : B(t) = 0\}$ is uncountable. This result makes sense: standard Brownian motion is a random process with normally distributed increments with mean 0. So, throughout the process, it makes sense that the function would cross of 0 many times. In fact, an uncountable number of times. We will prove this in two steps. First, we show that the zero set is closed with no isolated points. Then, we show that such sets are uncountable.

Theorem 4.7. *For a Brownian motion* $\{B(t) : t \geq 0\}$*, denote the zero set by:* $O = \{t \geq 0 : B(t) = 0\}$*. Then,* O *is a closed set with no isolated points.*

Proof. For this proof, we first need two definitions:

Definition 4.8. A *stopping time for a sequence of random variables,* $X_1, X_2, \ldots$*, is a random variable itself, call it* ρ*, such that* $\forall t \in \mathbb{N}$ *the event that* $\rho = t$ *is dependent only on* $X_1, X_2, \ldots, X_t$*.*

Definition 4.9. The *strong Markov property states that given a Brownian motion* $\{B(t) : t \geq 0\}$*, for every finite stopping time* ρ*, the process* $\{B(\rho+t) - B(\rho) : t \geq 0\}$ *is a standard Brownian motion.*

Now, O is going to be closed from the simple fact that a Brownian motion is continuous. Then, to show that there are no isolated points, we consider, for each non-negative $q \in \mathbb{Q}$, the first member of the zero set after q: $\phi_q = inf\{t \geq q :$

$B(t) = 0\}$. So, ϕ_q has a finite stopping time, and since O is closed, the inf is a minimum. Now, using the strong Markov property, we know that for each q, ϕ_q is not an isolated point from the right.

Thus, we just have to show that the remaining points are not isolated from the left. Now, we say that for any $0 < t \in \mathbb{Z}$ different from ϕ_q is not an isolated point form the left. We can see this if we examine the sequence $q_n \uparrow t$ with $q \in \mathbb{Q}$. Then, set $m_n = \phi_{q_n}$. Then, $q_n \leq m_n < t$, so $m_n \uparrow t$. So, t is not isolated form the left. And thus, O has no isolated points. □

Another term for a closed set with no isolated sets is a perfect set. Thus, we only need to show that perfect sets are uncountable, and we will have our result.

Theorem 4.10. *Perfect sets are uncountable.*

Proof. Take a perfect set, lets call it O since we've already proved that to be a perfect set. Well, it cannot be finite because it consists only of accumulation points. If O is countable, then we could write it as $O = \{o_1, o_2, \ldots\}$.

Then we'll consider an interval, $I_1 = (o_1 - 1, o_1 + 1)$. This interval is clearly the neighborhood around o_1. Then, since o_1 is an accumulation point, I_1 has infinitely many elements of O.

Now, take an element of I_1, o_2 for example and construct I_2 so that $\overline{I_2} \subset I_1$ and $o_1 \notin \overline{I_2}$.

Then, since o_2 is an accumulation point in O, I_2 has infinitely many elements of O.

Now take o_3, and I_3 with $\overline{I_3} \subset I_2$ and $o_1, o_2 \notin \overline{I_3}$. Continue with this. Then, we will have I_n and o_n with $\overline{I_{n+1}} \subset I_n$, $o_j \notin I_n$ for all $0 < j < n$, $x_n \in I_n$.

Then consider the intersection of the closure of the intervals with O:

$$A = \cap_{i=1}^{n} \overline{I_i} \cap O$$

Each set $\overline{I_i} \cap O$ is closed and bounded, and thus compact with $\overline{I_{n+1}} \cap O \subset \overline{I_n} \cap O$, so A is nonempty. Yet, o_1 is not in V since $o_1 \notin \overline{I_2}$ and the same holds for $o_2, o_3, \ldots$. So, O must have some elements not listed in $\{o_1, o_2, o_3, \ldots$ and thus O is uncountable. □

So we have seen what Brownian motion is and a few interesting properties concerning Brownian motion. Other aspects concerning the topic, while outside the scope of this paper, are topics entirely worth pursuing. Brownian motion, and the probability behind it, have numerous uses across many different fields in mathematics, probability, economics, chemistry, and so forth.

Acknowledgments. It is a pleasure to thank my mentors, David Chudzicki and Zachary Madden, for all of their support, patience, and help with this project. They have been an incredibly useful resource and one that I will miss once this project has been completed.

References

[1] David Williams. Probability With Martingales Cambridge University Press. 1991.
[2] Peter Mörters and Yuval Peres Brownian Motion http://www.stat.berkeley.edu/ peres/bmbook.pdf
[3] Bert G. Wachsmuth Interactive Real Analysis http://pirate.shu.edu/ wachsmut/ira/index.html