An Introduction to Brownian Motion

Anubhav Nanavaty

August 27, 2018

The purpose of this lecture is to provide motivation for and construct Brownian motion, given a rudimentary knowledge of probability and measure theory. A few properties of Brownian motion will be explored as well.

Brownian motion is named after the botanist Robert Brown, who observed that pollen grains exhibited a 'jittery' sort of motion when in water. He initially believed the the pollen was alive, until he observed the same behavior exhibited by inorganic matter. The first serious attempt to understand Brownian motion was made by Albert Einstein.

1 How Einstein Saved the Atomic Theory of Matter

At the turn of the 20th century, many physicists began to realize that results from the field of thermodynamics appeared to conflict with those of Newtonian Mechanics. In particular, if all objects are made up of atoms, then heat is just generated by collisions between these atoms. Newton's laws seemed to predict that any heat transfer should spontaneously be reversible, i.e., heat could spontaneously flow from colder to hotter regions, because it is equally likely for a set of atom collisions to happen in the reverse direction.

The observations which led to the creation of the second law of thermodynamics, however, did not follow the predictions made by Netwonian mechanics. The second law, in fact, loosely states that heat *cannot* spontaneously flow from colder to hotter regions. Consider an ice cube melting. Clearly, the ice cube cannot spontaneously form, thus showing that any object in a closed system that is losing heat cannot spontaneously gain heat.

These conflicting ideas led scientists to some crazy conclusions. In fact, some began to question whether atoms really existed. Others went even further, and supported the de-legitimization of the field of classical mechanics. Einstein, however, by using the concept of Brownian motion and mathematically calculating its effects, was not only able to unite both thermodynamics and classical mechanics but also introduce quantum mechanics to the mix.

Einstein showed that, if atoms existed, then they would undergo Brownian motion, (i.e. random movements), and one could observe the effects of such motion. He went about this calculation in the following way:

Imagine a drop of ink is situated at $x = 0$ in a (infinitesimally) thin tube stretching to infinity at both ends. If $f(x,t)$ denotes the density of the ink particles at time $t \geq 0$ and position $x \in \mathbb{R}$. So, $f(x,0) = \delta_0(x)$ i.e. the Dirac mass at $x = 0$. Now the goal is to calculate $f(x,t)$, given that the *probability density* of the event that a particle moves from x to $x + y$ in time τ, i.e. $\rho(\tau, y)$. Observe that if we want to calculate $f(x, t+\tau)$, then we have to integrate over the all positions on the real line at time t, multiplied by the probability density of the particle moving to x in time τ. Or, mathematically:

$$f(x,t+\tau) = \int_{\mathbb{R}} f(x-y,t)\rho(\tau,y)dy$$

Note that, since $\rho(x,t)$ is the probability density, $\int_{\mathbb{R}} \rho(y,t)dy = 1$, and $\int_{\mathbb{R}} y\rho(y,t)dy = 0$. Further, let us suppose suppose $Var(\rho) = \int_{\mathbb{R}} y^2\rho(y,\tau) = D\tau$ is linear in τ, which can be loosely thought of as an assumption that 'how spread out' the ink particles are linearly increases with time. If we assume that $f \in C^\infty([\mathbb{R} - 0] \times \mathbb{R})$, we can fix t and Taylor expand f in x around x to get:

$$\begin{aligned}
f(x,t+\tau) &= \int_{\mathbb{R}} \Big[f(x,t) + f_x y + \frac{1}{2} f_{xx} y^2 + o(y^2)\Big]\rho(\tau,y)dy \\
&= \int_{\mathbb{R}} f(x,t)\rho(\tau,y)dy + \int_{\mathbb{R}} f_x y\rho(\tau,y)dy + \int_{\mathbb{R}} \frac{1}{2} f_{xx} y^2 \rho(\tau,y)dy + \int_{\mathbb{R}} o(y^2)\rho(\tau,y)dy \\
&= f(x,t) + \frac{D\tau}{2} f_{xx} + \int_{\mathbb{R}} o(y^2)\rho(\tau,y)dy \\
\frac{f(x,t+\tau) - f(x,t)}{\tau} &= \frac{D}{2} f_{xx} + \int_{\mathbb{R}} o(y^2)\rho(\tau,y)dy \\
\lim_{\tau\to 0} \frac{f(x,t+\tau) - f(x,t)}{\tau} &= \lim_{\tau \to 0} \Big[\frac{D}{2} f_{xx} + \int_{\mathbb{R}} o(y^2)\rho(\tau,y)dy\Big] \\
f_t &= \frac{D}{2} f_{xx}
\end{aligned}$$

Out pops the heat diffusion equation! One can check that the solution to this equation is given as

$$f(x,t) = \frac{1}{(2\pi Dt)^{\frac{1}{2}}} e^{\frac{-x^2}{2Dt}}$$

Further, this implies that the probability density of the event that the particle is at $f(x,t)$ is $N(0, Dt)$. Observe that the distribution has nothing to do with position, but only the time.
Einstein used arguments from physics to show that $D \sim \frac{1}{N_A}$, where N_A is Avogadro's constant

(i.e. the number of atoms per mole). This allowed physicists to experimentally verify that atoms existed, ending a century old debate.

2 Mathematical Treatment of Brownian Motion

2.1 The Weiner Process

Definition 2.1. A real-valued Stochastic Process $W(t)$ is a Weiner Process, or a Brownian Motion if

1. $W(0) = 0$ almost everywhere
2. $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ if $t \geq s \geq 0$
3. For any finite increasing sequence $t_1, \ldots, t_n$ The random variables $W(t_1), W(t_2) - W(t_1), \ldots, W(t_{i+1}) - W(t_i)$ are independent.

Observe that $\mathbb{E}[W(t)] = 0$ and $\mathbb{E}[W^2(t)] = t$ because $W(t) \sim N(0,t)$

This definition of Brownian motion captures exactly the motion that Einstein was describing. The following proposition gives us more intuition about the process:

Proposition 2.2. $\mathbb{E}[W(t)W(s)] = s \wedge t = \min\{t, s\}$, *for* $t, s \geq 0$

Proof. Assume $t \geq s$. By expanding,

$$\begin{aligned}
\mathbb{E}[W(t)W(s)] &= \mathbb{E}[(W(s)W(t) - W(s))W(s)] \\
&= \mathbb{E}[W^2(s)] + \mathbb{E}[(W(t) - W(s))W(s)] \\
&= s + \mathbb{E}[(W(t) - W(s))]\mathbb{E}[W(s)] \\
&= s = s \wedge t
\end{aligned}$$

This tells us that the covariance of $W(t), W(s)$ is just the minimum of the variance of each variable.

□

2.2 The Lévy Ciesielski Construction of Brownian Motion

The next question might be: How do we explicitly construct Brownian motion? In short, we will first mathematically define "white noise," and then show that a motion that varies according to "white noise" is in fact a Weiner process. The way we do this is:

1. Cleverly pick a basis, $\{\varphi_n\}_{n \in \mathbb{N}}$ of $L_2[0,1]$
2. Define "white noise," or $\xi(t)$, to be some sum $\sum_{n=0}^{\infty} A_n \varphi_n$, where A_n are independent, Gaussian, and have mean 0.

3. We can construct $W(t) = \int_0^t \xi(s)ds$

Note that this process revolves around cleverly picking a basis, and cleverly defining the A_i coordinates of the white noise in the basis.

Definition 2.3. We define *Haar functions* as the family $\{h_k : [0,1] \to \mathbb{R} \mid k \in \mathbb{N}\}$ such that

- $h_0(t) := 1$
- $h_1(t) = \begin{cases} 1 & \text{if } t \in [0, \frac{1}{2}] \\ -1 & \text{if } t \in (\frac{1}{2}, 1] \end{cases}$
- For k such that $2^n \leq k \leq 2^{n+1}$, we define
$$h_k(t) = \begin{cases} 2^{\frac{n}{2}} & \text{if } t \in [\frac{k-2^n}{2^n}, \frac{k-2^n+1/2}{2^n}] \\ -2^{\frac{n}{2}} & \text{if } t \in (\frac{k-2^n+1/2}{2^n}, \frac{k-2^n+1}{2^n}] \\ 0 & \text{otherwise} \end{cases}$$

Lemma 2.4. $\{h_k\}_{k\in\mathbb{N}}$ *forms an orthonormal basis of* $L_2(0,1)$

Proof. To show normality, $\int_0^1 h_k^2 = 2^n(\frac{1}{2^{n+1}} + \frac{1}{2^{n+1}}) = 1$.
To show orthogonality, observe that for $l > k$ either $h_k h_l = 0$ or h_k is constant on *supp*(h_l) (suppose $h_k = 2^n$), so $\int_0^1 h_l h_k = 2^n \int_0^1 h_l = 0$ Therefore $\int_0^1 h_l h_k = \delta_{kl}$
To show that $\{h_k\}_{k\in\mathbb{N}}$ is a basis, we see that if $\langle f, h_k \rangle = 0$ for all $k \in \mathbb{N}$, then $\int_{\frac{k}{2^{n+1}}}^{\frac{k+1}{2^{n+1}}} f = 0$ for all $k \in \{0, \ldots 2^{n+1}\}$. This tells us that for any two diadic rationals, $r, s \in [0,1]$, $\int_r^s f = 0$. Since these rationals are dense in $[0,1]$, we see that $f \equiv 0$, proving that $\{h_k\}_{k\in\mathbb{N}}$ is a basis. $\square$

We now integrate these functions to a set of 'bump' functions.

Definition 2.5. A *k-th Schauder function*, where $k \in \mathbb{N}$ is defined as

$$s_k(t) = \int_0^t h_k(s)ds$$

Observe that a k'th Shauder function is just a triangular bump of height $2^{-n/2-1}$, over $[\frac{k-2^n}{2^n}, \frac{k-2^n+1}{2^n}]$. Now, we will define $W(t) = \sum_{k=0}^{\infty} A_k s_k(t)$, where A_k are $\mathcal{N}(0,1)$ random variables, but first we need prove two lemmas to show that such a sum converges.

Lemma 2.6. *If* $\{A_k\}_{k\in\mathbb{N}}$ *are* $\mathcal{N}(0,1)$ *independent random variables, then*

$$|A_k| = O(\sqrt{\log k})$$

Proof. The goal is to use Borel-Cantelli. For $k \geq 2$, we see that

$$\begin{aligned}\mathbb{P}[|A_k| > x] &= \frac{2}{\sqrt{2\pi}} \int_x^\infty e^{\frac{s^2}{2}} \, ds \\ &\leq \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{4}} \int_x^\infty e^{\frac{s^2}{2}} \, ds \\ &\leq C e^{-\frac{x^2}{4}}\end{aligned}$$

Now, if we set $x = 4\sqrt{\log(k)}$, then plugging this in:

$$\mathbb{P}[|A_k| > 4\sqrt{\log(k)}] \leq \frac{C}{k^4}$$
$$\sum_{k=2}^\infty \mathbb{P}[|A_k| > 4\sqrt{\log(k)}] \leq \sum_{k=2}^\infty \frac{C}{k^4} < \infty$$

Then by Borel-Cantelli, $\mathbb{P}[|A_k| > 4\sqrt{\log(k)}]$ infinitely often is 0, showing us that the random variables A_k satisfy the lemma with probability 1. □

Lemma 2.7. *If $\{a_k\}_{k\in\mathbb{N}}$ is a sequence of real numbers such that $|a_k| = O(k^\delta)$ for a certain $\delta \leq \frac{1}{2}$, then $\sum_{k=0}^\infty a_k s_k(t)$ converges uniformly in $t \in [0,1]$*

Proof. Fix $\epsilon > 0$. Since $\delta \leq \frac{1}{2}$ choose $m \in \mathbb{N}$ such that

$$\sum_{m=n}^\infty (2^{n+1})^\delta 2^{-n/2+1} < \epsilon$$

Remembering that a k'th Shauder function is just a triangular bump of height $2^{-n/2-1}$, we see that

$$\sum_{m=n}^\infty (2^{n+1})^\delta \sup_{2^n \leq k \leq 2^{n+1}} s_k(t) < \epsilon$$

Further, since for $2^n \leq k \leq 2^{n+1}$ we have that $|a_k| \leq C(2^{n+1})\delta$, so

$$\sum_{k=2^m} a_n s_n(t) \leq \sum_{m=n}^\infty \sup_{2^n \leq k \leq 2^{n+1}} |a_k| \sup_{2^n \leq k \leq 2^{n+1}} s_k(t) < \epsilon$$

Giving us that the sum uniformly converges. □

Theorem 2.8. *$W(t) := \sum_{n=1}^\infty A_n(\omega) s_k(t)$ converges uniformly in $t \in [0,1]$ for almost every ω.*

Proof. Apply Lemma 2.7 to Lemma 2.6. □

Now, we are one lemma away from proving that $W(t)$ is a Brownian motion.

Lemma 2.9. $\sum_{k=0}^{\infty} s_k(s)s_k(t) = t \wedge s$ *for* $s, t \in [0,1]$.

Proof. Define $\varphi_s(\tau) = \begin{cases} 1 & \text{if } \tau \in [0,s] \\ 0 & \text{if } \tau \in (s,1] \end{cases}$.
Then, if $s \le t$, we see that $s = \int_0^1 \varphi_s \varphi_t d\tau$. Since $\{h_k\}$ is an orthonormal basis, we can write:

$$\begin{aligned} s = \langle \varphi_s, \varphi_t \rangle &= \langle \sum_{k=0}^{\infty} \langle \varphi_s, h_k \rangle h_k, \sum_{j=0}^{\infty} \langle \varphi_t, h_j \rangle, h_j \rangle \\ &= \sum_{k=0}^{\infty} \langle \langle \varphi_s, h_k \rangle h_k, \langle \varphi_t, h_k \rangle h_k \rangle \\ &= \sum_{k=0}^{\infty} \langle \varphi_s, h_k \rangle \langle \varphi_t, h_k \rangle \\ &= \sum_{k=0}^{\infty} \int_0^1 \varphi_s h_k(\tau) d\tau \int_0^1 \varphi_t h_k(\tau) d\tau \\ &= \sum_{k=0}^{\infty} \int_0^s h_k(\tau) d\tau \int_0^t h_k(\tau) d\tau \\ &= \sum_{k=0}^{\infty} s_k(s) s_k(t) \end{aligned}$$

□

Now, we can finally show that our construction satisfies the necessary requirements for a Weiner process:

Theorem 2.10. $W(t) = \sum_{k=0}^{\infty} A_n s_n(t)$ *is a Brownian motion for* $t \in [0,1]$

Proof. One can straightforwardly check that $W(0) = 0$ almost everywhere. In order to show that $W(t) - W(s) \sim \mathcal{N}(0, t - s)$ for $t \geq s$, we calculate the characteristic function:

$$\begin{aligned}
\phi_{W(t)-W(s)}(\lambda) &= \mathbb{E}[e^{i\lambda(W(t)-W(s))}] \\
&= \mathbb{E}[e^{i\lambda \sum_{k=0}^{\infty}[A_k(s_k(t)-s_k(s))]}] \\
&= \prod_{k=0}^{\infty} \mathbb{E}[e^{i\lambda A_k(s_k(t)-s_k(s))}] \qquad \text{(due to independence)} \\
&= \prod_{k=0}^{\infty} e^{-\frac{\lambda^2}{2}(s_k(t)-s_k(s))^2} \qquad \text{(since } A_n \sim \mathcal{N}(0,1)) \\
&= e^{-\frac{\lambda^2}{2}\sum_{k=0}^{\infty}(s_k(t)-s_k(s))^2} \\
&= e^{-\frac{\lambda^2}{2}\sum_{k=0}^{\infty}(s_k^2(t)-2s_k(t)s_k(s)+s_k^2(s)} \\
&= e^{-\frac{\lambda^2}{2}(t-2s+s)} \qquad \text{(lemma 2.9)} \\
&= e^{-\frac{\lambda^2}{2}(t-s)}
\end{aligned}$$

This gives us that the characteristic function of $W(t)-W(s)$ is the same as the Gaussian $\mathcal{N}(0, t-s)$, telling us by uniqueness of the characteristic function that $W(t) - W(s) \sim \mathcal{N}(0, t - s)$ for $t \geq s$.

A calculation similar to the one just performed tells us that for an increasing finite sequence $0 = t_0 < t_1 < \cdots < t_n$:

$$\begin{aligned}
\phi_{W(t_1),\ldots,W(t_{m-1})-W(t_{m-2}),W(t_m)-W(t_{m-1})}(x_1, \ldots x_n) &= \mathbb{E}[e^{i\sum_j \lambda_j(W(t_j)-W_(t_{j-1}))}] \\
&= \prod_{j=1}^{n} e^{\frac{-\lambda_j^2}{2}(t_j-t_{j-1})} \\
&= \prod_{k=1}^{n} \phi_{W(t_i)-W(t_{i-1})}(x_i)
\end{aligned}$$

Which tells us that $W(t_i) - W(t_{i-1})$ are independent variables. This concludes the proof that $W(t)$ is a Brownian motion. □

Note that we have only constructed Brownian motion for $t \in [0, 1]$. However, we can extend this motion for all time $t \geq 0$ in the following way, assuming that there exist countably many independent random variables that are $\mathcal{N}(0, 1)$. Due to countability, create a countably infinite collection of countably infinite random variables. Then, we can define a countably infinite amount of different Brownian motions, W^n, for $t \in [0, 1]$. Now, we can inductively define a Brownian motion on all $t \geq 0$ as $W(t)$ such that

$$W(t) := W(n-1) + W^n(t-(n-1))$$

In this way, we have exploited countability to explicitly construct Brownian motion for all $t \geq 0$.