Arcsine Laws in Brownian Motion and Random Walks

Kailash Ranganathan

May 18 2026

Abstract

In this paper, we study the occupation statistics of Brownian motion, in particular how a random path distributes its time throughout space. For a simple random walk and Brownian motion, we arrive at a surprising answer, in particular that such occupation statistics aren't necessarily concentrated around their 1/2-means as symmetry might intuitively suggest. Instead, we see that these statistics, such as the last zero before a fixed horizon and the total time spent on one side of the origin are given by arcsine laws, where the distributions are largest near the extremes. We first derive these laws from the discrete walk to Brownian motion, emphasizing the reflection principle, stopping-time ideas, and the passage through Donsker's invariance principle. Next, we look at fine-grained statistics such as Brownian local time, and explain the occupation-density formula, the downcrossing viewpoint, and Tanaka's formula as complementary ways of seeing why local time is the right refinement. Then, we discuss the Ray-Knight theorem that shows that the local-time profile itself evolves in a spatial variable as a squared Bessel process. Lastly, we connect these ideas to directions in modern research, in particular reinforced random walks and self-interacting diffusions, where occupation statistics themselves influence the laws of future motion.

1 Setup

1.1 Brownian Motion

We begin with a standard one-dimensional Brownian motion $(B_t)_{t\geq 0}$ that starts at the origin. By the properties of Brownian motion, we have that $B_0 = 0$, the paths $t \mapsto B_t$ are almost surely continuous, the increments are independent, and for $0 \leq s < t$,

$$B_t - B_s \sim \mathcal{N}(0, t-s).$$

Recall that Brownian motion is symmetric about zero, formally, the processes $(B_t)_{t\geq 0}$ and $(-B_t)_{t\geq 0}$ are equal in distribution. It also satisfies the scaling relation

$$(B_{ct})_{t\geq 0} \overset{d}{=} (\sqrt{c}\, B_t)_{t\geq 0}, \qquad c > 0.$$

We will use these two features, symmetry and scaling, repeatedly in looking at occupation statistics.

As per the lecture notes, for any $a \in \mathbb{R}$, we write

$$T_a = \inf\{t \geq 0 : B_t = a\}$$

as the first hitting time as level a. We also define the running maximum

$$M_t = \sup_{0 \leq s \leq t} B_s.$$

The reflection principle then relates the distribution of M_t to that of B_t, where we have the relation. In particular,

$$\mathbb{P}(M_t \geq a) = 2\mathbb{P}(B_t \geq a), \qquad a > 0.$$

as given in the lecture notes. We will also use this identity in our derivation of the arcsine laws.

1.2 Occupation Statistics

Our main exploration is into how Brownian motion distributes its time across space. The simplest such quantity we can consider is the amount of time the path spends above zero. For $t > 0$, define

$$A_t^+ = \int_0^t \mathbf{1}_{\{B_s > 0\}} \, ds.$$

as the *positive occupation time* of Brownian motion on $[0, t]$. Then consequently $\frac{A_t^+}{t}$ is the fraction of this interval during which the path is positive. By symmetry, its expectation is $1/2$ (the path does not spend more time on average above zero than below zero). Thus, a natural assumption would be to expect that the random fraction A_t^+/t would concentrate its value near $1/2$. Counterintuitively, the arcsine law shows that the opposite is true, that Brownian motion is unusually likely to spend almost all of the time interval on one side of zero.

We will consider two other information metrics as well, in particular the "last zero before time t" as

$$G_t = \sup\{s \leq t : B_s = 0\},$$

and the "time of the running maximum" as

$$\Theta_t = \arg\max_{0 \leq s \leq t} B_s.$$

The maximizer is almost surely unique, so Θ_t is well-defined with probability one. In this paper, we will show that all three quantities $\frac{A_t^+}{t}, \frac{G_t}{t}, \frac{\Theta_t}{t}$ have the same distribution, namely the arcsine law on $[0, 1]$. Even though they only measure occupation in a more "global" way, we already notice that the Brownian time allocation is highly nonuniform.

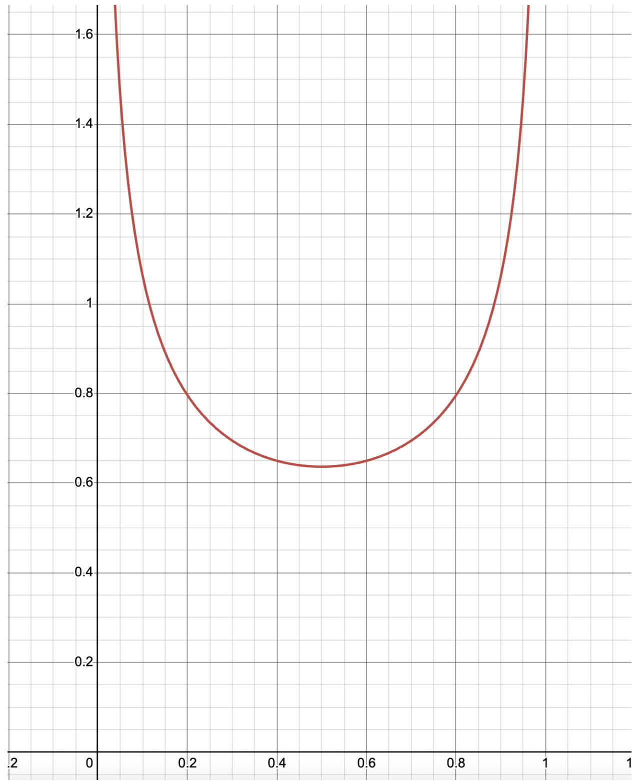

Figure 1: The arcsine density $\rho(x) = 1/(\pi\sqrt{x(1-x)})$ is smallest near $1/2$ and grows near 0 and 1. Through this we see that occupation statistics are not necessarily "balanced" on a typical path

2 Discrete Random Walk Derivation

We begin with a derivation of a discrete arcsine law for the random walk. Take $(S_n)_{n\geq 0}$ to be the simple symmetric random walk, where

$$S_0 = 0, \qquad S_n = X_1 + \cdots + X_n,$$

where $X_i \in \{-1, +1\}$ are iid uniform. Take the last return to zero before time $2n$:

We study the last return to zero before time $2n$:

$$G_{2n} := \max\{0 \leq k \leq 2n : S_k = 0\}.$$

G_{2n} is even as only even times can be a return to zero.

2.1 A reflection calculation

By explicitly summing over symmetric paths, we get that the probability of a walk having a 0 at step $2m$ is

$$u_{2m} := \mathbb{P}(S_{2m} = 0) = \binom{2m}{m} 2^{-2m}.$$

as each walk in this set has probability 2^{-m} and exactly m $+1$ increments and m -1 increments. We also need the probability that a walk started from zero does not return to zero during its first $2m$ steps:

$$q_{2m} := \mathbb{P}(S_1 \neq 0, S_2 \neq 0, \ldots, S_{2m} \neq 0).$$

The following lemma gives a derivation for q_{2m}:

Lemma 2.1 (No-return probability)**.** *For every $m \geq 0$,*

$$q_{2m} = u_{2m}.$$

Proof. A path that does not return to zero during its first $2m$ steps is either strictly positive or strictly negative following $n = 0$. By symmetry, these two sets have equal probability. Then, the next step is to determine $\mathbb{P}(S_1 > 0, \ldots, S_{2m} > 0)$. We condition on the final height $S_{2m} = 2k$, where $k \geq 1$. By the reflection principle, the number of $2m$-step paths that remain strictly positive after time 0 and end at height $2k$ is

$$\binom{2m-1}{m+k-1} - \binom{2m-1}{m+k}.$$

So, we have that as each path is equiprobable with probability 2^{-2m},

$$\begin{aligned}
\mathbb{P}(S_1 > 0, \ldots, S_{2m} > 0) &= 2^{-2m} \sum_{k=1}^{m} \left[\binom{2m-1}{m+k-1} - \binom{2m-1}{m+k} \right] \\
&= 2^{-2m} \binom{2m-1}{m} \\
&= \frac{1}{2} \binom{2m}{m} 2^{-2m} \\
&= \frac{1}{2}\, u_{2m}.
\end{aligned}$$

For the second equality, we have that the sum telescopes retaining only that term, and we also use the combinatorial relation

$$\binom{2m-1}{m} = \frac{1}{2} \binom{2m}{m}.$$

So, we have that

$$\mathbb{P}(S_1 > 0, \ldots, S_{2m} > 0) = \frac{1}{2} \mathbb{P}(S_{2m} = 0) = \frac{1}{2} u_{2m}.$$

Then, $q_{2m} = 2(\frac{1}{2}u_{2m}) = u_{2m}$, and so doubling for the negative case gives $q_{2m} = u_{2m}$. Mörters and Peres provides closely related reflection arguments for random walk and Brownian motion [3, Chapter 5]. $\square$

From this lemma, we can derive the law for the last-zero variable G_{2k} for the discrete random walk:

Proposition 2.2 (Discrete last-zero law)**.** *For* $0 \le k \le n$*,*

$$\mathbb{P}(G_{2n} = 2k) = u_{2k} u_{2n-2k}. \tag{1}$$

Proof. The event $\{G_{2n} = 2k\}$ has two independent parts. First, the walk must be at zero at time $2k$, an event with probability u_{2k}. Second, after time $2k$ it must avoid zero for the remaining $2n - 2k$ steps, an event with probability q_{2n-2k}. By the Markov property at the deterministic time $2k$ and the preceding lemma,

$$\mathbb{P}(G_{2n} = 2k) = u_{2k} q_{2n-2k} = u_{2k} u_{2n-2k}.$$

□

2.2 The arcsine limit

The discrete last-zero law already hints at the phenemenon we will see in the Brownian limit, namely that the walk is unusually likely to make its final return to zero either very early or very late, rather than near the midpoint. We now make this precise by rescaling the last-zero time by the total number of steps and passing to the large-n limit, where the exact combinatorial formula converges to the arcsine distribution. Stirling's formula gives an approximation for u_{2m}, as

$$u_{2m} = \binom{2m}{m} 2^{-2m} \sim \frac{1}{\sqrt{\pi m}}.$$

Hence, when k is proportional to n,

$$\mathbb{P}(G_{2n} = 2k) \approx \frac{1}{\pi \sqrt{k(n-k)}}.$$

Writing $k \approx nx$, this becomes

$$\mathbb{P}(G_{2n} = 2k) \approx \frac{1}{n} \cdot \frac{1}{\pi \sqrt{x(1-x)}}.$$

And so we note that the rescaled converges to a continuous RV with distribution given by the arcsine density, which we now state.

Theorem 2.3 (Discrete arcsine law)**.** *For every* $0 < a < b < 1$*,*

$$\mathbb{P}\left(\frac{G_{2n}}{2n} \in [a, b]\right) \longrightarrow \int_a^b \frac{1}{\pi \sqrt{x(1-x)}} \, \mathrm{d}x.$$

Proof. By Eq. (1),

$$\mathbb{P}\left(\frac{G_{2n}}{2n} \in [a, b]\right) = \sum_{k=\lceil an \rceil}^{\lfloor bn \rfloor} u_{2k} u_{2n-2k}.$$

Using Stirling's asymptotic uniformly away from the endpoints,

$$\sum_{k=\lceil an \rceil}^{\lfloor bn \rfloor} \frac{1}{\pi\sqrt{k(n-k)}} \sim \frac{1}{n} \sum_{k=\lceil an \rceil}^{\lfloor bn \rfloor} \frac{1}{\pi\sqrt{(k/n)(1-k/n)}}.$$

which converges to the desired integral in the high n limit as a Riemann sum. $\square$

Looking at this result, we note that the arcsine density does not appear from some complicated transform. Instead, it merely arises from multiplying two return probabilities, one for the portion of the walk before the last zero and one for the tail that avoids zeros afterward. Then both factors scale with inverse square root, which gives the $\frac{1}{\sqrt{x(1-x)}}$ resulting density shape.

3 Brownian arcsine laws

We now move to Brownian motion as the natural continuous analogue of the discrete random walk to show that the same arcsine behavior persists. In the discrete case, the last-zero fraction approaches a nonuniform limiting probability distribution concentrated near the endpoints, and we can show this in two complementary ways, either through going from random walk to Brownian motion via Donsker's invariance principle, or by deriving the Brownian arcsine law directly from reflection and scaling.

3.1 First arcsine law

We take

$$M_t := \sup_{0 \le s \le t} B_s$$

as the running maximum, and

$$\tau^* := \arg\max_{0 \le t \le 1} B_t$$

as the almost surely unique time at which the maximum on $[0,1]$ is attained. Mörters and Peres prove that the last zero before time one has the same distribution as τ^* [3, Theorem 5.26], so it suffices to compute the law for τ^*.

Theorem 3.1 (First Brownian arcsine law)**.** *The time τ^* of the Brownian maximum on $[0,1]$, and hence the last zero*

$$G_1 = \sup\{t \in [0,1] : B_t = 0\},$$

have distribution function

$$\mathbb{P}(\tau^* \le s) = \mathbb{P}(G_1 \le s) = \frac{2}{\pi} \arcsin(\sqrt{s}), \qquad 0 < s < 1.$$

Proof. Fix $0 < s < 1$. The event $\{\tau^* < s\}$ says that the maximum reached before time s is larger than the maximum reached after time s:

$$\mathbb{P}(\tau^* < s) = \mathbb{P}\left(\max_{0\le u\le s} B_u > \max_{s\le v\le 1} B_v\right).$$

Subtract B_s from both maxima. By time reversal, the process $(B_s - B_{s-u})_{0\le u\le s}$ is Brownian motion on $[0,s]$; hence

$$\max_{0\le u\le s}(B_u - B_s)$$

has the same distribution as a Brownian running maximum over time s.The process $B_s - B_{s-u}$ for $0 \le u \le s$ is a Brownian motion run for time s, and the process $B_{s+u} - B_s$ for $0 \le u \le 1-s$ is an independent Brownian motion run for time $1-s$. Thus

$$\mathbb{P}(\tau^* < s) = \mathbb{P}(M_s^{(1)} > M_{1-s}^{(2)}),$$

where the two maxima are independent.

By the reflection principle,

$$M_t \overset{d}{=} |B_t|.$$

By Brownian scaling,

$$M_s^{(1)} \overset{d}{=} \sqrt{s}\,|Z_1|, \qquad M_{1-s}^{(2)} \overset{d}{=} \sqrt{1-s}\,|Z_2|,$$

for independent standard normals Z_1, Z_2. Therefore

$$\mathbb{P}(\tau^* < s) = \mathbb{P}\left(\sqrt{s}|Z_1| > \sqrt{1-s}|Z_2|\right).$$

The Gaussian vector (Z_1, Z_2) is rotationally symmetric. Consider the geometry of the distributions: in polar coordinates, the event above corresponds to four angular wedges that pass through a total angle of $4\arcsin(\sqrt{s})$. Then dividing this by the full angle 2π yields

$$\mathbb{P}(\tau^* < s) = \frac{2}{\pi}\arcsin(\sqrt{s}).$$

So, the equality in law between τ^* and G_1 gives the desired last-zero statement. $\square$

The reflection principle is particularly handy in the above proof, reducing the maximum into a Gaussian calculation, where the arcsine law is derived through the geometry of the Gaussians.

3.2 Second Brownian arcsine law

The second Brownian arcsine law concerns the occupation variable

$$A_1 = \int_0^1 \mathbb{1}_{\{B_t>0\}}\,\mathrm{d}t.$$

Theorem 3.2 (Second Brownian arcsine law)**.** *The positive occupation time*

$$A_1 = \int_0^1 \mathbb{1}_{\{B_t > 0\}} \, \mathrm{d}t$$

has the arcsine distribution:

$$\mathbb{P}(A_1 \le x) = \frac{2}{\pi} \arcsin(\sqrt{x}), \qquad 0 < x < 1.$$

In this section I will provide a derivation of this theorem, in which our strategy is to begin with a derivation for the discrete random walk and move to Brownian motion to ultimately show that A_1 and τ are equal in distribution, meaning A_1 is also arcsine distributed.

We begin with the simple random walk (S_k). Define the number of positive partial sums

$$P_n := \#\{1 \le k \le n : S_k > 0\}.$$

Mörters and Peres prove an important identity, namely that P_n has the same distribution as the first time at which the walk reaches its overall maximum [3, Lemma 5.29]. The proof rearranges the increments of the walk: increments attached to positive partial sums are placed first in reverse order, and the remaining increments are placed afterward in forward order. Since all increment strings are equally likely, this rearrangement preserves the law.

The importance of this identity is that the right-hand side has a Brownian limit. For this, we need the additional result of Donsker's invariance principle, which allows us to move from the random walk to Brownian motion (this theorem is not in the written course notes, so we state it explicitly here).

Theorem 3.3 (Donsker's invariance principle)**.** *Take $(X_i)_{i \ge 1}$ to be independent and identically distributed random variables with*

$$\mathbb{E}[X_1] = 0, \qquad \mathrm{Var}(X_1) = 1,$$

and

$$S_n = X_1 + \cdots + X_n.$$

Define the linearly interpolated, diffusively rescaled process

$$W_n(t) = \frac{1}{\sqrt{n}} \left(S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor) X_{\lfloor nt \rfloor + 1} \right), \qquad 0 \le t \le 1.$$

Then, as $n \to \infty$,

$$W_n \Rightarrow B \qquad \text{in } C([0,1]),$$

where $B = (B_t)_{0 \le t \le 1}$ is standard Brownian motion and $C([0,1])$ is equipped with the uniform topology.

To now apply Donsker's invariance principle, we embed the random walk into a continuous path. For each n, define the linearly interpolated rescaled walk as above,

$$W_n(t) = \frac{1}{\sqrt{n}} \left(S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor)(S_{\lfloor nt \rfloor + 1} - S_{\lfloor nt \rfloor}) \right), \qquad 0 \leq t \leq 1.$$

Thus $W_n(k/n) = S_k/\sqrt{n}$ at the lattice times k/n, and between lattice times the path is joined by straight line segments. Donsker's invariance principle states that

$$W_n \Longrightarrow B \qquad \text{in } C[0,1],$$

where B is standard Brownian motion.

Now define

$$T(f) = \inf \left\{ t \in [0,1] : f(t) = \max_{0 \leq u \leq 1} f(u) \right\}$$

as the first time at which a continuous function f attains its maximum. Because each segment of W_n is linear, the maximum of W_n occurs at one of the interpolation points $\frac{k}{n}$. And so thus,

$$T(W_n) = \frac{1}{n} \min \left\{ 0 \leq k \leq n : S_k = \max_{0 \leq j \leq n} S_j \right\}.$$

The functional T is continuous at every continuous path having a unique maximizer, and Brownian motion has a unique time of maximum a.s.. And so we can apply the continuous mapping theorem to get the convergence

$$T(W_n) \Longrightarrow \tau^*,$$

where τ^* is our prior-defined time of max in the Brownian motion. But from the identity we established above, we know that P_n has the same distribution as the first index at which the random walk reaches its maximum. And so P_n and $T(W_n)$ are equal in distribution, and so

$$\frac{P_n}{n} \Longrightarrow \tau^*.$$

On the other hand, the functional

$$f \longmapsto \int_0^1 \mathbb{1}_{\{f(t) > 0\}} \, \mathrm{d}t$$

is continuous at any continuous function whose zero set has Lebesgue measure zero. Brownian motion satisfies continuity almost surely, so we then get the convergence

$$\frac{P_n}{n} \Longrightarrow A_1.$$

And we already showed τ^* is arcsine distributed in the first Brownian arcsine law, and so this completes the proof for A_1's arcsine distribution.

3.3 Intuition for Arcsine Distribution

In the past two sections, we uncover an overall pattern, that while in expectation the Brownian motion seems to hover around symmetry-obeying values (ie. $\mathbb{E}[A_1] = 1/2$), the arcsine law instead suggests that this mean is not representative of a typical path. Rather, a path that spends almost all of $[0, 1]$ positive and a path that spend almost all of its time negative balance each other in expectation, but the distribution is U-shaped to bias towards such paths.

A useful intuition for why this is is that Brownian motion does not cross zero at some steady frequency. Once the motion moves away from the origin, there is a meaningful chance that it remains on one side for a long stretch of time, and the probability of avoiding zero decays only on an inverse-square-root scale. For the last-zero statistic, this means it is relatively likely that the path makes its final return very early and then stays away from zero, or keeps returning until very late, as a last zero near the middle requires substantial zero-free behavior on both sides and is less favored. Likewise, the positive occupation time can be explained by this same behavior, as a path that spends almost all of the interval either above or below zero is more "typical" than one that repeatedly balances its time evenly between the two. In all these cases, the arcsine density captures this tendency of sorts towards long one-sided excursions.

4 Local Time

Up till now, we have focused on "global" values of the Brownian motion, in particular the time of maximum of the process and the total time it spends above the zero-line. In this section, we will look at local time as a more fine-grained occupation statistic, namely the amount of time Brownian motion will spend near a certain level x.

4.1 Occupation measures and densities

To make this notion precise, we define a measure for the time spent by Brownian motion across all different levels in space. In particular, we can define such an object that records the occupation of every Borel set $A \subseteq \mathbb{R}$ up to a time t. Then, local time is the density of this "occupation measure" wrt the Lebesgue measure.

We start by defining the Brownian occupation measure for a fixed t as

$$\mu_t(E) := \int_0^t \mathbb{1}_{\{B_s \in E\}} \, \mathrm{d}s.$$

In one dimension, this random measure is absolutely continuous with respect to Lebesgue measure, and its density is local time. Formally, local time at level x may be suggested by the limit

$$L_t^x = \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{\{|B_s - x| < \varepsilon\}} \, \mathrm{d}s. \tag{2}$$

The limit has a jointly continuous version in (t, x), and for sake of length developing this local-time field is deferred to Mörters and Peres in Chapter 6 [3].

Definition 4.1 (Brownian local time)**.** A jointly continuous random field $(L_t^x)_{t\geq 0, x\in\mathbb{R}}$ is Brownian local time if, for nonnegative measurable functions f, or more generally for integrable f,

$$\int_0^t f(B_s)\,\mathrm{d}s = \int_{\mathbb{R}} f(x) L_t^x\,\mathrm{d}x. \tag{3}$$

Equation Eq. (3) gives us the occupation-density formula, as

$$\mu_t(\,\mathrm{d}x) = L_t^x\,\mathrm{d}x.$$

Note we can re-arrive at the second arcsine variable as the mass of this density on the positive half-line:

$$A_1 = \mu_1((0,\infty)) = \int_0^\infty L_1^x\,\mathrm{d}x.$$

4.2 Downcrossing interpretation

The approximation in Eq. (2) has an equivalent pathwise interpretation through downcrossings. Take $a < b$, and define $D(a,b,t)$ as the count of the number of completed downcrossings of the interval $[a,b]$ before time t. Mörters and Peres show that if $a_n \uparrow 0$ and $b_n \downarrow 0$, then

$$2(b_n - a_n)D(a_n, b_n, t) \longrightarrow L_t^0 \tag{4}$$

almost surely, for every fixed $t > 0$ [3, Theorem 6.1].

We can offer some intuition on the factor $2(b-a)$: Brownian scaling says that a crossing of an interval of width $b-1$ will last on the order of $(b-a)^2$ units of time. So, a complete cycle consisting of a downcrossing and an upcrossing will contributes roughly $2(b-a)^2$ units of occupation inside the interval. To convert occupation time into occupation density, one divides by the interval width $b-a$. This then gives the scale

$$2(b-a)D(a,b,t),$$

which is precisely the quantity in Eq. (4).

5 Ray–Knight Theorem

The previous section introduced local time as a way to record Brownian occupation level by level. For a fixed time t, the map $x \to L_t^x$ describes how "intensely" the path has visited each spatial location. This profile ostensibly is a complicated summary of a single Brownian trajectory, but the Ray-Knight theorem shows that given a hitting time, this map instead has the familiar structure of a planar Brownian motion in the level variable x.

5.1 The squared Bessel process

As covered in Lecture 22 of the course notes [2], let

$$W_x = (W_x^{(1)}, W_x^{(2)}), \qquad x \geq 0,$$

be standard planar Brownian motion, indexed here by the parameter x. Define

$$Z_x := |W_x|^2 = (W_x^{(1)})^2 + (W_x^{(2)})^2.$$

The process $(Z_x)_{x\geq 0}$ is the squared two-dimensional Bessel process. For each fixed $x > 0$, Z_x has an exponential distribution with mean $2x$, and its infinitesimal generator is

$$\mathcal{G}f(z) = 2zf''(z) + 2f'(z).$$

Lecture 22 derives these facts directly, using the Gaussian density in polar coordinates and Itô's formula [2].

The form of this generator gives some intuition for why such a process can arise from local times. The coefficient $2z$ in front of f'' means that the random fluctuations become larger when the current value z is larger. This fits the local-time intuition, where (thinking very abstractly) if Brownian motion has already accumulated a large amount of local time at one level, then there are many excursions from that level that can contribute occupation to nearby levels. As the first-order drift term $2f'(z)$ supplies an additional deterministic growth component, we see all in all a branching-like evolution of "occupation mass" across space.

5.2 Theorem Statement

Witht his background, we now state the Ray–Knight theorem. Fix $a > 0$, let Brownian motion start from $B_0 = a$ and stop when it first hits zero, under the hitting time $T_0 := \inf\{t \geq 0 : B_t = 0\}$.

Theorem 5.1 (Ray–Knight theorem). *The local-time profile accumulated before hitting zero satisfies*

$$(L_{T_0}^x : 0 \leq x \leq a) \stackrel{d}{=} (|W_x|^2 : 0 \leq x \leq a).$$

Equivalently, the function

$$x \to L_{T_0}^x$$

has the law of a squared two-dimensional Bessel process in the spatial variable x.

This theorem provides a neat connection between x as a spatial level of the original one-dimensional Brownian motion under the local time profile and as a time parameter of a new stochastic process. In this sense, when stopping a Brownian motion at T_0, its accumulated occupation across levels behaves like the squared radial part of a planar Brownian process evolving in space.

5.3 Proof sketch

The proof in the lecture notes proceeds by approximating local time through downcrossings of very small intervals [2, Lecture 22]. The starting point is the interval hitting formula

$$\mathbb{P}_y(T_r < T_\ell) = \frac{y-\ell}{r-\ell}, \qquad \ell < y < r,$$

that follows from optional stopping for Brownian motion.

Fix $0 < x < a$, and let $D_n(x)$ denote the number of downcrossings of the interval

$$[x - 1/n, x]$$

before the stopping time T_0. The downcrossing approximation to local time gives

$$\frac{2}{n} D_n(x) \to L^x_{T_0}.$$

Now when one such downcrossing ends at $x - 1/n$, the path from that point either returns to x, which then allows another downcrossing to occur, or it hits zero first, which stops the process. By the above hitting-probability formula, we have

$$\mathbb{P}_{x-1/n}(T_0 < T_x) = \frac{1}{nx}.$$

After each downcrossing, the process could either return to x and make another downcrossing, or it could hit 0 before returning to x. By this formula, after every downcrossing there is some chance the process goes to 0, and likewise probability $1 - \frac{1}{nx}$ that it returns to x and starts another downcrossing. Thus, the number of downcrossings $D_n(x)$ follows a geometric distribution with parameter $1/(nx)$ (for k downcrossings, first have $k-1$ "failed" trials that create downcrossings, then after the kth, return to zero). Then after rescaling, by convergence of geometric distributions in the high-n limit

$$\frac{2}{n} D_n(x) \Longrightarrow \text{Exponential}(\lambda = \frac{1}{2x}).$$

This matches the distribution of $|W_x|^2$ for a fixed x, as the squared radius of planar Brownian motion at time x is exponential with mean $2x$.

It remains to understand the dependence between difference levels. To compare local time at x and at $x + h$, we can decompose the Brownian path into excursions away from level x. Some of these excursions rise high enough to reach $x + h$, and each such excursion contributes additional local time there. Conditional on the local-time mass already accumulated at level x, these contributions arise from independent excursion pieces. In the limit, their Laplace transform matches the transition law of the squared Bessel process.

Taking this comparison across finitely many levels gives that $(L^{x_1}_{T_0}, \ldots, L^{x_m}_{T_0})$ and $(|W_{x_1}|^2, \ldots, |W_{x_m}|^2)$ are equal in distribution. And lastly, continuity of the local-time profile and of the squared Bessel process upgrades this finite-dimensional equality to the full process-level equality of the theorem [2, Lecture 22].

6 Self-interacting Processes

So far, occupation has been purely descriptive, where given a path, we can count how long it has spent in certain regions, refine that into local time, and then study the resulting field. In self-interacting stochastic processes, the occupation profile influences future paths through different mechanisms.

In this section, we will explain the connection between existing occupation statistics we've discussed and self-interacting processes through the examples of the discrete reinforced walk and a continuous self-interacting diffusion.

6.1 Vertex-Reinforced Random Walk

Take X_n to be a nearest-neighbor walk on a graph and define its discrete local time at a vertex v by

$$L_n(v) := 1 + \sum_{k=0}^{n} \mathbb{1}_{\{X_k = v\}}.$$

A vertex-reinforced random walk chooses its next step with probabilities biased toward vertices that have been visited often. On a graph with neighbor relation $y \sim x$, the transitions are given by

$$\mathbb{P}(X_{n+1} = y \mid \mathcal{F}_n, X_n = x) = \frac{\mathbb{1}_{\{y \sim x\}} L_n(y)}{\sum_{z \sim x} L_n(z)}. \tag{5}$$

Pemantle introduced vertex-reinforced random walk in this form of occupation-dependent transition rule [4]. Here, the transition dynamics themselves depend on the occupation count of every vertex.

This "feedback" into the transition dynamics radically changes the long-term behavior. For the one-dimensional integer lattice, Tarrès proved that linearly vertex-reinforced random walk eventually localizes on five sites almost surely [5]. The actual localization theorem is beyond the scope of this paper, but we notice at a high level that with positive reinforcement, occupation is converted into a trapping mechanism of sorts for states rather than just a description of the random walk.

6.2 Self-Interacting Diffusions

A smooth continuous analog of the above idea is studied in the work of "self-interacting diffusion" by Benaïm, Ledoux, and Raimond [1]. They use the occupation measure

$$\mu_t := \frac{1}{t} \int_0^t \delta_{X_s} \, \mathrm{d}s$$

within the drift term of a stochastic differential equation. We can write the transition dynamics as

$$\mathrm{d}X_t = \mathrm{d}B_t - \nabla(V * \mu_t)(X_t) \, \mathrm{d}t, \tag{6}$$

where V is an interaction potential that is averaged against the occupation measure. If we look at the convolution term,

$$(V * \mu_t)(x) = \int V(x-y)\,\mu_t(\,\mathrm{d}y)$$

we can interpret this as how strongly the point x interacts with the regions that have been visited in the past. Then, the gradient $\nabla(V * \mu_t)(X_t)$ acts as a force determined by the accumulated spatial history of the path. Different forms for V can enable different behaviors, for example pulling the process back towards very visited regions or pushing away from them.

In the self-interacting diffusions setup, though the path remains random, the long-time evolution of the occupation measure can be compared to a deterministic dynamical system on the space of probability measures. Specifically, the empirical distribution of process occupation evolves asymptotically into a nonrandom flow.

Benaïm, Ledoux, and Raimond study broad classes of self-interacting diffusions of this type, where the drift depends on the current position and on the normalized occupation measure μ_t [1]. A key point of their work is that, although the path itself remains random, the long-time evolution of the occupation measure can be compared to a deterministic dynamical system on the space of probability measures. In other words, the empirical distribution of where the process has spent its time evolves, asymptotically, according to a more structured nonrandom flow.

6.3 True self-repelling motion

The strongest conceptual connection to local time comes from the true self-repelling motion of Tóth and Werner [6]. Their process is a one-dimensional continuous motion whose future behavior is shaped by its own accumulated occupation profile. Intuitively, the process is pushed away from regions it has already visited heavily and is drawn toward less occupied nearby regions.

A useful visualization is that the occupation profile forms a landscape over space as it evolves. As the process spends time near a point, the landcsape "rises," and then the motion tends to move in directions where this profile decreases, producing a local self-repulsive effect. Hence, the interaction is not determined by just the position of the point as in Brownian motion but rather by the entire landscape, and so the process is non-Markovian in its position alone (however, we can recover a Markovian description by enlarging the state space to include the full profile).

So in true self-repelling motion, we see that the full density actively influences the future trajectory. In this sense, the process turns local occupation from an observable into a dynamical force.

7 Conclusion

In this paper, we derive the arcsine laws for occupation statistics of Brownian motion, demonstrating that even though the process is symmetric about zero, its last zero before time one and its total time spent positive are not typically close to the midpoint of the interval. Instead, both are governed by the arcsine distribution, which favors values near the extremes. The discrete random walk derivation makes this phenomenon concrete, while the Brownian proofs show how it persists in the continuous limit.

Local time arises from a more detailed description of occupation, instead measuring how occupation is distributed across individual spatial levels. The occupation-density formula then makes this refinement precise, and both the downcrossing construction and Tanaka's formula demonstrate this local time is a natural object in Brownian analysis from multiple angles.

Next, the Ray-Knight theorem reveals an even deeper structure, where at a hitting time, the entire local-time profile becomes a familiar Markov process in the spatial variable. As a connection to more modern research, self-interacting processes push this idea one step further, where in these models, occupation no longer merely summarize where the path has been, but instead influence where the path is to go next.

8 Acknowledgements

The derivation of the arcsine laws follows closely from the discussion of occupation statistics in Mörters and Peres [3, Chapter 5]. We base our discussion in the material covered in the lecture notes [2], in particular the Brownian motion lectures and the treatment of the Ray-Knight theorem in Lecture 22. Our discussion of local time draws primarily from Mörters and Peres [3, Chapters 6-7], especially their presentations of occupation densities, downcrossings, and Tanaka's formula. The final section on self-interacting stochastic processes is intended as a brief extension into research areas, in particular the works of Pemantle, Tarrès, Benaïm-Ledoux-Raimond, and Tóth-Werner cited throughout.

References

[1] Michel Benaïm, Michel Ledoux, and Olivier Raimond. "Self-Interacting Diffusions". In: *Probability Theory and Related Fields* 122.1 (2002), pp. 1–41. DOI: 10.1007/s004400100161.

[2] Alan Hammond. *Probability Theory: Mathematics C218B/Statistics C205B*. Spring 2026 lecture notes. 2026.

[3] Peter Mörters and Yuval Peres. *Brownian Motion*. Paperback edition. Cambridge University Press, 2015.

[4] Robin Pemantle. "Vertex-Reinforced Random Walk". In: *Probability Theory and Related Fields* 92.1 (1992), pp. 117–136. DOI: 10.1007/BF01205239.

[5] Pierre Tarrès. "Vertex-Reinforced Random Walk on $\mathbb{Z}$ Eventually Gets Stuck on Five Points". In: *The Annals of Probability* 32.3B (2004), pp. 2650–2701. DOI: 10.1214/009117907000000694.

[6] Bálint Tóth and Wendelin Werner. "The True Self-Repelling Motion". In: *Probability Theory and Related Fields* 111.3 (1998), pp. 375–452. DOI: 10.1007/s004400050172.