Sample path properties of Brownian motion

by Peter Mörters (University of Bath)

This is a set of lecture notes based on a graduate course given at the Berlin Mathematical School in September 2011. The course is based on a selection of material from my book with Yuval Peres, entitled *Brownian motion*, which was published by Cambridge University Press in 2010.

1 Does Brownian motion hit points?

1.1 What is Brownian motion and why are we interested?

Much of probability theory is devoted to describing the *macroscopic picture* emerging in random systems defined by a host of *microscopic random effects*. Brownian motion is the macroscopic picture emerging from a particle moving randomly in d-dimensional space without making very big jumps. On the microscopic level, at any time step, the particle receives a random displacement, caused for example by other particles hitting it or by an external force, so that, if its position at time zero is S_0, its position at time n is given as $S_n = S_0 + \sum_{i=1}^n X_i$, where the displacements $X_1, X_2, X_3, \ldots$ are assumed to be independent, identically distributed random variables with values in $\mathbb{R}^d$. The process $\{S_n : n \geqslant 0\}$ is a random walk, the displacements represent the microscopic inputs. It turns out that not all the features of the microscopic inputs contribute to the macroscopic picture. Indeed, if they exist, only the *mean* and *covariance* of the displacements are shaping the picture. In other words, all random walks whose displacements have the same mean and covariance matrix give rise to the same macroscopic process, and even the assumption that the displacements have to be independent and identically distributed can be substantially relaxed. This effect is called *universality*, and the macroscopic process is often called a *universal object*. It is a common approach in probability to study various phenomena through the associated universal objects.

If the jumps of a random walk are sufficiently tame to become negligible in the macroscopic picture, in particular if it has finite mean and variance, any continuous time stochastic process $\{B(t) : t \geqslant 0\}$ describing the macroscopic features of this random walk should have the following properties:

(1) for all times $0 \leqslant t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n$ the random variables

$$B(t_n) - B(t_{n-1}),\, B(t_{n-1}) - B(t_{n-2}),\, \ldots,\, B(t_2) - B(t_1)$$

are independent; we say that the process has *independent increments*,

(2) the distribution of the increment $B(t+h) - B(t)$ does not depend on t; we say that the process has *stationary increments*,

(3) the process $\{B(t)\colon t \geqslant 0\}$ has almost surely continuous paths.

It follows (with some work) from the central limit theorem that these features imply that there exists a vector $\mu \in \mathbb{R}^d$ and a matrix $\Sigma \in \mathbb{R}^{d\times d}$ such that

(4) for every $t \geqslant 0$ and $h \geqslant 0$ the increment $B(t+h) - B(t)$ is multivariate normally distributed with mean $h\mu$ and covariance matrix $h\Sigma\Sigma^{\mathrm{T}}$.

Hence any process with the features (1)-(3) above is characterised by just three parameters,

- the *initial distribution*, i.e. the law of $B(0)$,
- the *drift vector* μ,
- the *diffusion matrix* Σ.

If the drift vector is zero, and the diffusion matrix is the identity we say the process is a *Brownian motion.* If $B(0) = 0$, i.e. the motion is started at the origin, we use the term *standard Brownian motion.*

Suppose we have a standard Brownian motion $\{B(t)\colon t \geqslant 0\}$. If X is a random variable with values in $\mathbb{R}^d$, μ a vector in $\mathbb{R}^d$ and Σ a $d \times d$ matrix, then it is easy to check that $\{\tilde{B}(t)\colon t \geqslant 0\}$ given by

$$\tilde{B}(t) = X + \mu t + \Sigma B(t), \text{ for } t \geqslant 0,$$

is a process with the properties (1)-(4) with initial distribution X, drift vector μ and diffusion matrix Σ. Hence the macroscopic picture emerging from a random walk with finite variance can be fully described by a standard Brownian motion.

1.2 Two basic properties of Brownian motion

A key property of Brownian motion is its *scaling invariance*, which we now formulate. We describe a transformation on the space of functions, which changes the individual Brownian random functions but leaves their distribution unchanged.

Lemma 1.1 (Scaling invariance)**.** *Suppose $\{B(t)\colon t \geqslant 0\}$ is a standard Brownian motion and let $a > 0$. Then the process $\{X(t)\colon t \geqslant 0\}$ defined by $X(t) = \frac{1}{a}B(a^2t)$ is also a standard Brownian motion.*

Proof. Continuity of the paths, independence and stationarity of the increments remain unchanged under the scaling. It remains to observe that $X(t) - X(s) = \frac{1}{a}(B(a^2t) - B(a^2s))$ is normally distributed with expectation 0 and variance $(1/a^2)(a^2t - a^2s) = t - s$. ■

A property that is immediate is that Brownian motion starts afresh at any given fixed time. This statement also holds for a class of random times called *stopping* *times*. A random variable T with values in $[0,\infty]$, defined on a probability space with filtration $(\mathcal{F}(t)\colon t \geqslant 0)$ is called a **stopping time** with respect to $(\mathcal{F}(t)\colon t \geqslant 0)$ if $\{T \leqslant t\} \in \mathcal{F}(t)$, for every $t \geqslant 0$. In the case of Brownian motion we choose the filtration given by

$$\mathcal{F}^+(t) = \bigcap_{\varepsilon>0} \sigma\{B(s)\colon s < t+\varepsilon\}.$$

This ensure that first entry times into open or closed sets are always stopping times. We can now state the *strong Markov property* for Brownian motion, which was rigorously established by Hunt and Dynkin in the mid fifties.

Theorem 1.2 (Strong Markov property)**.** *For every almost surely finite stopping time T, the process*

$$\{B(T+t) - B(T)\colon t \geqslant 0\}$$

is a standard Brownian motion independent of $\mathcal{F}^+(T)$.

We will see many applications of the strong Markov property later, however, the next result, the reflection principle, is particularly interesting. The reflection principle states that Brownian motion reflected at some stopping time T is still a Brownian motion.

Theorem 1.3 (Reflection principle)**.** *If T is a stopping time and $\{B(t)\colon t \geqslant 0\}$ is a standard Brownian motion, then the process $\{B^*(t)\colon t \geqslant 0\}$ called* **Brownian motion reflected at** *T and defined by*

$$B^*(t) = B(t)1_{\{t\leqslant T\}} + (2B(T) - B(t))1_{\{t>T\}}$$

is also a standard Brownian motion.

Proof. If T is finite, by the strong Markov property both paths

$$\{B(t+T) - B(T)\colon t \geqslant 0\} \text{ and } \{-(B(t+T) - B(T))\colon t \geqslant 0\} \tag{1}$$

are Brownian motions and independent of the beginning $\{B(t)\colon 0 \leqslant t \leqslant T\}$. The process arising from glueing the first path in (1) to $\{B(t)\colon 0 \leqslant t \leqslant T\}$ and the process arising from glueing the second path in (1) to $\{B(t)\colon 0 \leqslant t \leqslant T\}$ have the same distribution. The first is just $\{B(t)\colon t \geqslant 0\}$, the second is $\{B^*(t)\colon t \geqslant 0\}$, as introduced in the statement. ■

Now we apply the reflection principle in the case of *linear* Brownian motion. Let $M(t) = \max_{0\leqslant s\leqslant t} B(s)$. A priori it is not at all clear what the distribution of this random variable is, but we can determine it as a consequence of the reflection principle.

Lemma 1.4. $\mathbb{P}_0\{M(t) > a\} = 2\mathbb{P}_0\{B(t) > a\} = \mathbb{P}_0\{|B(t)| > a\}$ *for all $a > 0$.*

Proof. Let $T = \inf\{t \geqslant 0\colon B(t) = a\}$ and let $\{B^*(t)\colon t \geqslant 0\}$ be Brownian motion reflected at the stopping time T. Then

$$\{M(t) > a\} = \{B(t) > a\} \cup \{M(t) > a,\ B(t) \leqslant a\}.$$

This is a disjoint union and the second summand coincides with event $\{B^*(t) \geqslant a\}$. Hence the statement follows from the reflection principle. ■

1.3 The range of planar Brownian motion has zero area

Suppose $\{B(t) : t \geqslant 0\}$ is planar Brownian motion. We denote the Lebesgue measure on $\mathbb{R}^d$ by $\mathcal{L}_d$. In this section we prove Lévy's theorem on the area of planar Brownian motion.

Theorem 1.5 (Lévy 1940)**.** *Almost surely,* $\mathcal{L}_2(B[0,1]) = 0$.

For a set $A \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$ we write $A + x := \{a + x\colon a \in A\}$.

Lemma 1.6. *If $A_1, A_2 \subset \mathbb{R}^2$ are Borel sets with positive area, then*

$$\mathcal{L}_2\big(\{x \in \mathbb{R}^2\colon \mathcal{L}_2(A_1 \cap (A_2 + x)) > 0\}\big) > 0.$$

The proof of the lemma is an **exercise**. To prove Theorem 1.5 we let $X = \mathcal{L}_2(B[0,1])$ denote the area of $B[0,1]$. First we check that $\mathbb{E}[X] < \infty$. Note that $X > a$ only if the Brownian motion leaves the square centred in the origin of side length $\sqrt{a}$. Hence, using Theorem 1.4 and a standard tail estimate for standard normal variables,

$$\mathbb{P}\{X > a\} \leqslant 2\,\mathbb{P}\{\max_{t\in[0,1]} |W(t)| > \sqrt{a}/2\} = \leqslant 8\,\mathbb{P}\{W(1) > \sqrt{a}/2\} \leqslant 8e^{-a/8},$$

for $a > 1$, where $\{W(t)\colon t \geqslant 0\}$ is standard one-dimensional Brownian motion. Hence,

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}\{X > a\}\, da \leqslant 4 \int_1^\infty e^{-a/8} da + 1 < \infty.$$

Note that $B(3t)$ and $\sqrt{3}B(t)$ have the same distribution, and hence

$$\mathbb{E}\mathcal{L}_2(B[0,3]) = 3\mathbb{E}\mathcal{L}_2(B[0,1]) = 3\mathbb{E}[X]\,.$$

Note that we have $\mathcal{L}_2(B[0,3]) \leqslant \sum_{j=0}^2 \mathcal{L}_2(B[j,j+1])$ with equality if and only if for $0 \leqslant i < j \leqslant 2$ we have $\mathcal{L}_2(B[i,i+1] \cap B[j,j+1]) = 0$. On the other hand, for $j = 0,1,2$, we have $\mathbb{E}\mathcal{L}_2(B[j,j+1]) = \mathbb{E}[X]$ and

$$3\mathbb{E}[X] = \mathbb{E}\mathcal{L}_2(B[0,3]) \leqslant \sum_{j=0}^2 \mathbb{E}\mathcal{L}_2(B[j,j+1]) = 3\mathbb{E}[X]\,,$$

whence, almost surely, the intersection of any two of the $B[j,j+1]$ has measure zero. In particular, $\mathcal{L}_2(B[0,1] \cap B[2,3]) = 0$ almost surely.

Now we can use the Markov property to define two Brownian motions, $\{B_1(t)\colon t \in [0,1]\}$ by $B_1(t) = B(t)$, and $\{B_2(t)\colon t \in [0,1]\}$ by $B_2(t) = B(t+2)-B(2)+B(1)$. The random variable $Y := B(2) - B(1)$ is independent of both Brownian motions. For $x \in \mathbb{R}^2$, let $R(x)$ denote the area of the set $B_1[0,1] \cap (x + B_2[0,1])$, and note that $\{R(x)\colon x \in \mathbb{R}^2\}$ is independent of Y. Then

$$0 = \mathbb{E}[\mathcal{L}_2(B[0,1] \cap B[2,3])] = \mathbb{E}[R(Y)] = (2\pi)^{-1} \int_{\mathbb{R}^2} e^{-|x|^2/2}\, \mathbb{E}[R(x)]\, dx,$$

where we are averaging with respect to the Gaussian distribution of $B(2)-B(1)$. Thus, for $\mathcal{L}_2$-almost all x, we have $R(x) = 0$ almost surely and hence,

$$\mathcal{L}_2\big(\{x \in \mathbb{R}^2\colon R(x) > 0\}\big) = 0, \qquad \text{almost surely.}$$

From Lemma 1.6 we get that, almost surely, $\mathcal{L}_2(B[0,1]) = 0$ or $\mathcal{L}_2(B[2,3]) = 0$. The observation that $\mathcal{L}_2(B[0,1])$ and $\mathcal{L}_2(B[2,3])$ are identically distributed and independent completes the proof that $\mathcal{L}_2(B[0,1]) = 0$ almost surely.

Remark 1.7. How big is the path of Brownian motion? We have seen that the Lebesgue measure of a planar Brownian path is zero almost surely, but a more precise answer needs the concept of Hausdorff measure and dimension, which we develop in the next lecture. $\diamond$

The following corollary holds for Brownian motion in any dimension $d \geqslant 2$ and answers the question of the title in the negative.

Corollary 1.8. *For any points $x, y \in \mathbb{R}^d$, $d \geqslant 2$, we have $\mathbb{P}_x\{y \in B(0,1]\} = 0$.*

Proof. Observe that, by projection onto the first two coordinates, it suffices to prove this result for $d = 2$. Note that Theorem 1.5 holds for Brownian motion with arbitrary starting point $y \in \mathbb{R}^2$. By Fubini's theorem, for any fixed $y \in \mathbb{R}^2$,

$$\int_{\mathbb{R}^2} \mathbb{P}_y\{x \in B[0,1]\}\, dx = \mathbb{E}_y \mathcal{L}_2(B[0,1]) = 0.$$

Hence, for $\mathcal{L}_2$-almost every point x, we have $\mathbb{P}_y\{x \in B[0,1]\} = 0$. By symmetry of Brownian motion,

$$\mathbb{P}_y\{x \in B[0,1]\} = \mathbb{P}_0\{x - y \in B[0,1]\} = \mathbb{P}_0\{y - x \in B[0,1]\} = \mathbb{P}_x\{y \in B[0,1]\}.$$

We infer that $\mathbb{P}_x\{y \in B[0,1]\} = 0$, for $\mathcal{L}_2$-almost every point x. For any $\varepsilon > 0$ we thus have, almost surely, $\mathbb{P}_{B(\varepsilon)}\{y \in B[0,1]\} = 0$. Hence,

$$\mathbb{P}_x\{y \in B(0,1]\} = \lim_{\varepsilon \downarrow 0} \mathbb{P}_x\{y \in B[\varepsilon,1]\} = \lim_{\varepsilon \downarrow 0} \mathbb{E}_x \mathbb{P}_{B(\varepsilon)}\{y \in B[0,1-\varepsilon]\} = 0,$$

where we have used the Markov property in the second step. $\blacksquare$

2 How big is the path of Brownian motion?

Dimensions are a tool to measure the size of mathematical objects on a crude scale. For example, in classical geometry one can use dimension to see that a line segment (a one-dimensional object) is smaller than the surface of a ball (a two-dimensional object), but there is no difference between line-segments of different lengths. It may therefore come as a surprise that dimension is able to distinguish the size of so many objects in probability theory.

For every $\alpha \geqslant 0$ the α-**value** of a sequence $E_1, E_2, \dots$ of sets in a metric space is (with $|E_i|$ denoting the diameter of E_i)

$$\sum_{i=1}^{\infty} |E_i|^{\alpha} .$$

For every $\alpha \geqslant 0$ denote

$$\mathcal{H}_{\delta}^{\alpha}(E) := \inf \Big\{ \sum_{i=1}^{\infty} |E_i|^{\alpha} \colon E_1, E_2, \dots \text{ is a covering of } E \text{ with } |E_i| \leqslant \delta \Big\} .$$

The α-**Hausdorff measure** of E is defined as

$$\mathcal{H}^{\alpha}(E) = \lim_{\delta \downarrow 0} \mathcal{H}_{\delta}^{\alpha}(E),$$

informally speaking the α-value of the most efficient covering by small sets. If $0 \leqslant \alpha < \beta$, and $\mathcal{H}^{\alpha}(E) < \infty$, then $\mathcal{H}^{\beta}(E) = 0$. If $0 \leqslant \alpha < \beta$, and $\mathcal{H}^{\beta}(E) > 0$, then $\mathcal{H}^{\alpha}(E) = \infty$. Thus we can define

$$\dim E = \inf \Big\{ \alpha \geqslant 0 \colon \mathcal{H}^{\alpha}(E) < \infty \Big\} = \sup \Big\{ \alpha \geqslant 0 \colon \mathcal{H}^{\alpha}(E) > 0 \Big\},$$

the **Hausdorff dimension** of the set E.

Theorem 2.1. *For $d \geqslant 2$, then $\mathcal{H}^2\big(B([0,1])\big) < \infty$ and hence dim $B[0,1] \leqslant 2$, almost surely.*

Proof. For any $n \in \mathbb{N}$, we look at the covering of $B([0,1])$ by the closure of the balls

$$\mathcal{B}\Big(B(\tfrac{k}{n}), \max_{\frac{k}{n} \leqslant t \leqslant \frac{k+1}{n}} \big|B(t) - B(\tfrac{k}{n})\big|\Big), \quad k \in \{0, \dots, n-1\}.$$

By the uniform continuity of Brownian motion on the unit interval, the maximal diameter in these coverings goes to zero, as $n \to \infty$. Moreover, we have

$$\mathbb{E}\Big[\Big(\max_{\frac{k}{n} \leqslant t \leqslant \frac{k+1}{n}} \big|B(t) - B(\tfrac{k}{n})\big| \Big)^2\Big] = \mathbb{E}\Big[\Big(\max_{0 \leqslant t \leqslant \frac{1}{n}} |B(t)| \Big)^2\Big] = \frac{1}{n}\, \mathbb{E}\Big[\Big(\max_{0 \leqslant t \leqslant 1} |B(t)| \Big)^2\Big],$$

using Brownian scaling. The expectation on the right is finite by the reflection principle. Indeed,

$$\begin{aligned}\mathbb{E}\Big[\Big(\max_{0\leqslant t\leqslant 1}|B(t)|\Big)^2\Big] &= \int_0^\infty \mathbb{P}\big\{\max_{0\leqslant t\leqslant 1}|B(t)|^2 > x\big\}\,dx\\ &\leqslant 4\int_0^\infty \mathbb{P}\Big\{\max_{0\leqslant t\leqslant 1}W(t) > \sqrt{x/\sqrt{2}}\Big\}\,dx\\ &\leqslant 8\int_0^\infty \mathbb{P}\Big\{B(t) > \sqrt{x/\sqrt{2}}\Big\}\,dx < \infty.\end{aligned}$$

Hence the expected 2-value of the nth covering is bounded from above by

$$4\mathbb{E}\Big[\sum_{k=0}^{n-1}\Big(\max_{\frac{k}{n}\leqslant t\leqslant\frac{k+1}{n}}\big|B(t)-B(\tfrac{k}{n})\big|\Big)^2\Big] = 4\,\mathbb{E}\Big[\Big(\max_{0\leqslant t\leqslant 1}|B(t)|\Big)^2\Big],$$

which implies, by Fatou's lemma, that

$$\mathbb{E}\Big[\liminf_{n\to\infty} 4\sum_{k=0}^{n-1}\Big(\max_{\frac{k}{n}\leqslant t\leqslant\frac{k+1}{n}}\big|B(t)-B(\tfrac{k}{n})\big|\Big)^2\Big] < \infty.$$

Hence the liminf is almost surely finite, as required. ■

Remark 2.2. *With some extra effort it can be shown that* $\mathcal{H}^2(B([0,1]))=0$.

From the definition of the Hausdorff dimension it is plausible that in many cases it is relatively easy to give an upper bound on the dimension: just find an efficient cover of the set and find an upper bound to its α-value. However it looks more difficult to give lower bounds, as we must obtain a lower bound on α-values of *all* covers of the set. The energy method is a way around this problem, which is based on the existence of a nonzero measure on the set. The basic idea is that, if this measure distributes a positive amount of mass on a set E in such a manner that its local concentration is bounded from above, then the set must be large in a suitable sense. Suppose μ is a measure on a metric space (E,ρ) and $\alpha\geqslant 0$. The **α-potential** of a point $x\in E$ with respect to μ is defined as

$$\phi_\alpha(x) = \int \frac{d\mu(y)}{\rho(x,y)^\alpha}\,.$$

In the case $E=\mathbb{R}^3$ and $\alpha=1$, this is the Newton gravitational potential of the mass μ. The **α-energy** of μ is

$$I_\alpha(\mu) = \int \phi_\alpha(x)\,d\mu(x) = \iint \frac{d\mu(x)\,d\mu(y)}{\rho(x,y)^\alpha}\,.$$

Measures with $I_\alpha(\mu)<\infty$ spread the mass so that at each place the concentration is sufficiently small to overcome the singularity of the integrand. This is only possible on sets which are large in a suitable sense.

Theorem 2.3 (Energy method)**.** *Let $\alpha \geqslant 0$ and μ be a nonzero measure on a metric space E. Then, for every $\varepsilon > 0$, we have*

$$\mathcal{H}_\varepsilon^\alpha(E) \geqslant \frac{\mu(E)^2}{\iint_{\rho(x,y)<\varepsilon} \frac{d\mu(x)\, d\mu(y)}{\rho(x,y)^\alpha}}\,.$$

Hence, if $I_\alpha(\mu) < \infty$ then $\mathcal{H}^\alpha(E) = \infty$ and, in particular, $\dim E \geqslant \alpha$.

Proof. (Due to O. Schramm) If $\{A_n : n = 1, 2, \ldots\}$ is any pairwise disjoint covering of E consisting of closed sets of diameter $< \varepsilon$, then

$$\iint_{\rho(x,y)<\varepsilon} \frac{d\mu(x)\, d\mu(y)}{\rho(x,y)^\alpha} \geqslant \sum_{n=1}^\infty \iint_{A_n \times A_n} \frac{d\mu(x)\, d\mu(y)}{\rho(x,y)^\alpha} \geqslant \sum_{n=1}^\infty \frac{\mu(A_n)^2}{|A_n|^\alpha},$$

and moreover,

$$\mu(E) \leqslant \sum_{n=1}^\infty \mu(A_n) = \sum_{n=1}^\infty |A_n|^{\frac{\alpha}{2}} \frac{\mu(A_n)}{|A_n|^{\frac{\alpha}{2}}}$$

Given $\delta > 0$ choose a covering as above such that additionally

$$\sum_{n=1}^\infty |A_n|^\alpha \leqslant \mathcal{H}_\varepsilon^\alpha(E) + \delta.$$

Using now the Cauchy–Schwarz inequality, we get

$$\mu(E)^2 \leqslant \sum_{n=1}^\infty |A_n|^\alpha \sum_{n=1}^\infty \frac{\mu(A_n)^2}{|A_n|^\alpha} \leqslant \Big(\mathcal{H}_\varepsilon^\alpha(E) + \delta\Big) \iint_{\rho(x,y)<\varepsilon} \frac{d\mu(x)\, d\mu(y)}{\rho(x,y)^\alpha}.$$

Letting $\delta \downarrow 0$ gives the stated inequality. Further, letting $\varepsilon \downarrow 0$, if $\mathbb{E} I_\alpha(\mu) < \infty$ the integral converges to zero, so that $\mathcal{H}_\varepsilon^\alpha(E)$ diverges to infinity. ■

Remark 2.4. To get a lower bound on the dimension from this method it suffices to show finiteness of a single integral. In particular, in order to show for a random set E that $\dim E \geqslant \alpha$ almost surely, it suffices to show that $\mathbb{E} I_\alpha(\mu) < \infty$ for a (random) measure on E. ◇

Theorem 2.5 (Taylor 1953)**.** *Let $\{B(t) : 0 \leqslant t \leqslant 1\}$ be d-dimensional Brownian motion, $d \geqslant 2$, then* $\dim B[0, 1] = 2$ *almost surely.*

Proof. Recall that we already know the upper bound. We now look at the lower bound. A natural measure on $B[0, 1]$ is the occupation measure μ defined by $\mu(A) = \mathcal{L}(B^{-1}(A) \cap [0, 1])$, for all Borel sets $A \subset \mathbb{R}^d$, or, equivalently,

$$\int_{\mathbb{R}^d} f(x)\, d\mu(x) = \int_0^1 f\big(B(t)\big)\, dt,$$

for all bounded measurable functions f. We want to show that, for $0 < \alpha < 2$,

$$\mathbb{E} \iint \frac{d\mu(x)\, d\mu(y)}{|x-y|^\alpha} = \mathbb{E} \int_0^1 \int_0^1 \frac{ds\, dt}{|B(t) - B(s)|^\alpha} < \infty. \tag{2}$$

Let us evaluate the expectation

$$\mathbb{E}|B(t) - B(s)|^{-\alpha} = \mathbb{E}\big[(|t-s|^{1/2}|B(1)|)^{-\alpha}\big] = |t-s|^{-\alpha/2} \int_{\mathbb{R}^d} \frac{c_d}{|z|^\alpha} e^{-|z|^2/2} dz.$$

The integral can be evaluated using polar coordinates, but all we need is that it is a finite constant c depending on d and α only. Substituting this expression into (2) and using Fubini's theorem we get

$$\mathbb{E} I_\alpha(\mu) = c \int_0^1 \int_0^1 \frac{ds\, dt}{|t-s|^{\alpha/2}} \leqslant 2c \int_0^1 \frac{du}{u^{\alpha/2}} < \infty. \tag{3}$$

Therefore $I_\alpha(\mu) < \infty$ and hence $\dim B[0,1] > \alpha$, almost surely. The lower bound on the range follows by letting $\alpha \uparrow 2$. ■

We define the α-**capacity**, of a metric space (E, ρ) as

$$\mathrm{Cap}_\alpha(E) := \sup \Big\{ I_\alpha(\mu)^{-1} \colon \mu \text{ a probability measure on } E \Big\}.$$

In the case of the Euclidean space $E = \mathbb{R}^d$ with $d \geqslant 3$ and $\alpha = d - 2$ the α-capacity is also known as the **Newtonian capacity**. Theorem 2.3 states that a set of positive α-capacity has dimension at least α. The famous Frostman's lemma states that in Euclidean spaces this method is sharp, i.e., for any closed (or, more generally, analytic) set $A \subset \mathbb{R}^d$,

$$\dim A = \sup \big\{ \alpha : \ \mathrm{Cap}_\alpha(A) > 0 \big\}.$$

We omit the proof, but we'll discuss in Exercise 4 how this result can be exploited to give McKean's theorem.

3 Which sets are hit by Brownian motion?

One of our ideas to measure the size of sets was based on the notion of capacity. While this notion appeared to be useful, but maybe a bit artificial at the time, we can now understand its true meaning. This is linked to the notion of polarity, namely whether a set has a positive probability of being hit by a suitably defined random set.

More precisely, we call a Borel set $A \subset \mathbb{R}^d$ is *polar* for Brownian motion if, for all x,

$$\mathbb{P}_x\big\{B(t) \in A \text{ for some } t > 0\big\} = 0.$$

By Corollary 1.8 points are polar for Brownian motion in all dimensions $d \geqslant 2$. The general characterisation of polar sets requires an extension of the notion of capacities to a bigger class of kernels.

Suppose $A \subset \mathbb{R}^d$ is a Borel set and $K \colon \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty]$ is a kernel. Then the K-energy of a measure μ is defined to be

$$I_K(\mu) = \iint K(x, y)\, d\mu(x)\, d\mu(y),$$

and the K-capacity of A is defined as

$$\mathrm{Cap}_K(A) = \big[\inf\big\{I_K(\mu) \colon \mu \text{ a probability measure on } A\big\}\big]^{-1}.$$

Recall that the α-energy of a measure and the α-capacity Cap_α of a set correspond to the kernel $K(x, y) = |x - y|^{-\alpha}$.

Theorem 3.1 (Kakutani's theorem)**.** *A closed set Λ is polar for d-dimensional Brownian motion if and only if it has zero K-capacity for the* **potential kernel** *K defined by*

$$K(x, y) = \begin{cases} \big|\log\big(\frac{1}{|x-y|}\big)\big| & \text{if } d = 2, \\ |x - y|^{2-d} & \text{if } d \geqslant 3. \end{cases}$$

Instead of proving Kakutani's theorem directly, we aim for a stronger, quantitative result. To this end we take $\{B(t) \colon 0 \leqslant t \leqslant T\}$ to be d-dimensional Brownian motion killed at time T, and either $d \geqslant 3$ and $T = \infty$, or $d \geqslant 2$ and T is the first exit time from a bounded domain D containing the origin. This result gives, for compact sets $\Lambda \subset D$, a quantitative estimate of

$$\mathbb{P}_0\big\{\exists 0 < t < T \text{ such that } B(t) \in \Lambda\big\}$$

in terms of capacities. However, even if $d = 3$ and $T = \infty$, one cannot expect that

$$\mathbb{P}_0\big\{\exists t > 0 \text{ such that } B(t) \in \Lambda\big\} \asymp \mathrm{Cap}_K(\Lambda)$$

for the potential kernel K in Theorem 3.1. Observe, for example, that the left hand side depends strongly on the starting point of Brownian motion, whereas the right hand side is translation invariant. Similarly, if Brownian motion is starting at the origin, the left hand side is invariant under scaling, i.e. remains the same when Λ is replaced by $\lambda\Lambda$ for any $\lambda > 0$, whereas the right hand side is not. For a direct comparison of hitting probabilities and capacities, it is therefore necessary to use a capacity function with respect to a scale-invariant modified kernel. To this end define the Green kernel $G(x, \cdot)$, as the density of the expected occupation measure

$$\int G(x,y) f(y)\, dy = \mathbb{E}_x \int_0^T f(B(t))\, dt,$$

which in the case $d \geqslant 3$ is known as

$$G(x,y) = \frac{\Gamma(d/2-1)}{2\pi^{d/2}} \, |x-y|^{2-d},$$

and hence agrees up to a constant multiple with the potential kernel. We define the **Martin kernel** $M \colon D \times D \to [0,\infty]$ by

$$M(x,y) := \frac{G(x,y)}{G(0,y)} \qquad \text{for } x \neq y,$$

and otherwise by $M(x,x) = \infty$.

The following theorem of Benjamini, Pemantle and Peres (1995) shows that Martin capacity is indeed a good estimate of the hitting probability.

Theorem 3.2. *Let $\{B(t) \colon 0 \leqslant t \leqslant T\}$ be a Brownian motion killed as before, and $A \subset D$ closed. Then*

$$\tfrac{1}{2} \operatorname{Cap}_M(A) \leqslant \mathbb{P}_0\{\exists 0 < t \leqslant T \text{ such that } B(t) \in A\} \leqslant \operatorname{Cap}_M(A) \tag{4}$$

Proof. Let μ be the (possibly sub-probability) distribution of $B(\tau)$ for the stopping time $\tau = \inf\{0 < t \leqslant T \,:\, B(t) \in A\}$. Note that the total mass of μ is

$$\mu(A) = \mathbb{P}_0\{\tau \leqslant T\} = \mathbb{P}_0\{B(t) \in A \text{ for some } 0 < t \leqslant T\}. \tag{5}$$

The idea for the upper bound is that if the harmonic measure μ is nonzero, it is an obvious candidate for a measure of finite M-energy. Recall from the definition of the Green's function, for any $y \in D$,

$$\mathbb{E}_0 \int_0^T \mathbf{1}\{|B(t)-y| < \varepsilon\}\, dt = \int_{\mathcal{B}(y,\varepsilon)} G(0,z)\, dz. \tag{6}$$

By the strong Markov property applied to the first hitting time τ of A,

$$\begin{aligned}
\mathbb{P}_0\big\{|B(t)-y| < \varepsilon \text{ and } t \leqslant T\big\} &\geqslant \mathbb{P}_0\big\{|B(t)-y| < \varepsilon \text{ and } \tau \leqslant t \leqslant T\big\} \\
&= \mathbb{E}\mathbb{P}\big\{|B(t-\tau)-y| < \varepsilon \mid \mathcal{F}(\tau)\big\}.
\end{aligned}$$

Integrating over t and using Fubini's theorem yields

$$\mathbb{E}_0 \int_0^T 1\{|B(t)-y|<\varepsilon\}\,dt \geqslant \int_A \int_{\mathcal{B}(y,\varepsilon)} G(x,z)\,dz\,d\mu(x).$$

Combining this with (6) we infer that

$$\int_{\mathcal{B}(y,\varepsilon)} \int_A G(x,z)\,d\mu(x)\,dz \leqslant \int_{\mathcal{B}(y,\varepsilon)} G(0,z)\,dz\,.$$

Dividing by $\mathcal{L}(\mathcal{B}(0,\varepsilon))$ and letting $\varepsilon \downarrow 0$ we obtain

$$\int_A G(x,y)\,d\mu(x) \leqslant G(0,y),$$

i.e. $\int_A M(x,y)\,d\mu(x) \leqslant 1$ for all $y \in D$. Therefore, $I_M(\mu) \leqslant \mu(A)$ and thus if we use $\mu/\mu(A)$ as a probability measure we get

$$\mathrm{Cap}_M(A) \geqslant [I_M(\mu/\mu(A))]^{-1} \geqslant \mu(A),$$

which by (5) yields the upper bound on the probability of hitting A.

To obtain a lower bound for this probability, a second moment estimate is used. It is easily seen that the Martin capacity of A is the supremum of the capacities of its compact subsets, so we may assume that A is a compact subset of the domain $D \setminus \{0\}$. We take $\varepsilon > 0$ smaller than half the distance of A to $D^{\mathrm{c}} \cup \{0\}$. For $x, y \in A$ let

$$h_\varepsilon(x,y) = \int_{\mathcal{B}(y,\varepsilon)} G(x,\xi)\,d\xi$$

denote the expected time which a Brownian motion started in x spends in the ball $\mathcal{B}(y,\varepsilon)$. Also define

$$h^*_\varepsilon(x,y) = \sup_{|x-z|<\varepsilon} \int_{\mathcal{B}(y,\varepsilon)} G(z,\xi)\,d\xi.$$

Given a probability measure ν on A, and $\varepsilon > 0$, consider the random variable

$$Z_\varepsilon = \int_A \int_0^T \frac{1\{B(t) \in \mathcal{B}(y,\varepsilon)\}}{h_\varepsilon(0,y)}\,dt\,d\nu(y)\,.$$

Clearly $\mathbb{E}_0 Z_\varepsilon = 1$. By symmetry, the second moment of Z_ε can be written as

$$\begin{aligned}
\mathbb{E}_0 Z_\varepsilon^2 &= 2\mathbb{E}_0 \int_0^T ds \int_s^T dt \iint \frac{1\{B(s) \in \mathcal{B}(x,\varepsilon),\, B(t) \in \mathcal{B}(y,\varepsilon)\}}{h_\varepsilon(0,x)h_\varepsilon(0,y)}\,d\nu(x)\,d\nu(y) \\
&\leqslant 2\mathbb{E}_0 \iint \int_0^T ds\, 1\{B(s) \in \mathcal{B}(x,\varepsilon)\} \frac{h^*_\varepsilon(x,y)}{h_\varepsilon(0,x)h_\varepsilon(0,y)}\,d\nu(x)\,d\nu(y) \\
&= 2 \iint \frac{h^*_\varepsilon(x,y)}{h_\varepsilon(0,y)}\,d\nu(x)\,d\nu(y).
\end{aligned} \tag{7}$$

Observe that, for all fixed $x, y \in A$ we have $\lim_{\varepsilon\downarrow 0} \mathcal{L}(\mathcal{B}(0,\varepsilon))^{-1}\, h^*_\varepsilon(x,y) = G(x,y)$ and $\lim_{\varepsilon\downarrow 0} \mathcal{L}(\mathcal{B}(0,\varepsilon))^{-1}\, h_\varepsilon(0,y) = G(0,y)$. Hence, we obtain,

$$\lim_{\varepsilon\downarrow 0} \mathbb{E}Z_\varepsilon^2 \leqslant 2 \iint \frac{G(x,y)}{G(0,y)}\, d\nu(x)\, d\nu(y) = 2 I_M(\nu), \tag{8}$$

using dominated convergence (which has to be justified by a separate argument, which we omit here). Clearly, the hitting probability $\mathbb{P}\{\exists 0 < t < T, y \in A$ such that $B(t) \in \mathcal{B}(y,\varepsilon)\}$ is at least

$$\mathbb{P}\{Z_\varepsilon > 0\} \geqslant \frac{(\mathbb{E}Z_\varepsilon)^2}{\mathbb{E}Z_\varepsilon^2} = (\mathbb{E}Z_\varepsilon^2)^{-1}\,,$$

where we have used the Paley–Zygmund inequality. Compactness of A, together with transience and continuity of Brownian motion, imply that if the Brownian path visits every ε-neighbourhood of the compact set A then it intersects A itself. Therefore, by (8),

$$\mathbb{P}\{\exists 0 < t < T \text{ such that } B(t) \in A\} \geqslant \lim_{\varepsilon\downarrow 0} (\mathbb{E}Z_\varepsilon^2)^{-1} \geqslant \frac{1}{2 I_M(\nu)}.$$

Since this is true for all probability measures ν on A, we get the conclusion. ■

From Theorem 3.2 we now readily obtain Kakutani's theorem. It suffices, by taking countable unions, to consider compact sets Λ which have positive distance from the origin. First consider the case $d \geqslant 3$. Then $G(0,x)$ is bounded away from zero and infinity. Hence the set Λ is polar if and only if its K-capacity vanishes. In the case $d = 2$ we choose a large ball $\mathcal{B}(0,R)$ containing Λ. The Green's function for the Brownian motion stopped upon leaving $\mathcal{B}(0,R)$ satisfies

$$G(x,y) = -\tfrac{1}{\pi} \log|x-y| + \mathbb{E}_x\big[\tfrac{1}{\pi} \log|B(T) - y|\big]\,.$$

The second summand of $G(x,y)$ is bounded from above if $x, y \in \Lambda$, and $G(0,y)$ is bounded from zero. Hence only the contribution from $-\log|x-y|$ decides about finiteness of the Martin energy of a probability measure. Therefore, any probability measure on Λ with finite Martin energy has finite K-energy, and vice versa. This completes the proof.

Suppose that $\{B_1(t) \colon t \geqslant 0\}$ and $\{B_2(t) \colon t \geqslant 0\}$ are two independent d-dimensional Brownian motions started in arbitrary points. The question we ask now is, in which dimensions the paths of the two motions have a nontrivial intersection, in other words whether there exist times $t_1, t_2 > 0$ such that $B_1(t_1) = B_2(t_2)$. Keeping the path $\{B_1(t) \colon t \geqslant 0\}$ fixed, we have to decide whether it is a polar set for the second Brownian motion. By Kakutani's theorem, Theorem 3.1, this question depends on its capacity with respect to the potential kernel.

Theorem 3.3.

(a) *For $d \geqslant 4$, almost surely, two independent Brownian paths in $\mathbb{R}^d$ have an empty intersection, except for a possible common starting point.*

(b) *For $d \leqslant 3$, almost surely, the intersection of two independent Brownian paths in $\mathbb{R}^d$ is nontrivial, i.e. contains points other than a possible common starting point.*

Proof. **(a).** Note that it suffices to look at one Brownian motion and show that its path is, almost surely, a set of capacity zero with respect to the potential kernel. If $d \geqslant 4$, the capacity with respect to the potential kernel is a multiple of the $(d-2)$-capacity. By Theorem 2.3 this capacity is zero for sets of finite $(d-2)$-dimensional Hausdorff measure. Now note that if $d \geqslant 5$ the dimension of a Brownian path is two, and hence strictly smaller than $d-2$, so that the $(d-2)$-dimensional Hausdorff measure is zero, which shows that the capacity must be zero. If $d = 4$ the situation is only marginally more complicated, although the dimension of the Brownian path is $2 = d-2$ and the simple argument above does not apply. However, we know that $\mathcal{H}^2(B[0,1]) < \infty$ almost surely, which implies that $\mathrm{Cap}_2(B[0,1]) = 0$ by Theorem 2.3. This implies that an independent Brownian motion almost surely does not hit any of the segments $B[n, n+1]$, and therefore avoids the path entirely.
(b). If $d = 3$, the capacity with respect to the potential kernel is a multiple of the 1-capacity. As the Hausdorff dimension of a path is two, this capacity is positive by Frostman's lemma. Therefore two Brownian paths in $d = 3$ intersect with positive probability. It is an exercise to complete the proof and show that this happens almost surely. ■

4 Does Brownian motion have multiple points?

Having found that two Brownian motions intersect, we now ask whether for $d = 2, 3$ a collection of p independent d-dimensional Brownian motions

$$\{B_1(t)\colon t \geqslant 0\}, \ldots, \{B_p(t)\colon t \geqslant 0\}$$

intersect, i.e. if there exist times $t_1, \ldots, t_p > 0$ such that $B_1(t_1) = \cdots = B_p(t_p)$.

Theorem 4.1.

(a) *For $d \geqslant 3$, almost surely, three independent Brownian paths in $\mathbb{R}^d$ have an empty intersection, except for a possible common starting point.*

(b) *For $d = 2$, almost surely, the intersection of any finite number p of independent Brownian paths in $\mathbb{R}^d$ is nontrivial, i.e. contains points other than a possible common starting point.*

In the light of our discussion of the case $p = 2$, it is natural to approach the question about the existence of intersections of p paths, by asking for the Hausdorff dimension and measure of the intersection of $p-1$ paths. This leads to an easy proof of (a).

Lemma 4.2. *Suppose $\{B_i(t)\colon t \geqslant 0\}$, for $i = 1, 2$, are two independent Brownian motions in $d = 3$. Then, almost surely, for every compact set $\Lambda \subset \mathbb{R}^3$ not containing the starting points of the Brownian motions, we have $\mathcal{H}^1(B_1[0,\infty) \cap B_2[0,\infty) \cap \Lambda) < \infty$.*

Proof. Fix a cube $\mathsf{Cube} \subset \mathbb{R}^3$ of unit side length not containing the starting points. It suffices to show that, almost surely, $\mathcal{H}^1(B_1[0,\infty) \cap B_2[0,\infty) \cap \mathsf{Cube}) < \infty$. For this purpose let $\mathfrak{C}_n$ be the collection of dyadic subcubes of Cube of side length 2^{-n}, and $\mathfrak{I}_n$ be the collection of cubes in $\mathfrak{C}_n$ which are hit by both motions. There exists $C > 0$ such that, for any cube $E \in \mathfrak{C}_n$,

$$\mathbb{P}\big\{E \in \mathfrak{I}_n\big\} = \mathbb{P}\big\{\exists s > 0 \text{ with } B(s) \in E\big\}^2 \leqslant C 2^{-2n}.$$

Now, for every n, the collection $\mathfrak{I}_n$ is a covering of $B_1[0,\infty) \cap B_2[0,\infty) \cap \mathsf{Cube}$, and

$$\mathbb{E}\Big[\sum_{E \in \mathfrak{I}_n} |E|\Big] = 2^{3n}\, \mathbb{P}\big\{E \in \mathfrak{I}_n\big\} \sqrt{3} 2^{-n} \leqslant C\sqrt{3}.$$

Therefore, by Fatou's lemma, we obtain

$$\mathbb{E}\Big[\liminf_{n\to\infty} \sum_{E \in \mathfrak{I}_n} |E|\Big] \leqslant \liminf_{n\to\infty} \mathbb{E}\Big[\sum_{E \in \mathfrak{I}_n} |E|\Big] \leqslant C\sqrt{3}.$$

Hence the liminf is finite almost surely, and we infer from this that $\mathcal{H}^1(B_1[0,\infty) \cap B_2[0,\infty) \cap \mathsf{Cube})$ is finite almost surely. ∎

To prove Theorem 4.1 (a) it suffices to show that, for any cube Cube of unit side length which does not contain the origin, we have $\mathrm{Cap}_1(B_1[0,\infty) \cap B_2[0,\infty) \cap \mathsf{Cube}) = 0$. This follows directly from Lemma 4.2 and the energy method, Theorem 2.3. For Theorem 4.1 (b) it suffices to show that the Hausdorff dimension of the set $B_1(0,\infty) \cap \ldots \cap B_{p-1}(0,\infty)$ is positive in the case $d = 2$. This problem was raised by Itô and McKean in the first edition of their influential book and has since been resolved by Taylor and Fristedt. The problem of finding lower bounds for the Hausdorff dimension of the intersection sets is best approached using the technique of *stochastic co-dimension*, which we discuss now.

Given a set A, the idea behind the stochastic co-dimension approach is to take a suitable random test set Θ, and check whether $\mathbb{P}\{\Theta \cap A \neq \emptyset\}$ is zero or positive. In the latter case this indicates that the set is large, and we should therefore get a lower bound on the dimension of A. The approach we use here is based on using the family of percolation limit sets as test sets.

Suppose that $C \subset \mathbb{R}^d$ is a fixed compact unit cube. We denote by $\mathfrak{C}_n$ the collection of compact dyadic subcubes (relative to C) of side length 2^{-n}. We also let

$$\mathfrak{C} = \bigcup_{n=0}^{\infty} \mathfrak{C}_n.$$

Given $\gamma \in [0,d]$ we construct a random compact set $\Gamma[\gamma] \subset C$ inductively as follows: We keep each of the 2^d compact cubes in $\mathfrak{C}_1$ independently with probability $p = 2^{-\gamma}$. Let $\mathfrak{S}_1$ be the collection of cubes kept in this procedure and $\mathsf{S}(1)$ their union. Pass from $\mathfrak{S}_n$ to $\mathfrak{S}_{n+1}$ by keeping each cube of $\mathfrak{C}_{n+1}$, which is not contained in a previously rejected cube, independently with probability p. Denote by $\mathfrak{S} = \bigcup_{n=1}^{\infty} \mathfrak{S}_n$ and let $\mathsf{S}(n+1)$ be the union of the cubes in $\mathfrak{S}_{n+1}$. Then the random set

$$\Gamma[\gamma] := \bigcap_{n=1}^{\infty} \mathsf{S}(n)$$

is called a **percolation limit set**. The usefulness of percolation limit sets in fractal geometry comes from the following theorem.

Theorem 4.3 (Hawkes 1981). *For every $\gamma \in [0,d]$ and every closed set $A \subset C$ the following properties hold*

- **(i)** *if* $\dim A < \gamma$, *then almost surely,* $A \cap \Gamma[\gamma] = \emptyset$,
- **(ii)** *if* $\dim A > \gamma$, *then* $A \cap \Gamma[\gamma] \neq \emptyset$ *with positive probability.*

Remark 4.4. *The stochastic co-dimension technique and the energy method are closely related: A set A is called* polar for the percolation limit set*, if*

$$\mathbb{P}\{A \cap \Gamma[\gamma] \neq \emptyset\} = 0.$$

A set is polar for the percolation limit set if and only if it has γ-capacity zero. For $d \geqslant 3$, the criterion for polarity of a percolation limit set with $\gamma = d-2$

therefore agrees with the criterion for the polarity for Brownian motion, recall Theorem 3.1. This 'equivalence' between percolation limit sets and Brownian motion has a quantitative strengthening due to Peres.

The proof of part (i) in Hawkes' theorem is based on the *first moment method*, which means that we essentially only have to calculate an expectation. Because $\dim A < \gamma$ there exists, for every $\varepsilon > 0$, a covering of A by countably many sets $D_1, D_2, \dots$ with $\sum_{i=1}^{\infty} |D_i|^\gamma < \varepsilon$. As each set is contained in no more than a constant number of dyadic cubes of smaller diameter, we may even assume that $D_1, D_2, \dots \in \mathfrak{C}$. Suppose that the side length of D_i is 2^{-n}, then the probability that $D_i \in \mathfrak{S}_n$ is $2^{-n\gamma}$. By picking from $D_1, D_2, \dots$ those cubes which are in $\mathfrak{S}$ we get a covering of $A \cap \Gamma[\gamma]$. Let N be the number of cubes picked in this procedure, then

$$\mathbb{P}\{A \cap \Gamma[\gamma] \neq \emptyset\} \leqslant \mathbb{P}\{N > 0\} \leqslant \mathbb{E}N = \sum_{i=1}^{\infty} \mathbb{P}\{D_i \in \mathfrak{S}\} = \sum_{i=1}^{\infty} |D_i|^\gamma < \varepsilon.$$

As this holds for all $\varepsilon > 0$ we infer that, almost surely, we have $A \cap \Gamma[\gamma] = \emptyset$.

The proof of part (ii) is based on the *second moment method*, which means that a variance has to be calculated. We also use Frostman's lemma, which states that, as $\dim A > \gamma$, there exists a probability measure μ on A such that $I_\gamma(\mu) < \infty$. Now let n be a positive integer and define the random variables

$$Y_n = \sum_{C \in \mathfrak{S}_n} \frac{\mu(C)}{|C|^\gamma} = \sum_{C \in \mathfrak{C}_n} \mu(C) 2^{n\gamma}\, 1_{\{C \in \mathfrak{S}_n\}}.$$

Note that $Y_n > 0$ implies $\mathsf{S}(n) \cap A \neq \emptyset$ and, by compactness, if $Y_n > 0$ for all n we even have $A \cap \Gamma[\gamma] \neq \emptyset$. As $Y_{n+1} > 0$ implies $Y_n > 0$, we get that

$$\mathbb{P}\big\{A \cap \Gamma[\gamma] \neq \emptyset\big\} \geqslant \mathbb{P}\big\{Y_n > 0 \text{ for all } n\big\} = \lim_{n \to \infty} \mathbb{P}\big\{Y_n > 0\big\}.$$

It therefore suffices to give a positive lower bound for $\mathbb{P}\{Y_n > 0\}$ independent of n. A straightforward calculation gives for the first moment $\mathbb{E}[Y_n] = \sum_{C \in \mathfrak{C}_n} \mu(C) = 1$. For the second moment we find

$$\mathbb{E}[Y_n^2] = \sum_{C \in \mathfrak{C}_n} \sum_{D \in \mathfrak{C}_n} \mu(C)\mu(D)\, 2^{2n\gamma}\, \mathbb{P}\{C \in \mathfrak{S}_n \text{ and } D \in \mathfrak{S}_n\}.$$

The latter probability depends on the dyadic distance of the cubes C and D: if 2^{-m} is the side length of the smallest dyadic cube which contains both C and D, then the probability in question is $2^{-2\gamma(n-m)} 2^{-\gamma m}$. The value m can be estimated in terms of the Euclidean distance of the cubes, indeed if $x \in C$ and $y \in D$ then $|x - y| \leqslant \sqrt{d} 2^{-m}$. This gives a handle to estimate the second moment in terms of the energy of μ. We find that

$$\mathbb{E}[Y_n^2] = \sum_{C \in \mathfrak{C}_n} \sum_{D \in \mathfrak{C}_n} \mu(C)\mu(D) 2^{\gamma m} \leqslant d^{\gamma/2} \iint \frac{d\mu(x)\, d\mu(y)}{|x - y|^\gamma} = d^{\gamma/2} I_\gamma(\mu).$$

Plugging these moment estimates into the Paley–Zygmund inequality gives $\mathbb{P}\{Y_n > 0\} \geqslant d^{-\gamma/2} I_\gamma(\mu)^{-1}$, as required.

We now apply Hawkes' theorem to intersections of Brownian paths.

Theorem 4.5. *Suppose $d = 2, 3$ and $\{B_1(t)\colon t \geqslant 0\}$, $\{B_2(t)\colon t \geqslant 0\}$ are independent d-dimensional Brownian motions. Then, almost surely,*

$$\dim\big(B_1[0,\infty) \cap B_2[0,\infty)\big) = 4 - d.$$

Proof. We focus on $d = 3$ (the case $d = 2$ being similar) and note that the upper bound follows from Lemma 4.2, and hence only the lower bound remains to be proved. Suppose $\gamma < 1$ is arbitrary, and pick $\beta > 1$ such that $\gamma + \beta < 2$. Let $\Gamma[\gamma]$ and $\Gamma[\beta]$ be two independent percolation limit sets, independent of the Brownian motions. Note that $\Gamma[\gamma] \cap \Gamma[\beta]$ is a percolation limit set with parameter $\gamma+\beta$. Hence, by Theorem 4.3 (ii) and the fact that $\dim(B_1[0,\infty)) = 2 > \gamma + \beta$, we have

$$\mathbb{P}\big\{B_1[0,\infty) \cap \Gamma[\gamma] \cap \Gamma[\beta] \neq \emptyset\big\} > 0.$$

Interpreting $\Gamma[\beta]$ as the test set and using Theorem 4.3 (i) we obtain

$$\dim\big(B_1[0,\infty) \cap \Gamma[\gamma]\big) \geqslant \beta \qquad \text{with positive probability.}$$

As $\beta > 1$, given this event, the set $B_1[0,\infty) \cap \Gamma[\gamma]$ has positive capacity with respect to the potential kernel in $\mathbb{R}^3$ and is therefore nonpolar with respect to the independent Brownian motion $\{B_2(t)\colon t \geqslant 0\}$. We therefore have

$$\mathbb{P}\big\{B_1[0,\infty) \cap B_2[0,\infty) \cap \Gamma[\gamma] \neq \emptyset\big\} > 0.$$

Using Theorem 4.3 (i) we infer that $\dim(B_1[0,\infty) \cap B_2[0,\infty)) \geqslant \gamma$ with positive probability. A zero-one law shows that this must in fact hold almost surely, and the result follows as $\gamma < 1$ was arbitrary. ■

A point $x \in \mathbb{R}^d$ has **multiplicity** p for a Brownian motion $\{B(t)\colon t \geqslant 0\}$ in $\mathbb{R}^d$, if there exist times $0 < t_1 < \cdots < t_p$ with $x = B(t_1) = \cdots = B(t_p)$.

Theorem 4.6. *Suppose $d \geqslant 2$ and $\{B(t)\colon t \in [0,1]\}$ is a d-dimensional Brownian motion. Then, almost surely,*

- *if $d \geqslant 4$ no double points exist, i.e. Brownian motion is injective,*
- *if $d = 3$ double points exist, but triple points fail to exist,*
- *if $d = 2$ points of any finite multiplicity exist.*

Proof. We only discuss the problem of double points. To show *nonexistence* of double points in $d \geqslant 4$ it suffices to show that for any rational $\alpha \in (0,1)$, almost surely, there exists no times $0 \leqslant t_1 < \alpha < t_2 \leqslant 1$ with $B(t_1) = B(t_2)$. Fixing such an α, the Brownian motions $\{B_1(t)\colon 0 \leqslant t \leqslant 1-\alpha\}$

and $\{B_2(t)\colon 0 \leqslant t \leqslant \alpha\}$ given by $B_1(t) = B(\alpha + t) - B(\alpha)$ and $B_2(t) = B(\alpha - t) - B(\alpha)$ are independent and hence, by Theorem 3.3, they do not intersect, almost surely, proving the statement.

To show *existence* of double points in $d \leqslant 3$ we apply Theorem 3.3 to the independent Brownian motions $\{B_1(t)\colon 0 \leqslant t \leqslant \frac{1}{2}\}$ and $\{B_2(t)\colon\ 0 \leqslant t \leqslant \frac{1}{2}\}$ given by $B_1(t) = B(\frac{1}{2} + t) - B(\frac{1}{2})$ and $B_2(t) = B(\frac{1}{2} - t) - B(\frac{1}{2})$, to see that, almost surely, the two paths intersect. ∎

Knowing that planar Brownian motion has points of arbitrarily large *finite* multiplicity, it is an interesting question whether there are points of *infinite* multiplicity. The following deep result was first proved by Dvoretzky, Erdös and Kakutani.

Theorem 4.7. *Let $\{B(t)\colon t \geqslant 0\}$ be a planar Brownian motion. Then, almost surely, there exists a point $x \in \mathbb{R}^2$ such that the set $\{t \geqslant 0\colon B(t) = x\}$ is uncountable.*

5 How many times can Brownian motion visit a point?

How big can the sets $T(x) = \big\{t \geqslant 0 \colon B(t) = x\big\}$ of times mapped by d-dimensional Brownian motion onto the same point x possibly be? We have seen in the previous section that, almost surely,

- in dimension $d \geqslant 4$ all sets $T(x)$ consist of at most one point,
- in dimension $d = 3$ all sets $T(x)$ consist of at most two points,
- in dimension $d = 2$ at least one of the sets $T(x)$ is uncountable.

We now give an upper bound on the size of $T(x)$, simultaneously for all x, in the planar case.

Theorem 5.1. *Suppose $\{B(t) \colon t \geqslant 0\}$ is a planar Brownian motion. Then, almost surely, for all $x \in \mathbb{R}^2$, we have $\dim T(x) = 0$.*

Proof. Define $\tau_R^* = \min\big\{t : |B(t)| = R\big\}$. It is sufficient to show that, almost surely,

$$\dim(T(x) \cap [0, \tau_R^*]) = 0 \text{ for all } x \in \mathbb{R}^2, |x| < R.$$

Lemma 5.2. *Consider a cube $Q \subset \mathbb{R}^2$ centred at a point x and having diameter $2r$, and assume that the cube Q is inside the ball of radius R about the origin. Define recursively*

$$\begin{aligned} \tau_1^Q &= \inf\{t \geqslant 0 : B(t) \in Q\}\,, \\ \tau_{k+1}^Q &= \inf\{t \geqslant \tau_k^Q + r^2 : B(t) \in Q\}, \qquad \text{for } k \geqslant 1, \end{aligned}$$

There exists $c = c(R) > 0$ such that, with $2^{-m-1} < r \leqslant 2^{-m}$, for any $z \in \mathbb{R}^2$,

$$\mathbb{P}_z\big\{\tau_k^Q < \tau_R^*\big\} \leqslant \Big(1 - \frac{c}{m}\Big)^k \leqslant e^{-ck/m}. \tag{9}$$

Proof. It suffices to bound $\mathbb{P}_z\{\tau_{k+1}^Q \geqslant \tau_R^* \,|\, \tau_k^Q < \tau_R^*\}$ from below by

$$\begin{aligned} &\mathbb{P}_z\big\{\tau_{k+1}^Q \geqslant \tau_R^* \,\big|\, |B(\tau_k^Q + r^2) - x| > 2r, \tau_k^Q < \tau_R^*\big\} \\ &\qquad\qquad \times \mathbb{P}_z\big\{|B(\tau_k^Q + r^2) - x| > 2r \,\big|\, \tau_k^Q < \tau_R^*\big\}. \end{aligned}$$

The second factor can be bounded from below by a positive constant, which does not depend on r and R. The first factor is bounded from below by the probability that planar Brownian motion started at any point in $\partial\mathcal{B}(0, 2r)$ hits $\partial\mathcal{B}(0, 2R)$ before $\partial\mathcal{B}(0, r)$. This probability is given by

$$\frac{\log 2r - \log r}{\log 2R - \log r} \geqslant \frac{1}{\log_2 R + 2 + m}.$$

This is at least c/m for some $c > 0$ which depends on R only. ■

Denote by $\mathfrak{C}_m$ the set of dyadic cubes of side length 2^{-m} inside $\mathsf{Cube} = [-\frac{1}{2}, \frac{1}{2}]^d$.

Lemma 5.3. *There exists a random variable $C = C(\omega)$ such that, almost surely, for all m and for all cubes $Q \in \mathfrak{C}_m$ we have $\tau^Q_{\lceil m^2 C+1\rceil} > \tau^*_R$.*

Proof. From (9) we get that

$$\sum_{m=1}^{\infty} \sum_{Q\in\mathfrak{C}_m} \mathbb{P}\{\tau^Q_{\lceil cm^2+1\rceil} < \tau^*_R\} \leqslant \sum_{m=1}^{\infty} 2^{dm}\theta^{cm}.$$

Now choose c so large that $2^d\theta^c < 1$. Then, by the Borel–Cantelli lemma, for all but finitely many m we have $\tau^Q_{\lceil cm+1\rceil} > \tau^*_R$ for all $Q \in \mathfrak{C}_m$. Finally, we can choose a random $C(\omega) > c$ to handle the finitely many exceptional cubes. ■

The idea is to verify $\dim T(x) = 0$ for all paths satisfying Lemma 5.3 using completely deterministic reasoning. As this set of paths has full measure, this verifies the statement.

Fix a path $\{B(t)\colon t \geqslant 0\}$ satisfying Lemma 5.3 for a constant $C > 0$. Fix m and let $Q \in \mathfrak{C}_m$ be the cube containing a given $x \in \mathsf{Cube}$. Lemma 5.3 yields a covering of $T(x) \cap [0, \tau^*_R]$, which uses at most $Cm^2 + 1$ intervals of length 2^{-2m}. For any $\gamma > 0$ we hence find

$$\mathcal{H}^\gamma_{2^{-2m}}(T(x) \cap [0, \tau^*_R]) \leqslant (Cm^2 + 1)\,(2^{-2m})^\gamma.$$

Letting $m \uparrow \infty$ gives $\mathcal{H}^\gamma(T(x) \cap [0, \tau^*_R]) = 0$. Thus $T(x) \cap [0, \tau^*_R]$ and hence $T(x)$ itself has Hausdorff dimension zero. ■

This result leaves us with the informal question how often does a planar Brownian motion visit its most visited site. There are several ways to formalise this and I close this lecture series with a brief survey on recent progress on this question.

Large occupation times: Dembo, Peres, Rosen, Zeitouni

An idea of Dembo, Peres, Rosen and Zeitouni (2001) is to study the most visited small balls of planar Brownian motion. Let

$$\mu(A) = \int_0^T 1\{B_s \in A\}\, ds$$

be the total occupation time of Brownian motion up to the first exit time T of the disc of radius one. Then

$$\lim_{\varepsilon\downarrow 0} \sup_{x\in\mathbb{R}^2} \frac{\mu(\mathcal{B}(x,\varepsilon))}{\varepsilon^2 \log(1/\varepsilon)} = \sup_{x\in\mathbb{R}^2} \limsup_{\varepsilon\downarrow 0} \frac{\mu(\mathcal{B}(x,\varepsilon))}{\varepsilon^2 \log(1/\varepsilon)} = 2 \qquad \text{almost surely.}$$

Loosely, speaking the most visited small ball is visited for $2\varepsilon^2 \log(1/\varepsilon)$ time units, where ε is its radius.

This result is one the one hand not a satisfactory answer to our question, as the points x where we are close to the supremum above need not be points of infinite multiplicity. On the other hand it does answer the analogous question for simple random walks on the planar lattice. Denoting by $T_n(x)$ the number of visits of the random walk to site x during the first n steps, we have

$$\lim_{n\to\infty} \max_{x\in\mathbb{Z}^2} \frac{T_n(x)}{(\log n)^2} = \frac{1}{\pi} \qquad \text{almost surely.}$$

This can be derived from the Brownian motion result by strong approximation.

Large number of excursions: Bass, Burdzy, Khoshnevisan

An idea much closer to the original question is due to Bass, Burdzy, and Khoshnevisan (1994). Again we stop planar Brownian motion at first exit from the unit disc. Given any point x we denote by N_ε^x the number of excursions of the Brownian motion from x which travel at least distance ε from x. For almost every point x we have $N_\varepsilon^x = 0$, but a point x has infinite multiplicity if and only if

$$\lim_{\varepsilon\downarrow 0} N_\varepsilon^x = \infty.$$

The idea of Bass et al. is to quantify the number of visits to x by the rate of increase of N_ε^x as $\varepsilon \downarrow 0$. They show that, for every $0 < a < \frac{1}{2}$, almost surely there exist points $x \in \mathbb{R}^2$ with

$$\lim_{\varepsilon\downarrow 0} \frac{N_\varepsilon^x}{\log(1/\varepsilon)} = a.$$

On the other hand it is not hard to show (using a variant of the proof of Theorem 5.1) that the set of points with

$$\liminf_{\varepsilon\downarrow 0} \frac{N_\varepsilon^x}{\log(1/\varepsilon)} > 2e$$

is almost surely empty. The question precisely for which values of a there exists a point in the plane with $\lim_{\varepsilon\downarrow 0} \frac{N_\varepsilon^x}{\log(1/\varepsilon)} = a$ is still open.

Large set of visits: Cammarota, Mörters

Another idea is to measure the size of the sets $T(x)$ using a finer notion of dimension. Instead of evaluating coverings using the α-value

$$\sum_{i=1}^{\infty} |E_i|^\alpha$$

we look at a general nondecreasing (gauge) function $\varphi\colon [0,\varepsilon) \to [0,\infty)$ with $\varphi(0) = 0$ and define the φ-value

$$\sum_{i=1}^{\infty} \varphi(|E_i|).$$

The φ-Hausdorff measure is then defined by

$$\mathcal{H}^{\varphi}(E) = \lim_{\delta\downarrow 0} \inf \Big\{ \sum_{i=1}^{\infty} \varphi(|E_i|) \colon E_1, E_2, E_3, \ldots \text{ cover } E, \text{ and } |E_i| \leqslant \delta \Big\} .$$

This coincides with the α-Hausdorff measure for $\varphi(r) = r^{\alpha}$. Many of the measures occuring in probability theory (e.g. in the form of local times or occupation measures) turn out to be Hausdorff measures for special gauge functions restricted to particular sets.

Slowly varying functions can be used to measure the size of sets of dimension zero. The aim is to find a smallest gauge function φ (with repect to the partial ordering $\phi \ll \psi$ if $\lim_{\varepsilon\downarrow 0} \phi(\varepsilon)/\psi(\varepsilon) = 0$) for which almost surely exists a point x in the plane so that $\mathcal{H}^{\varphi}(T(x)) > 0$. In work currently in progress we show that almost surely there exists such a point for

$$\varphi(\varepsilon) = \frac{\log\log\log(1/\varepsilon)}{\log(1/\varepsilon)} .$$

We also show (and again this is not hard) that, almost surely, for any gauge function ϕ with

$$\lim_{\varepsilon\downarrow 0} \phi(\varepsilon) \log(1/\varepsilon) = 0$$

we have $\mathcal{H}^{\phi}(T(x)) = 0$ for all $x \in \mathbb{R}^2$. There is a gap between the two bounds which is subject of ongoing work.

6 Exercises

Exercise 1. Prove Lemma 1.6: If $A_1, A_2 \subset \mathbb{R}^2$ are Borel sets with positive area, then

$$\mathcal{L}_2\big(\{x \in \mathbb{R}^2 \colon \mathcal{L}_2(A_1 \cap (A_2 + x)) > 0\}\big) > 0.$$

Exercise 2. Let $0 < r < s < t$. Show that the probability that d-dimensional Brownian motion started at any point in $\partial\mathcal{B}(0,s)$ hits $\partial\mathcal{B}(0,t)$ before $\partial\mathcal{B}(0,r)$ is given by

- $\dfrac{\log s - \log r}{\log t - \log r}$ if $d = 2$,
- $\dfrac{s^{2-d} - r^{2-d}}{t^{2-d} - r^{2-d}}$ if $d > 2$.

Exercise 3. Prove the *Paley-Zygmund inequality*

$$\mathbb{P}\{X > 0\} \geqslant \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}.$$

Exercise 4. Prove *McKean's theorem*: Let $A \subset [0,\infty)$ be a closed subset and $\{B(t) \colon t \geqslant 0\}$ a d-dimensional Brownian motion. Then, almost surely,

$$\dim B(A) = 2 \dim A \wedge d.$$

Exercise 5. Complete the proof of Theorem 3.3 (b): For $d \leqslant 3$, the intersection of two independent Brownian paths in $\mathbb{R}^d$ is nontrivial with probability one.

Exercise 6. Find the Hausdorff dimension of a percolation limit set $\Gamma[\gamma] \subset \mathbb{R}^d$.

Exercise 7. Prove *Kaufman's theorem*: Let $\{B(t) \colon t \geqslant 0\}$ be Brownian motion in dimension $d \geqslant 2$. Almost surely, for any set $A \subset [0,\infty)$, we have

$$\dim B(A) = 2 \dim A.$$

Compare this result with McKean's theorem.